Contents lists available at ScienceDirect

American Journal of Preventive Cardiology
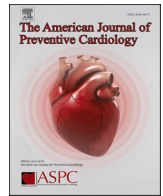
State-of-the-Art Review

Medicine 2032: The future of cardiovascular disease prevention with machine learning and digital health technology

Aamir Javaid [a], Fawzi Zghyer [a], Chang Kim [a], Erin M. Spaulding [a], Nino Isakadze [a], Jie Ding [a], Daniel Kargillis [a], Yumin Gao [a], Faisal Rahman [b], Donald E. Brown [c], Suchi Saria [d], Seth S. Martin [a], Christopher M. Kramer [e], Roger S. Blumenthal [a], Francoise A. Marvel [a,*]

[a] Johns Hopkins Ciccarone Center for the Prevention of Cardiovascular Disease, Division of Cardiology, Department of Medicine, Johns Hopkins University School of Medicine, 600 N. Wolfe St, Carnegie 591, Baltimore, MD 21287, USA
[b] Division of Cardiology, Department of Medicine, Baylor College of Medicine, Houston, TX, USA
[c] School of Data Science, University of Virginia, Charlottesville, VA, USA
[d] Machine Learning and Healthcare Laboratory, Departments of Computer Science, Statistics, and Health Policy, Malone Center for Engineering in Healthcare, and Armstrong Institute for Patient Safety and Quality, Johns Hopkins University, Baltimore, MD, USA
[e] Cardiovascular Division, Department of Medicine, University of Virginia Health, Charlottesville, VA, USA

ARTICLE INFO

ABSTRACT

Machine learning (ML) refers to computational algorithms that iteratively improve their ability to recognize patterns in data. The digitization of our healthcare infrastructure is generating an abundance of data from electronic health records, imaging, wearables, and sensors that can be analyzed by ML algorithms to generate personalized risk assessments and promote guideline-directed medical management. ML's strength in generating insights from complex medical data to guide clinical decisions must be balanced with the potential to adversely affect patient privacy, safety, health equity, and clinical interpretability. This review provides a primer on key advances in ML for cardiovascular disease prevention and how they may impact clinical practice.

1. Introduction

Artificial intelligence (AI) refers to computational systems that are able to perform tasks that normally require human intelligence and decision-making [1]. While traditional computer algorithms solve tasks by following instructions hard-coded by humans, ML is a subset of AI that enables computer algorithms to improve themselves based on experience [2]. Deep learning is a specific type of ML that can formulate complex feature representations by linking simple mathematical functions together in an interconnected network, often called a neural network [3]. Neural networks were originally inspired by the biology of the human brain in which interconnected neurons send and receive signals. Neurons or nodes within artificial neural networks contain non-linear mathematical functions to control which signals are sent to subsequent layers of the model, which are adjusted during training to optimize for an outcome of interest (Fig. 1) [3]. Deep learning uses neural networks with many layers of artificial neurons that can learn highly non-linear patterns from data [4].

One way to further categorize ML is supervised and unsupervised learning [2,5]. In supervised learning, a model is trained using human-labeled input-output pairs to determine the outputs for unpaired inputs [6]. These algorithms need large datasets to produce meaningful predictions, which often require significant human time for data labeling. Unsupervised learning, in contrast, refers to algorithms that identify data organizing patterns or data dimensionality reduction without focusing on outcome prediction [7]. These and hybrid approaches, such as semi-supervised and self-supervised methods [8], can reduce the time-intensive data labeling phase and elucidate hidden structures within datasets, which may be especially useful in complicated datasets with missing labels such as electronic health records (EHR).

Cardiology is a data-driven and evidence-based field that has been an early adopter of ML [9–11]. Numerous studies have demonstrated improved performance of ML compared to traditional risk assessments using well established cardiovascular disease (CVD) risk factors [12,13].

* Corresponding author.
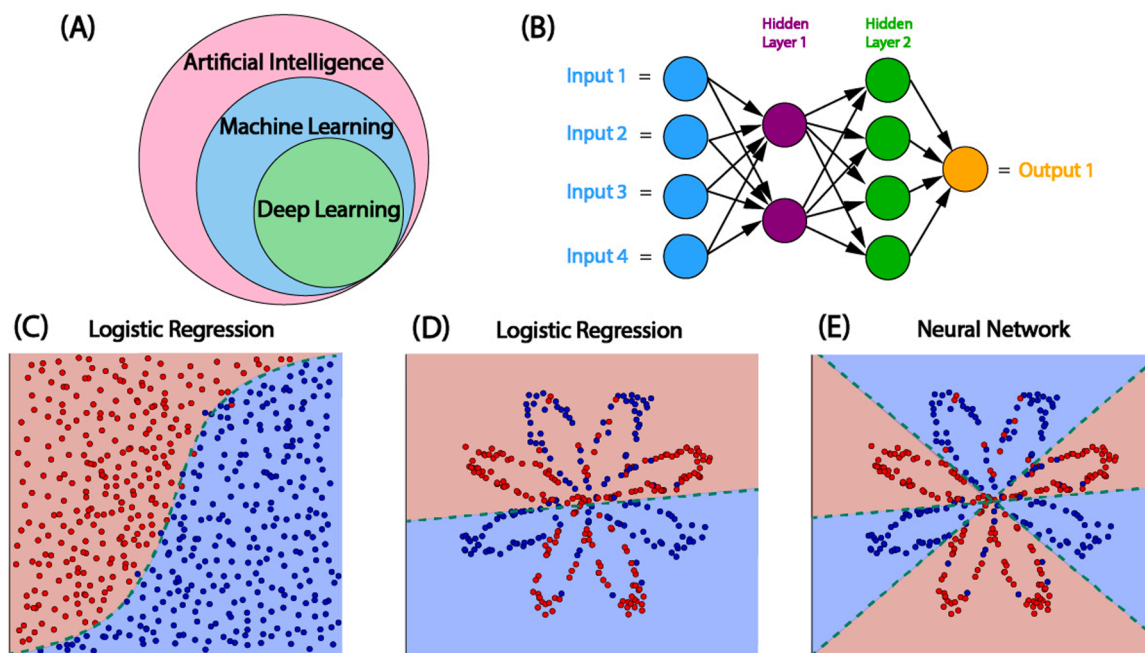 *E-mail address:* fmarvel1@jhmi.edu (F.A. Marvel).

Fig. 1. Machine learning overview. (A) Deep learning is a subset of machine learning (ML), which is a subset of artificial intelligence. (B) Architecture of a ML convolutional neural network (CNN) with two hidden layers. >3 layers qualifies as deep learning. (C-E) Logistic regression vs. CNN. Traditional cardiovascular risk estimates use logistic regression models, which excel when data is linearly separable (C) but not as well in complex situations (D). ML can generate more complex decision boundaries (E) [112].

Patient-specific information collected from the EHR, cardiovascular imaging, applications (apps), wearables, and sensors provide opportunities for generating longitudinal cardiovascular data for early disease detection and management. Additionally, social determinants of health, which are among the most significant factors for reducing death and disability from CVD [14], can be analyzed by ML algorithms. We believe

these innovations will drive a shift towards precision CVD prevention led by insights derived from big data, ML, and digital health technologies. In this review, we describe advances in ML as they apply to CVD prevention.

Table 1
Selected studies using machine learning for cardiovascular risk prediction.

Study	Dataset	Model Inputs	Outcomes	Top predictors	External Validation
Alaa et al. [13]	423,604 UK Biobank participants	104 lab variables, 369 clinical variables	CVD events over 5 years	Age, smoking, walking pace	No
Kakadiaris et al. [12]	6459 MESA participants	8 ACC/AHA risk calculator inputs	"Hard" and All CVD events over 13 years	N/A	FLEMENGHO (1348 White participants)
Sánchez-Cabo et al. [113]	4184 PESA participants	115 variables including demographics, systolic blood pressure, blood/urine tests, diet	Subclinical atherosclerosis over 3 years	Age, Hba1c, total cholesterol to HDL ratio, leukocyte volume, and hemoglobin	AWHS (1240 participants)
Ambale-Venkatesh et al. [15]	6814 MESA participants	735 variables from imaging and noninvasive tests, questionnaires, biomarker panels	CHD and all CVD over 12 years	CAC, TNF-α, Cardiac Troponin-T, N-tp-BNP	No
Weng et al. [16]	378,256 patients from UK	8 ACC/AHA risk calculator inputs and additional 22 variables (labs, past medical history, medications)	All CVD over 10 years	Age, sex, race, smoking	No
Kennedy et al. [114]	114,000 US Veterans Health Administration patients	Past medical history, medications, labs, vital signs	CVA and CVD death over 5 years	N/A	No
Dogan et al. [85]	2295 FHS participants	4 genetic and epigenetic loci	CHD in 5 years	N/A	No

All studies involved asymptomatic adults free of CVD.
FLEMENGHO - the Flemish Study of Environment, Genes and Health Outcomes.
MESA - Multiethnic Study of Atherosclerosis.
PESA - Progression of Early Subclinical Atherosclerosis.
AWHS - Aragon Workers' Health Study.
FHS – Framingham Heart Study.
ACC – American College of Cardiology.
AHA – American Heart Association.
CVD - cardiovascular disease.
CHD - coronary heart disease.
CVA - cerebrovascular accident.

2. Machine learning enhanced cardiovascular risk assessments

Current evidence suggests ML outperforms traditional CVD risk prediction tools [12,13,15,16]. In the Multi-Ethnic Study of Atherosclerosis (MESA), Kakadiaris et al. developed an ML model with the same inputs as the American College of Cardiology (ACC)/American Heart Association (AHA) atherosclerotic CVD risk estimator [12]. The ML model had an Area Under the Receiver Operating Characteristic curve (AUROC) of 0.94 vs. 0.72 for the traditional calculator in predicting CVD events. Additionally, the ACC/AHA estimator recommended statin therapy for 46% of the study population, despite 24% of hard CVD events occurring in patients not recommended statin treatment. In contrast, the ML model recommended statin therapy to 11% of the cohort and only 14% of hard CVD events occurred in participants who were not recommended statin therapy. The ML model would have simultaneously prescribed fewer statin medications and missed fewer events than the ACC/AHA estimator, potentially leading to more targeted therapy and enhanced CVD risk prediction [12]. Of note, the validation cohort for this model consisted of only non-Hispanic White participants due to low event rates, underscoring the need for validation in more diverse populations [12]. Still, in comparison to traditional risk estimators with fixed effect size estimates and assumed linear relationships between variables, the strength of ML is its ability to detect complex, nonlinear relationships and to iteratively improve its ability to model them with more data allowing for more precise risk estimates and fewer false alerts [17].

ML can integrate clinical data beyond traditional risk factors to incrementally improve predictive performance. In a prospective study of 423,604 UK Biobank patients without baseline CVD, investigators created an algorithm which automatically selects and optimizes ensembles of ML pipelines for predicting clinical prognosis. Using the same inputs as the Framingham score, the ML model had a higher AUROC (0.744 ± 0.005) than the Framingham Score (0.724 ± 0.004) and a Cox proportional hazards model (0.734 ± 0.005) [13]. The addition of 104 lab variables and 369 other clinical variables improved the AUROC of the ML model to 0.774 ± 0.005, in comparison to a Cox model with the same variables (0.758 ± 0.005) [13]. Using the all-variable model, the top three predictors were age, smoking, and walking pace, which were better predictors than traditional cardiovascular risk factors such as hypertension and dyslipidemia [13]. Limitations of the study include lack of cholesterol laboratory values (the BMI version of the Framingham Score was used) and lack of racial/ethnic data.

Another study in MESA explored risk estimation using a variety of ML model inputs including traditional risk factors and demographics, anthropometry, medication use, electrocardiogram (ECG), serum biomarkers, and cardiac imaging. The ML model performed better than established risk scores with increased prediction accuracy (decreased Brier score by 10–25%). Coronary artery calcium (CAC) was the strongest predictor of coronary heart disease and all CVD, while laboratory values such as tissue necrosis factor-α and N-terminal pro-Brain Natriuretic Peptide were among the top 5 predictors. The only traditional CVD risk factors in the top 20 predictors were age and smoking [15]. A summary of studies using ML to estimate cardiovascular risk is shown in Table 1.

These results suggest that the addition of clinical data stored in the EHR may improve risk prediction beyond traditional risk factors. This is particularly likely in those cases where the model is developed specific to a healthcare system. To evaluate the generalizability of these models and their impact on clinical outcomes, external validation and clinical trials comparing ML to the standard of care are needed.

3. Machine learning enhanced cardiovascular imaging

Imaging data such as coronary computed tomographic angiography (cCTA) [18], CAC [15], and epicardial adipose tissue [19] have improved ML predictive performance and are often among the strongest correlated variables with CVD outcomes such as myocardial infarction and cardiac death [19]. While early studies included these inputs as human-calculated values, recent models have automated this process by implementing ML computer vision techniques on the raw images. A Stanford-based team published two ML models that automated CAC scoring for dedicated gated coronary CT scans and non-gated chest CTs. The gated-chest CT model showed near perfect agreement with conventional manual scoring and reduced average scoring time with ML taking 3.5 s compared to manual scoring which required 261 s. Using CAC scores from paired gated-CT scans as the gold standard, a second model was trained to score CAC on non-gated chest CTs on an internal dataset and was externally validated in MESA. This model reported high sensitivity (80–100%) and positive predictive value (87–100%)[20] and is currently utilized clinically at a few US-based medical centers [21].

The application of this technology to CAC scoring, a leading diagnostic tool, is a powerful example of leveraging ML to advance preventive cardiology. ML can facilitate increased efficiency and access to CAC testing, particularly considering the information gain from non-gated CTs performed for other indications that also inform clinicians on a CAC score to improve CVD risk stratification. A future state of CVD preventive screening may automatically calculate CAC score, aortic valve calcification, epicardial adipose tissue, and other "radiomic" features to predict CVD risk from routine chest CTs [22,23]. Despite the promise of ML models applied to routine testing, the challenges to consider include incidental findings and nonessential downstream testing. ML-derived imaging results need to be carefully evaluated and clinical oversight is required to guide recommendations for appropriate clinical decision-making. In general, the higher the clinical acuity, the more extensive the validation process of the ML algorithm will need to be, given the shorter amount of time available for human review. In this sense, image analysis in preventive cardiology may become the first field for ML integration into clinical cardiology practice.

4. Machine learning enhanced cardiovascular wearables and sensors

Most CVD prevention guidelines recommend the use of 10-year absolute CVD risk estimates to guide therapeutic recommendations [24]. These assessments are largely driven by the age and sex of a patient; hence, most young patients are at low risk despite the majority going on to develop CVD in their lifetimes [25,26]. Since many risk factors and causes of CVD are well established and often begin decades before they manifest as an event, proposed alternative approaches are to use 30-year or lifetime risk assessments [27] and thereby prevent events by initiating therapy as soon as the causes of disease are identified [28]. These approaches are based on primary prevention, but are limited by the nature of the data used to estimate risk. As CVD risk estimations advance in accuracy, future assessments will likely include personalized cardiovascular health data collected from consumer-grade wearable devices.

Smart wearables are consumer-grade electronics that can be worn on the body as accessories such as watches, rings, and wristbands or embedded in clothing [29]. Data streams generated by these sensors may provide new insights into early detection of CVD, discovery of new CVD risk markers, and help facilitate primordial prevention. An estimated 20% of US residents own smart wearables and the global market is expected to undergo a compound annual growth rate of 25% through 2025 [30,31]. Additionally, novel loaner device programs have emerged to improve equitable access to technology. The iShare program, for example, provided a loaner Apple Watch, blood pressure monitor, and Apple iPhone with data access to patients who did not own smart technology to help them participate in a CVD health promotion program [32]. These efforts may be a first step towards long-term comprehensive health technology access and literacy.

In the future, smart wearable technology will allow clinicians to monitor trends in cardiovascular health metrics and develop actionable insights informed by an ML model at an individual- and population-level

by collecting biometrics on heart rate and rhythm, physical activity, blood pressure, sleep, glucose, and more (Central Illustration). The volume, variety, and velocity of this data is suited for analysis by ML algorithms, which excel in analyzing large datasets to detect trends that enhance predictive power. While these parameters are associated with traditional cardiovascular risk factors, they may also help uncover new CVD predictive markers. In the next section, we cover biologic data that can be measured by smart wearables with regards to utility for CVD prevention.

4.1. Heart rate and rhythm

Heart rate is measured by calculating the time interval between beats detected using photoplethysomography (PPG) or ECG [33]. For PPG, an emitter on the back of the device sends a continuous pulse of light through the skin and a photodetector detects the variation in the intensity of reflected photons because of microvascular changes in blood volume [33]. This can be used to detect subtle changes in heart rate trends over time, such as resting heart rate, heart rate variability, and recovery after exercise, all of which are associated with cardiovascular risk. Remote photo PPG, which can estimate pulse from microvascular changes detected on video, is a contactless approach which may see more widespread implementation in the coming decade [34]. Specific smartwatches and handheld devices also have single-lead ECG recording capability [35,36]. Based on these capabilities, ML enhanced by data from smart wearables may detect heart rate or ECG changes predictive of acute illness. Examples of ML-enhanced prediction include infections [37], arrhythmias [38], cardiac arrest [39], and thromboembolic events [40]. Heart rate metrics also have long-term prognostic utility for CVD, although these are not included in traditional risk assessments. For example, a high resting heart rate in populations free of CVD has been independently associated with increased risk of coronary artery disease and stroke [41]. Among 91 patients who were monitored using wrist-wearables in the 30 day post-discharge period after myocardial infarction, patients who were readmitted had elevated heart rates manifesting 3 days prior to readmission [42]. Low heart rate variability is independently associated with increased likelihood for myocardial ischemia, and the addition of heart rate variability to traditional CVD risk factors significantly improves the pretest probability for myocardial ischemia [43]. Conversely, higher heart rate variability during recovery from exercise has been associated with increased risk of cardiovascular mortality and all-cause mortality [44]. While most of these data are from single time checks during study visits, near-continuous PPG recording may present opportunities to further risk stratify and refine predictive metrics using ML, such as by monitoring changes in trends over time or capturing paroxysmal events.

In addition to prediction of acute disease states, ML models using ECG data have also been used to estimate long-term cardiac mortality. In an analysis of 7067 participants in the Third National Health and Nutrition Examination Survey (NHANES III), Kim et al. assessed whether ML algorithms trained on demographic, clinical, and ECG data could predict long-term cardiac mortality. Multiple ML models only required demographic and ECG data to achieve comparable performance to the pooled cohort equation and additional clinical data did not significantly improve predictive performance, suggesting that the 12-lead ECG may contain most of the prognostic information that can be derived from traditional cardiovascular risk factors. Interestingly, important ECG features were clustered in inferior and lateral leads in this study [45]. In another study of 1558,415 patients where ML analysis of ECG was used to predict an individual's age, those with ML-predicted age more than 8 years of chronological age had a 1.79 higher hazard of mortality [46]. The prognostic abilities of the single-lead smartwatch ECG for CVD remains to be evaluated. One study showed it is possible to detect left ventricular systolic dysfunction using asynchronous two-lead smartwatch ECGs [47]. Availability of consumer facing six- and 12-lead ECG devices may further improve the scalability of such algorithms.

4.2. Physical activity

Physical activity is one of the most powerful interventions to reduce risk of CVD [48,49] and all-cause mortality [50]. ML risk prediction models that included metrics of physical activity, such as walking pace, have found them to be among the strongest predictors of CVD events, even above traditional risk factors like hypertension and diabetes [13]. Most early studies have assessed these metrics using retrospective questionnaires at clinic visits. A potentially more objective and reproducible assessment of physical fitness could be cardiopulmonary exercise testing (CPET) which is minimally utilized in CVD prevention [51], although recent work has shown the utility of ML for improving the interpretability of CPET results for clinical use [52].

In contrast, newer smart technologies cannot only measure physical activity in real-time but also document trends over time. Most wearables track physical activity using a three-axis accelerometer to measure linear acceleration and a gyroscope to measure angular motion. GPS and barometers are also included in many devices to track position and record distance covered during a workout. Energy expenditure calculations and their accuracies differ based on the activity via a combination of the accelerometer, GPS, and heart rate, as determined by PPG [53]. Wrist devices, for example, are more accurate at calculating energy expenditure during activities such as running and walking, but less so during higher intensity exercise [54] or for activities with less arm motion, such as cycling and resistance training [55]. Some wearables estimate VO2max based on heart rate during exercise vs. at rest [56]. Wearables also provide an opportunity to guide activity through smart device reminders, recovery scores based on biometric parameters, and real-time biometric feedback during exercise.

Physical activity type, duration, and intensity may be particularly important for cardiac health [57], but it is unclear which wearable metrics are important for risk estimation. The added predictive value of device-measured activity metrics compared to previous methods and optimal monitoring timeframe are not established. While some wearables use trends in physical activity over days to weeks, in addition to patterns in other metrics like heart rate and sleep, to calculate a recovery index score suggesting readiness to take on physical activity or stress, none of these devices are FDA-approved and should be interpreted with caution [58,59]. Currently, there is lack of consensus on specific validation criteria for device performance and most medical societies have not established standards for clinical use of digital health technologies. Clinical trials are necessary to determine the appropriate inputs for risk or recovery scores and to aid the development of clinical practice guidelines in this area.

4.3. Blood pressure

Hypertension has been recognized as a CVD risk factor for over half a century [60] and is a leading cause of global morbidity and mortality. Emerging methods of digital home-based blood pressure measurements may improve monitoring frequency, including nocturnally and during exercise which have been associated with poorer outcomes [61]. Studies have found low reliability of wrist blood pressure cuffs compared to upper-arm readings [62], but there is increasing interest in cuff-less methods. These use a variety of methods such as pulse transit time, pulse wave analysis, facial video processing, ultrasound, and volume control, some of which employ ML methods [63–65]. Unfortunately, variable accuracy between devices and lack of validation has precluded most societies from endorsing their clinical use [64,66]. Cuffless devices often need individual user calibration on a weekly to monthly basis and require different standards for validation than cuff devices. A recent randomized clinical trial among 2101 patients with hypertension that found no significant difference in blood pressure control or user satisfaction over a six month period using a home blood pressure cuff compared to the same cuff paired with a smartphone app [67]. Overall, further work is required to determine the clinical utility of

digitally-enhanced blood pressure monitoring.

4.4. Sleep

The American Academy of Sleep Medicine and the Sleep Research Society recommend ≥7 h of sleep per night to promote optimal health [68]. Both short sleep (<7 h) and long sleep (>9 h) have been associated with increased risk of CVD, but ideal sleep duration varies between individuals [69]. Consumer sleep technologies (CSTs) use computer-based systems to measure and improve parameters of sleep, such as sleep stages, duration, quality, and wake-time. Due to their low cost, CSTs may soon achieve widespread availability and improve precision sleep medicine. Compared to the gold standard of polysomnography, in which electrodes are placed on the scalp to measure brain activity, most CSTs are peripherally-located wearables which use an accelerometer to measure motion and PPG for heart rate [70]. Some wearable CSTs include additional components such as oximetry, chest motion, and peripheral arterial tone [71]. The agreement of peripheral CSTs with polysomnography is strongest for 2-stage categorization of sleep vs wakefulness, whereas agreement for 4-stage sleep categorization ranges from 50 to 75% agreement [70].

Newer contactless CSTs employ a variety of methods. The Google Nest Hub, for instance, uses a sliding radar to measure motion and a microphone to detect sources of sleep disturbance (e.g. coughs and snoring), which are continuously analyzed by an ML model. The model was trained using over 10,000 labeled polysomnography sessions from MESA and the Sleep Heart Health Study (SHHS) [72]. In addition to increased user comfort of non-contactless methods, this could detect disorders such as obstructive sleep apnea (OSA). OSA is highly prevalent in the general population and associated with many forms of CVD but an estimated 85% of patients with clinically significant OSA remain undiagnosed [73,74]. CSTs may also help improve sleep with minimal user input. Embedded sensors in mattresses track heart rate, breathing, and movement, which are analyzed by an ML algorithm that adjusts the mattress temperature to optimize sleep quality using heated or cooled water [75]. These examples demonstrate how ML may be integrated into daily life, making silent adjustments to improve human experience and health. However, many have not been validated in clinical trials or cleared for use by the FDA, which remains a major hurdle to clinical adoption.

4.5. Interstitial glucose

Diabetes mellitus is the 7th highest cause of death in the US. Based on NHANES III, an estimated 28 million (10%) U.S. adults had diagnosed diabetes, 10 million adults had undiagnosed diabetes, and 114 million adults (46%) were prediabetic [76]. ML may enhance our understanding of diabetes by improving diagnosis, identifying new disease subtypes, modeling complex relationships with other comorbidities for cardiovascular risk, monitoring treatment response, and providing personalized guidance for lifestyle behaviors or other treatments in real-time.

Digital health tools may help improve diagnosis for the large proportion of adults living with undiagnosed diabetes. For example, an ML model trained on smartphone-based PPG readings from 37,709 participants using the smartphone camera achieved a AUROC of 0.766 for detecting diabetes and 0.740 in a validation cohort of 7806 individuals, indicating a potential widely available non-invasive digital biomarker [77]. Furthermore, large strides in continuous glucose monitoring (CGM) have expanded tools available for diabetes management [78–80]. Newer wearable interstitial fluid sensors are even able to measure additional metabolites at the same time, such as lactate or alcohol [81]. As one example of ML-enhanced CGM, Bent et al. developed a dataset of 25,000 simultaneous interstitial glucose and smartwatch readings in addition to a 10-day food log from 16 patients with prediabetic A1c% [82]. An ML model was created that determines an individual's normal, high, and low blood glucose values based on one

standard deviation from their mean [82]. They then trained another ML model to predict the glucose category based on 69 non-glucose input variables. The top predictors were the time of reading, sugar intake in last 24 h, hemoglobin A1c, carbohydrate intake in previous 2 h, sex, and physical activity in last 24 h [82]. This study highlights two key concepts enabled by ML: personalized glucose ranges rather than using standardized cut-offs and accurate prediction of abnormal blood glucose levels based on demographics and biometric data captured by digital health tools. This suggests a way digital health technology could help improve personalized management of diabetes by potentially anticipating factors leading to abnormal glucose levels based on real-time behaviors and making appropriate recommendations to avoid them.

5. Machine learning-enhanced poly-omic risk assessments

Multi-omics, cardiometabolics, and polygenic risk scores (PRS) are among the most promising future areas for ML risk assessments. Among the many insights they give into the pathophysiology of disease, they also provide an avenue to quantitate the known role of family history in CVD and identify higher-risk individuals long before the earliest stages of atherogenesis [83]. Omics approaches are still in their infancy due to their novelty and the need for further validation before implementation into routine clinical care. Despite this, studies have found improved prediction of coronary heart disease by PRS compared to traditional cardiovascular risk factors, even in adults aged 70 or older [84]. ML has been used in a few PRS studies. In one study, an ensemble of random forest ML models using four blood-based genetic and four epigenetic loci had a sensitivity of 0.70 for predicting incident coronary heart disease within five years, as compared to the Framingham risk score and ACC/AHA risk estimators which had sensitivities of 0.20 and 0.38, respectively [85]. The addition of a single nucleotide polymorphism associated with coronary artery disease (ID3 rs11574) to traditional risk factors improved ML model performance (AUROC 0.84) compared to the Framingham score (AUROC 0.72). These early findings may suggest the benefit of incorporating omics data into future ML-based CVD risk assessments.

6. Road to clinical implementation: challenges for ML in CVD prevention

Several recent discussion papers have reviewed challenges for implementation of digital health and ML in healthcare in great depth [86,87]. In the next section we focus on major implementation challenges and potential solutions with respect to CVD prevention.

6.1. Data availability

Despite the abundance of data collected by healthcare systems, medicine has several unique challenges compared to other data-driven industries where ML has thrived. The Health Insurance Portability and Accountability Act (HIPAA) requires national standards to protect sensitive health information with rigorous, center-specific Institutional Review Boards (IRBs) to regulate how patient data is used. This importantly protects patient privacy but has unintentionally created fragmented data silos across the country. As such, most published healthcare ML models use locally-obtained datasets and lack external validation. The Tufts Predictive Analytics and Comparative Effectiveness Cardiovascular Prediction Model Registry estimates 58% of cardiovascular prediction models have never been externally validated [88]. Even fewer models have been evaluated prospectively. Out of 130 FDA-approved ML-based devices, only 4 were evaluated prospectively, none of which included the 54 Class III high-risk devices, and only 37 devices were evaluated at multiple sites [89]. Lack of a continuous method for rigorous re-evaluation over time can lead to overfitting, dataset shift, and poor performance when applied in the real world [90].

6.2. Model development

For ML to perform well in a clinical context, training data must be high volume, diverse, and reproducible. In many situations, cardiovascular datasets only record traditional risk factors as binary variables; thus, relevant features may not be recorded at all or lack the granularity and volume required by ML to elucidate relevant relationships [91]. Data streams from digital health technologies may be the perfect match for the volume and granularity needed for ML. Furthermore, their ubiquitous nature may help overcome the fragmented nature of the U.S. healthcare system. However, uncoordinated data assembly and sharing practices can result in massive time for curation by humans for building ML models.

A coordinated approach, with the goal to improve priority health needs, is required to record and store data on platforms that are universally accessible and with which all devices can interface. There are several methods under investigation to address these limitations, such as creation of large anonymized clinical consortiums for ML training, new technology that allows interfacing between different EHRs to facilitate model validation, open-source publication of computational processes and datasets, and more [87]. The federal National Artificial Intelligence Research Resource Task force (NAIRR) was established in 2021 to create a National Research Cloud that would facilitate broad access to data and computing resources to boost ML research. In addition to increasing ML research on a national level, this could also result in a broader range of research topics being investigated which may not have a clear immediate financial benefit but rather substantial long-term returns, as has been the case with basic science that is mostly funded at the federal level [92]. The National Institute of Health (NIH) is also requiring researchers to include a data-management plan in grant applications by 2023 to eventually make all data publicly available [93].

6.3. Diversity and health equity

Striving for health equity is a priority across healthcare. While ML may seem more objective than humans on the surface, biases present in data used for model training will be exponentially amplified in actual use. Recent examples from tech demonstrate the pitfalls of training ML models using data lacking representation of historically underrepresented groups [94–96]. To address this, we must ensure diversity in medical datasets used to train ML models. Digital health technology may provide opportunities to reach new populations that traditionally have been underrepresented in medical research cohorts and in whom current risk assessment tools underperform. Further, there are now opportunities for quantitatively measuring and correcting for biases in ML models, which is harder to do with fixed and premeditated rule-based calculators [90,97].

By leveraging the accessibility of mobile technology, a more streamlined, equitable healthcare system could be a reality; however, challenges remain, particularly in the era of the COVID-19 pandemic [98]. The Pew Research Center found that Black and Hispanic individuals have similar smartphone ownership to the general population [99]; however, reports have varied on whether their utilization of smartphones to access health information is more or less than White individuals [100,101]. In a survey of 926 respondents, most respondents were very concerned or somewhat concerned about AI's unintended consequences, including misdiagnosis (92%), privacy breaches (71%), less time with clinicians (70%), and higher health care costs (68%) [102]. A higher proportion of respondents who self-identified as being members of underrepresented racial/ethnic groups indicated high levels of concern about these issues compared to White respondents [102]. Incorporating stakeholder input from varied groups in the development of digital health systems may help improve user engagement [103]. Devices, especially wearables and sensors, must be optimized for all skin-colors and body types [104]. With PPG, for example, there are efforts to implement dynamic light intensity at the microprocessor level

which will allow individuals of darker skin tone to transmit the same average number of photons as individuals with light skin through the blood vessels. Another approach is implementing multi-wavelength techniques and opto-mechanics to measure the amount of melanin in the skin and adjust the algorithms accordingly [105]. Considerable efforts are being dedicated to address these priority needs, such as through projects funded by the AHA Strategically Focused Research Network on Health Technologies and Innovation [106].

6.4. Integration with clinical systems

Digital health technology and ML are potentially transformative tools, but interoperability is key to success in clinical care delivery. While ML can enhance screening and risk assessments, it must seamlessly integrate into clinical pathways that have been independently proven to mitigate those risks, rather than new workstreams for healthcare teams. Early clinical implementation success stories exemplify this, such as ML-augmented detection of diabetic retinopathy [107] or CAC [20]. In contrast, any ML approach that requires a new clinical pathway, rather than enhancing an existing one, will need rigorous scientific evaluation and testing prior to considering implementation [91]. Clinical trials that fail to show improved outcomes using digital health interventions compared to the standard of care should be expected as the true utility of various interventions are determined [67]. A negative trial with a particular device does not rule out the possibility that the same outcome may be improved with a different device or software. Amidst the present arms race of digital health interventions competing for footholds in the healthcare landscape, inter-device performance will also need to be validated.

The FDA has issued a statement on steps to advance digital health policies with a proposed regulatory framework for ML-based software as medical devices [88]. Development of a ML model and providing proof of concept studies would underpin a successful IRB submission to conduct clinical trials to prove the safety and efficacy of these systems in a clinical environment [88]. Evaluating performance of ML models at multiple clinical sites with diverse populations and encouraging prospective head-to-head clinical trials against the standard of care will help ensure model performance and safety [89]. Additionally, algorithms should integrate with clinical pathways in ways that optimize and enhance workflows rather than contributing to alarm or pop-up fatigue [87].

6.5. Building clinician and patient trust

Even if models are effective, clinician and public trust must also be earned. Giving individuals control over their own health data may be one way to increase participation and interoperability, which is promoted by laws such as the 21st Century Cures Act, but more needs to be done to address historic breaches of trust related to over-surveillance and cybersecurity failures [86]. Clinicians are trained to seek clear diagnostic labels, whereas ML may look at data as a continuum of phenotypes and arrive at conclusions that are not compatible with the traditional medical schema. While we have metrics to evaluate model performance, it may be difficult to interpret how a model arrived at its predictions even if they are highly accurate, leading to the moniker of ML as a black box [108]. A survey found that 31% of nonclinical respondents were very uncomfortable and 40.5% somewhat uncomfortable with receiving a diagnosis from a ML model that was accurate 90% of the time but incapable of explaining its rationale [102]. Considerable efforts are being dedicated to develop methods for interpretation and visualization of model features which may help gain the trust of clinicians [109]. We believe that ML will enhance rather than replace the clinician's ability to make data-driven decisions together with the patient. In order to achieve this, education to better understand ML and its limitations should be integrated with clinical training so clinicians can serve as effective communicators and patient advocates [110]. Finally,
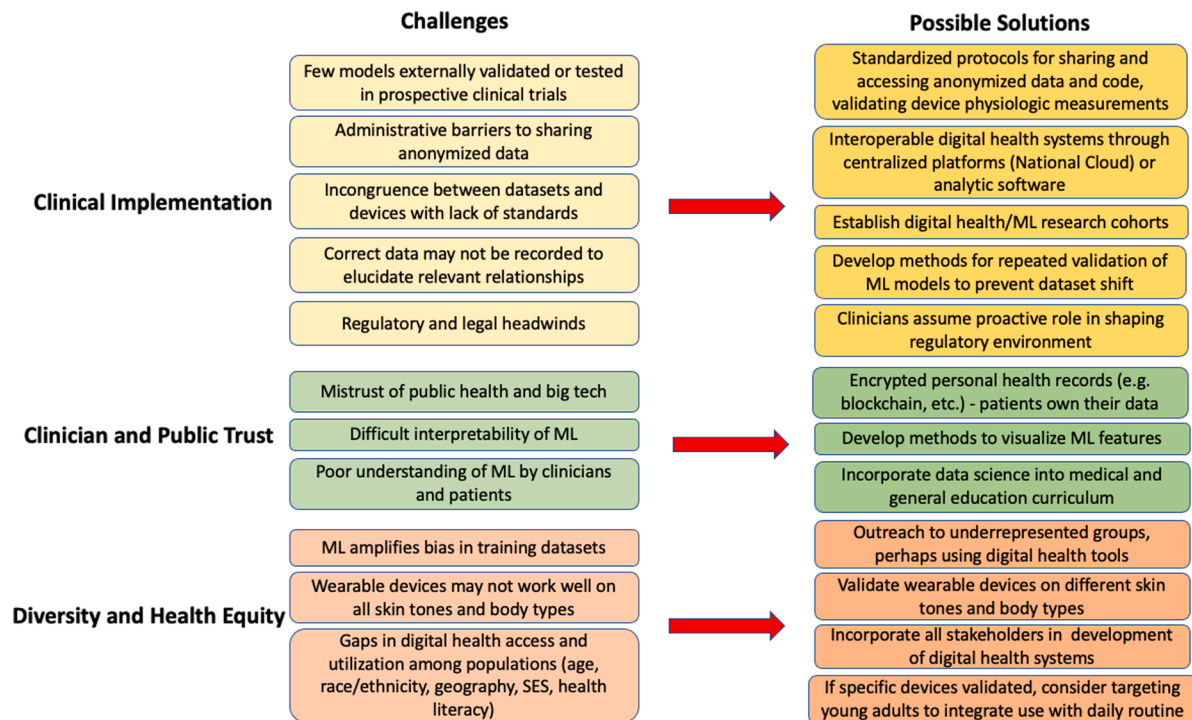
Fig. 2. Challenges and possible solutions for clinical implementation of machine learning (ML) in cardiovascular disease prevention. SES = socioeconomic status.

for clinicians to fully embrace ML algorithms into clinical practice, relevant regulatory and legal frameworks must be in place to delineate shared responsibilities and liabilities between the ML algorithm and the clinician. While the current legal framework incentivizes physicians to minimize the use of ML/AI to avoid liability [111], clinicians should take an active role in shaping the regulatory and legal environment as it is expected to evolve considerably in the coming years. Proposed challenges and solutions for building and implementing a successful ML model are outlined in Fig. 2.

7. Conclusions

ML has the potential to improve preventive care for CVD, as demonstrated through models that improve CVD risk prediction using traditional risk factors, clinical and laboratory values, imaging, wearable and sensor data, and omics. In coming years, we will likely see ML integrated within current healthcare systems and new systems built from the ground up based on these technologies. This gives us an opportunity to rethink our values and restructure healthcare in a way that improves patient outcomes, increases value of care, and is more equitable. At this early stage of ML in preventive care, it is incumbent upon us to set frameworks now that will guide us into the upcoming decades. We need to ensure ML models are trained and validated on diverse datasets, develop standardized practices for data sharing and rigorous re-evaluation of models, promote interoperability, and give individuals control of their own health data. Additionally, clinicians need to understand when ML can and cannot be applied and interpret model recommendations in the full clinical context of each patient. For these reasons, rather than replacing clinicians, ML will be an additional tool in the clinical armamentarium to enhance human-led decision-making and delivery of care. Collaboration between all stakeholders, including clinicians, researchers, patients, and industry partners, is essential to increasing trust in ML, digital health, and shaping the path forward for CVD prevention.

CRediT authorship contribution statement

Aamir Javaid: Conceptualization, Investigation, Writing – original draft, Writing – review & editing. **Fawzi Zghyer:** Writing – review & editing. **Chang Kim:** Software, Writing – review & editing. **Erin M. Spaulding:** Writing – review & editing. **Nino Isakadze:** Writing – review & editing. **Jie Ding:** Writing – review & editing. **Daniel Kargillis:** Methodology. **Yumin Gao:** Supervision. **Faisal Rahman:** Conceptualization, Writing – review & editing. **Donald E. Brown:** Data curation, Writing – review & editing. **Suchi Saria:** Data curation, Writing – review & editing. **Seth S. Martin:** Writing – review & editing. **Christopher M. Kramer:** Writing – review & editing. **Roger S. Blumenthal:** Writing – review & editing. **Francoise A. Marvel:** Investigation, Writing – review & editing, Writing – original draft.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Sung JJ, Stewart CL, Freedman B. Artificial intelligence in health care: preparing for the fifth Industrial Revolution. Med J Aust 2020;213:253–255.e251. https://doi.org/10.5694/mja2.50755.

[2] Scott IA, Cook D, Coiera EW, Richards B. Machine learning in clinical practice: prospects and pitfalls. Med J Aust 2019;211:203–205.e201. https://doi.org/10.5694/mja2.50294.

[3] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521:436–44. https://doi.org/10.1038/nature14539.

[4] Goodfellow I, Bengio Y, Courville A. Deep learning. MIT Press; 2016.

[5] Rashidi HH, Tran NK, Betts EV, Howell LP, Green R. Artificial intelligence and machine learning in pathology: the present landscape of supervised methods. Acad Pathol 2019;6. https://doi.org/10.1177/2374289519873088. 2374289519873088-2374289519873088.

[6] Rajkomar A, Dean J, Kohane I. Machine learning in medicine. N Engl J Med 2019; 380:1347–58. https://doi.org/10.1056/NEJMra1814259.

[7] Lundervold AS, Lundervold A. An overview of deep learning in medical imaging focusing on MRI. Z Med Phys 2019;29:102–27. https://doi.org/10.1016/j.zemedi.2018.11.002.

[8] Sarker IH. Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions. SN Comput Sci 2021;2. https://doi.org/10.1007/s42979-021-00815-1. 420-420.

[9] Attia ZI, et al. Screening for cardiac contractile dysfunction using an artificial intelligence-enabled electrocardiogram. Nat Med 2019;25:70–4. https://doi.org/10.1038/s41591-018-0240-2.

[10] Zhang J, et al. Fully automated echocardiogram interpretation in clinical practice. Circulation 2018;138:1623–35. https://doi.org/10.1161/circulationaha.118.034338.

[11] Madani A, Arnaout R, Mofrad M, Arnaout R. Fast and accurate view classification of echocardiograms using deep learning. npj Digit Med 2018;1. https://doi.org/10.1038/s41746-017-0013-1.

[12] Kakadiaris IA, et al. Machine learning outperforms ACC /AHA CVD risk calculator in MESA. J Am Heart Assoc 2018;7:e009476. https://doi.org/10.1161/jaha.118.009476.

[13] Alaa AM, Bolton T, Di Angelantonio E, 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. https://doi.org/10.1371/journal.pone.0213653.

[14] Havranek EP, et al. Social determinants of risk and outcomes for cardiovascular disease. Circulation 2015;132:873–98. https://doi.org/10.1161/CIR.0000000000000228.

[15] Ambale-Venkatesh B, et al. Cardiovascular event prediction by machine learning: the multi-ethnic study of atherosclerosis. Circ Res 2017;121:1092–101. https://doi.org/10.1161/CIRCRESAHA.117.311312.

[16] Weng SF, Reps J, Kai J, Garibaldi JM, Qureshi N. Can machine-learning improve cardiovascular risk prediction using routine clinical data? PLoS One 2017;12: e0174944. https://doi.org/10.1371/journal.pone.0174944.

[17] Saria S, Goldenberg A. Subtyping: what it is and its role in precision medicine. IEEE Intell Syst 2015;30:70–5.

[18] Motwani M, et al. Machine learning for prediction of all-cause mortality in patients with suspected coronary artery disease: a 5-year multicentre prospective registry analysis. Eur Heart J 2017;38:500–7. https://doi.org/10.1093/eurheartj/ehw188.

[19] 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. https://doi.org/10.1093/cvr/cvz321.

[20] Eng D, et al. Automated coronary calcium scoring using deep learning with multicenter external validation. npj Digit Med 2021;4:88. https://doi.org/10.1038/s41746-021-00460-1.

[21] Mongan, J. & Kallianos, K. (Center for intelligent imaging: UCSF radiology, 2021).

[22] Eslami P, et al. Radiomics of coronary artery calcium in the framingham heart study. Radiol Cardiothorac Imaging 2020;2:e190119. https://doi.org/10.1148/ryct.2020190119. e190119.

[23] Jávorszky N, et al. Deep learning–based atherosclerotic coronary plaque segmentation on coronary CT angiography. Eur Radiol 2022. https://doi.org/10.1007/s00330-022-08801-8.

[24] Arnett DK, et al. 2019 ACC/AHA guideline on the primary prevention of cardiovascular disease. J Am Coll Cardiol 2019;140:e596–646.

[25] Canto JG, et al. Number of coronary heart disease risk factors and mortality in patients with first myocardial infarction. JAMA 2011;306:2120–7. https://doi.org/10.1001/jama.2011.1654.

[26] Khot UN, et al. Prevalence of conventional risk factors in patients with coronary heart disease. JAMA 2003;290:898–904. https://doi.org/10.1001/jama.290.7.898.

[27] Thanassoulis G, Sniderman AD, Pencina MJ. A long-term benefit approach vs standard risk-based approaches for statin eligibility in primary prevention. JAMA Cardiol 2018;3:1090–5. https://doi.org/10.1001/jamacardio.2018.3476.

[28] Sniderman AD, et al. The causal exposure model of vascular disease. Clin Sci 2012;122:369–73. https://doi.org/10.1042/CS20110449 (Lond).

[29] Bayoumy K, et al. Smart wearable devices in cardiovascular care: where we are and how to move forward. Nat Rev Cardiol 2021. https://doi.org/10.1038/s41569-021-00522-7.

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

[31] Wearable Medical Devices Market Share, Size, Trends, Industry Analysis Report. (Polaris Market Research, 2020).

[32] Yang WE, et al. Strategies for the successful implementation of a novel iPhone Loaner System (iShare) in mHealth interventions: prospective study. JMIR Mhealth Uhealth 2019;7:e16391.

[33] Kamišalić A, Fister Jr I, Turkanović M, Karakatič S. Sensors and functionalities of non-invasive wrist-wearable devices: a review. Sensors 2018;18. https://doi.org/10.3390/s18061714 (Basel).

[34] Prakash SKA, Tucker CS. Bounded Kalman filter method for motion-robust, non-contact heart rate estimation. Biomed Opt Express 2018;9:873–97.

[35] Samol A, et al. Single-lead ECG recordings including Einthoven and Wilson leads by a smartwatch: a new era of patient directed early ECG differential diagnosis of cardiac diseases? Sensors 2019;19:4377. https://doi.org/10.3390/s19204377 (Basel).

[36] Kardia, < https://www.kardia.com/ >(2022).

[37] Mitratza M, et al. The performance of wearable sensors in the detection of SARS-CoV-2 infection: a systematic review. Lancet Digit Health 2022;4:e370–83. https://doi.org/10.1016/S2589-7500(22)00019-X.

[38] Perez MV, et al. Large-scale assessment of a smartwatch to identify atrial fibrillation. N Engl J Med 2019;381:1909–17.

[39] Narayan SM, Wang PJ, Daubert JP. New concepts in sudden cardiac arrest to address an intractable epidemic: JACC state-of-the-art review. J Am Coll Cardiol 2019;73:70–88. https://doi.org/10.1016/j.jacc.2018.09.083.

[40] Shade, J.K. et al. COVID-HEART: development and validation of a multi-variable model for real-time prediction of cardiovascular complications in hospitalized patients with COVID-19. *medRxiv*, 2021.2001.2003.21249182, doi:10.1101/2021.01.03.21249182 (2021).

[41] Zhang D, Wang W, Li F. Association between resting heart rate and coronary artery disease, stroke, sudden death and noncardiovascular diseases: a meta-analysis. CMAJ 2016;188:E384–92. https://doi.org/10.1503/cmaj.160050.

[42] Weng D, et al. Heart rate trajectories in patients recovering from acute myocardial infarction: a longitudinal analysis of apple watch heart rate recordings. Cardiovasc Digit Health J 2021;2:270–81. https://doi.org/10.1016/j.cvdhj.2021.05.003.

[43] Goldenberg I, et al. Heart rate variability for risk assessment of myocardial ischemia in patients without known coronary artery disease: the HRV-DETECT (Heart rate variability for the detection of myocardial ischemia) study. J Am Heart Assoc 2019;8:e014540. https://doi.org/10.1161/JAHA.119.014540. e014540.

[44] Singh N, et al. Heart rate variability: an old metric with new meaning in the era of using mHealth technologies for health and exercise training guidance. Part two: prognosis and training. Arrhythm Electrophysiol Rev 2018;7:247–55. https://doi.org/10.15420/aer.2018.30.2.

[45] Kim, C.H. et al. Machine learning to predict 10-year cardiovascular mortality from the electrocardiogram: analysis of the third national health and nutrition examination survey (NHANES III). *medRxiv*, 10.1101/2021.09.09.21263327 (2021).

[46] Lima EM, et al. Deep neural network-estimated electrocardiographic age as a mortality predictor. Nat Commun 2021;12:1–10.

[47] 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. https://doi.org/10.3390/diagnostics12030654 (Basel).

[48] Zhang X, Cash RE, Bower JK, Focht BC, Paskett ED. Physical activity and risk of cardiovascular disease by weight status among U.S adults. PLoS One 2020;15: e0232893. https://doi.org/10.1371/journal.pone.0232893.

[49] Georgousopoulou EN, et al. Physical activity level improves the predictive accuracy of cardiovascular disease risk score: the ATTICA study (2002-2012). Int J Prev Med 2016;7:52. https://doi.org/10.4103/2008-7802.178346.

[50] Blond K, Brinkløv CF, Ried-Larsen M, Crippa A, Grøntved A. Association of high amounts of physical activity with mortality risk: a systematic review and meta-analysis. Br J Sports Med 2020;54:1195–201. https://doi.org/10.1136/bjsports-2018-100393.

[51] Guazzi, M. et al. EACPR/AHA joint scientific statement. clinical recommendations for cardiopulmonary exercise testing data assessment in specific patient populations. (2012).

[52] Portella J, et al. Using machine learning to identify organ system specific limitations to exercise via cardiopulmonary exercise testing. IEEE J Biomed Health Inform 2022. https://doi.org/10.1109/jbhi.2022.3163402. **Pp**.

[53] Yang CC, Hsu YL. A review of accelerometry-based wearable motion detectors for physical activity monitoring. Sensors 2010;10:7772–88.

[54] Shcherbina A, et al. Accuracy in wrist-worn, sensor-based measurements of heart rate and energy expenditure in a diverse cohort. J Pers Med 2017;7:3.

[55] Boudreaux BD, et al. Validity of wearable activity monitors during cycling and resistance exercise. Med Sci Sports Exerc 2018;50:624–33. https://doi.org/10.1249/mss.0000000000001471.

[56] Passler S, Bohrer J, Blöchinger L, Senner V. Validity of wrist-worn activity trackers for estimating VO(2max) and energy expenditure. Int J Environ Res Public Health 2019;16. https://doi.org/10.3390/ijerph16173037.

[57] Arvidsson D, Fridolfsson J, Börjesson M. Measurement of physical activity in clinical practice using accelerometers. J Intern Med 2019;286:137–53. https://doi.org/10.1111/joim.12908.

[58] Whoop, < https://www.whoop.com/ >(2022).

[59] Oura Ring, < https://ouraring.com/ >(2022).

[60] Kannel WB, Dawber TR, Kagan A, Revotskie N, Stokes J. Factors of risk in the development of coronary heart disease—six-year follow-up experience. Ann Intern Med 1961;55:33–50. https://doi.org/10.7326/0003-4819-55-1-33.

[61] Whelton PK, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA guideline for the prevention, detection, evaluation, and management of high blood pressure in adults: executive summary: a report of the American College of Cardiology/American Heart Association task force on clinical practice guidelines. Hypertension 2018;71:1269–324. https://doi.org/10.1161/hyp.0000000000000066.

[62] Zweiker R, Schumacher M, Fruhwald FM, Watzinger N, Klein W. Comparison of wrist blood pressure measurement with conventional sphygmomanometry at a cardiology outpatient clinic. J Hypertens 2000;18:1013–8. https://doi.org/10.1097/00004872-200018080-00004.

[63] Bard DM, Joseph JI, van Helmond N. Cuff-less methods for blood pressure telemonitoring. Front Cardiovasc Med 2019;6:40. https://doi.org/10.3389/fcvm.2019.00040.

[64] Stergiou GS, et al. Cuffless blood pressure measuring devices: review and statement by the European Society of Hypertension Working Group on blood pressure monitoring and cardiovascular variability. J Hypertens 2022;40: 1449–60.

[65] Li YH, Harfiya LN, Purwandari K, Lin YD. Real-time cuffless continuous blood pressure estimation using deep learning model. Sensors 2020;20. https://doi.org/10.3390/s20195606 *(Basel)*.

[66] van Helmond N, Day W, Martin SS, Plante TB. Cuff-less, personal, ambulatory blood pressure devices and disruption of existing blood pressure measurement paradigms. Am J Hypertens 2020;33:813–5. https://doi.org/10.1093/ajh/hpaa090.

[67] Pletcher MJ, et al. Effectiveness of standard vs enhanced self-measurement of blood pressure paired with a connected smartphone application: a randomized clinical trial. JAMA Intern Med 2022. https://doi.org/10.1001/jamainternmed.2022.3355.

[68] Watson N, Badr M, Belenky G, Bliwise D, Buxton OM, Buysee D, Tasali E. Recommended amount of sleep for a healthy adult: a joint consensus statement of the American academy of sleep medicine and sleep research society. Sleep 2015;38:843–4.

[69] Yin J, et al. Relationship of sleep duration with all-cause mortality and cardiovascular events: a systematic review and dose-response meta-analysis of prospective cohort studies. J Am Heart Assoc 2017;6:e005947.

[70] Miller DJ, et al. A validation study of the WHOOP strap against polysomnography to assess sleep. J Sports Sci 2020;38:2631–6. https://doi.org/10.1080/02640414.2020.1797448.

[71] Tauman R, et al. Watch-PAT is useful in the diagnosis of sleep apnea in patients with atrial fibrillation. Nat Sci Sleep 2020;12:1115–21. https://doi.org/10.2147/NSS.S278752.

[72] Dixon, M. & Lee, R.S.*Enhanced sleep sensing in nest hub*, < https://ai.googleblog.com/2021/11/enhanced-sleep-sensing-in-nest-hub.html >(2021).

[73] Tietjens JR, et al. Obstructive sleep apnea in cardiovascular disease: a review of the literature and proposed multidisciplinary clinical management strategy. J Am Heart Assoc 2019;8:e010440. https://doi.org/10.1161/JAHA.118.010440.

[74] Motamedi KK, McClary AC, Amedee RG. Obstructive sleep apnea: a growing problem. Ochsner J 2009;9:149–53.

[75] Eight Sleep, https://www.eightsleep.com/ (2022).

[76] Tsao CW, et al. Heart disease and stroke statistics-2022 update: a report from the American Heart Association. Circulation 2022;145:e153–639. https://doi.org/10.1161/CIR.0000000000001052.

[77] Avram R, et al. A digital biomarker of diabetes from smartphone-based vascular signals. Nat Med 2020;26:1576–82. https://doi.org/10.1038/s41591-020-1010-5.

[78] Fokkert M, et al. Improved well-being and decreased disease burden after 1-year use of flash glucose monitoring (FLARE-NL4). BMJ Open Diabetes Res Care 2019;7:e000809.

[79] Bolinder J, Antuna R, Geelhoed-Duijvestijn P, Kröger J, Weitgasser R. Novel glucose-sensing technology and hypoglycaemia in type 1 diabetes: a multicentre, non-masked, randomised controlled trial. Lancet 2016;388:2254–63. https://doi.org/10.1016/s0140-6736(16)31535-5.

[80] Dunn TC, Xu Y, Hayter G, Ajjan RA. Real-world flash glucose monitoring patterns and associations between self-monitoring frequency and glycaemic measures: a European analysis of over 60 million glucose tests. Diabetes Res Clin Pract 2018;137:37–46. https://doi.org/10.1016/j.diabres.2017.12.015.

[81] Tehrani F, et al. An integrated wearable microneedle array for the continuous monitoring of multiple biomarkers in interstitial fluid. Nat Biomed Eng 2022. https://doi.org/10.1038/s41551-022-00887-1.

[82] Bent B, et al. Engineering digital biomarkers of interstitial glucose from noninvasive smartwatches. npj Digit Med 2021;4:89. https://doi.org/10.1038/s41746-021-00465-w.

[83] Muse ED, Chen SF, Torkamani A. Monogenic and polygenic models of coronary artery disease. Curr Cardiol Rep 2021;23:107. https://doi.org/10.1007/s11886-021-01540-0.

[84] Neumann JT, et al. Prognostic value of a polygenic risk score for coronary heart disease in individuals aged 70 years and older. Circ Genom Precis Med 2022;15:e003429. https://doi.org/10.1161/CIRCGEN.121.003429.

[85] Dogan MV, et al. Blood-based biomarkers for predicting the risk for five-year incident coronary heart disease in the framingham heart study via machine learning. Genes 2018;9:641. https://doi.org/10.3390/genes9120641 *(Basel)*.

[86] Abernethy A, et al. The promise of digital health: then, now, and the future. NAM Perspect 2022. https://doi.org/10.31478/202206e.

[87] Saria S. Not all AI is created equal: strategies for safe and effective adoption. NEJM Catal Innov Care Deliv 2022;3.

[88] Wessler BS, et al. Abstract 130: the Tufts PACE clinical predictive model registry: update 1990 through 2015. Circ Cardiovasc Qual Outcomes 2017;10. https://doi.org/10.1161/circoutcomes.10.suppl_3.130. A130-A130.

[89] Wu E, et al. How medical AI devices are evaluated: limitations and recommendations from an analysis of FDA approvals. Nat Med 2021;27:582–4. https://doi.org/10.1038/s41591-021-01312-x.

[90] Finlayson SG, et al. The clinician and dataset shift in artificial intelligence. N Engl J Med 2021;385:283.

[91] London AJ. Artificial intelligence in medicine: overcoming or recapitulating structural challenges to improving patient care? Cell Rep Med 2022. https://doi.org/10.1016/j.xcrm.2022.100622.

[92] Ho, D.E., King, J., Wald, R.C. & Wald, C.*Building a national AI research resource: a blueprint for the national research cloud*, < https://hai.stanford.edu/white-paper-building-national-ai-research-resource >(2021).

[93] Kozlov, M. (Nature Publishing Group, 2022).

[94] Simonite, T. Machines taught by photos learn sexist view of women. (2017). < https://www.wired.com/story/machines-taught-by-photos-learn-a-sexist-view-of-women/ >.

[95] Vincent, J. Google 'fixed' its racist algorithm by removing gorillas from its image-labeling tech. (2018). < https://www.theverge.com/2018/1/12/16882408/google-racist-gorillas-photo-recognition-algorithm-ai >.

[96] Lohr, S. in *The New York Times* (2018).

[97] Wang HE, et al. A bias evaluation checklist for predictive models and its pilot application for 30-day hospital readmission models. J Am Med Inform Assoc 2022. https://doi.org/10.1093/jamia/ocac065.

[98] Giansanti D. The artificial intelligence in digital pathology and digital radiology: where are we? Healthcare 2021;9:30.

[99] Mobile Fact Sheet, < https://www.pewresearch.org/internet/fact-sheet/mobile/ >(2020).

[100] Anderson, M. Racial and ethnic differences in how people use mobile technology. (2015).

[101] Mitchell UA, Chebli PG, Ruggiero L, Muramatsu N. The digital divide in health-related technology use: the significance of race/ethnicity. Gerontologist 2019;59:6–14.

[102] Khullar D, et al. Perspectives of patients about artificial intelligence in health care. JAMA Netw. Open 2022;5:e2210309. e2210309.

[103] Isakadze N, et al. Engaging diverse patients, caregivers, and clinicians as partners to identify disease management related challenges in atrial fibrillation: a qualitative study utilizing human-centered design methodology. Circulation 2021;144:A11258. A11258.

[104] Dagher L, Shi H, Zhao Y, Marrouche NF. Wearables in cardiology: here to stay. Heart Rhythm 2020;17:889–95. https://doi.org/10.1016/j.hrthm.2020.02.023.

[105] Colvonen PJ. Response to: investigating sources of inaccuracy in wearable optical heart rate sensors. npj Digit Med 2021;4:1–2.

[106] *More than $14 million in research grants awarded for health technology solutions focused on heart and brain health, including special projects related to COVID-19 and CVD*, <https://newsroom.heart.org/news/more-than-14-million-in-research-grants-awarded-for-health-technology-solutions-focused-on-heart-and-brain-health-including-special-projects-related-to-covid-19-and-cvd>(2020).

[107] Abràmoff MD, Lavin PT, Birch M, Shah N, Folk JC. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. npj Digit Med 2018;1:39. https://doi.org/10.1038/s41746-018-0040-6.

[108] McCoy LG, Brenna CT, Chen SS, Vold K, Das S. Believing in black boxes: machine learning for healthcare does not need explainability to be evidence-based. J Clin Epidemiol 2022;142:252–7.

[109] Montavon G, Samek W, Müller KR. Methods for interpreting and understanding deep neural networks. Digit Signal Process 2018;73:1–15. https://doi.org/10.1016/j.dsp.2017.10.011.

[110] Russak AJ, et al. Machine learning in cardiology-ensuring clinical impact lives up to the hype. J Cardiovasc Pharmacol Ther 2020;25:379–90. https://doi.org/10.1177/1074248420928651.

[111] Price WN, Gerke S, Cohen IG. Potential liability for physicians using artificial intelligence. JAMA 2019;322:1765–6.

[112] Ng, A. in *Deep learning specialization* Vol. 2022 (deeplearning.ai).

[113] Sánchez-Cabo F, et al. Machine learning improves cardiovascular risk definition for young, asymptomatic individuals. J Am Coll Cardiol 2020;76:1674–85. https://doi.org/10.1016/j.jacc.2020.08.017.

[114] Kennedy EH, Wiitala WL, Hayward RA, Sussman JB. Improved cardiovascular risk prediction using nonparametric regression and electronic health record data. Med Care 2013;51:251–8. https://doi.org/10.1097/MLR.0b013e31827da594.