OMICS A Journal of Integrative Biology
Volume 24, Number 5, 2020

DOI: 10.1089/omi.2019.0142

Implementing Artificial Intelligence and Digital Health in Resource-Limited Settings? Top 10 Lessons We Learned in Congenital Heart Defects and Cardiology

Nicholas Ekow Thomford,[1–3] Christian Domilongo Bope,[1–4] Francis Edem Agamah,[1,2] Kevin Dzobo,[2,5] Richmond Owusu Ateko,[6] Emile Chimusa,[1,2] Gaston Kuzamunu Mazandu,[1] Simon Badibanga Ntumba,[4] Collet Dandara,[1,2] and Ambroise Wonkam[1,2]

Abstract

Artificial intelligence (AI) is one of the key drivers of digital health. Digital health and AI applications in medicine and biology are emerging worldwide, not only in resource-rich but also resource-limited regions. AI predates to the mid-20th century, but the current wave of AI builds in part on machine learning (ML), big data, and algorithms that can learn from massive amounts of online user data from patients or healthy persons. There are lessons to be learned from AI applications in different medical specialties and across developed and resource-limited contexts. A case in point is congenital heart defects (CHDs) that continue to plague sub-Saharan Africa, which calls for innovative approaches to improve risk prediction and performance of the available diagnostics. Beyond CHDs, AI in cardiology is a promising context as well. The current suite of digital health applications in CHD and cardiology include complementary technologies such as neural networks, ML, natural language processing and deep learning, not to mention embedded digital sensors. Algorithms that build on these advances are beginning to complement traditional medical expertise while inviting us to redefine the concepts and definitions of expertise in molecular diagnostics and precision medicine. We examine and share here the lessons learned in current attempts to implement AI and digital health in CHD for precision risk prediction and diagnosis in resource-limited settings. These top 10 lessons on AI and digital health summarized in this expert review are relevant broadly beyond CHD in cardiology and medical innovations. As with AI itself that calls for systems approaches to data capture, analysis, and interpretation, both developed and developing countries can usefully learn from their respective experiences as digital health continues to evolve worldwide.

Keywords: artificial intelligence, digital health, machine learning, deep learning, congenital heart defects, eHealth

Introduction

At the time of The Great Depression in the United States, the American sociologist William F. Ogburn has proposed the theory of ‘‘cultural lags,’’ the idea that as technological changes leap forward, they create a cultural lag in society, which then need to adapt to the new realities introduced by innovations (Ogburn, 1922).

Yet, history has shown that the relationship between society and technological change is not a one-way interaction, and that society and individuals are not simply passive adapters of emerging technologies. That is, new technologies not only shape but are also shaped by society.

While scientific communities have long subscribed to the cultural lag theory of technology development, society and individuals can shape technology development through their

[1]Division of Human Genetics, Department of Pathology, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa.
[2]Institute for Infectious Disease and Molecular Medicine, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa.
[3]School of Medical Sciences, Department of Medical Biochemistry, University of Cape Coast, Cape Coast, Ghana.
[4]Department of Mathematics and Computer Sciences, Faculty of Sciences, University of Kinshasa, Kinshasa, D.R. Congo.
[5]Division of Medical Biochemistry, Department of Integrative Biomedical Sciences, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa.
[6]University of Ghana Medical School, Department of Chemical Pathology, University of Ghana, Accra, Ghana.

''agency'' or put in other words, humans' ability to shape their sociotechnical environment.

If we accept the view that there is a bidirectional relationship between technology and society, what are the prospects and challenges across the continuum of developed and developing countries as artificial intelligence (AI) and digital health are becoming trending research topics, not to mention applications in medicine, biology, and ecology?

The Rise of AI and Digital Health

AI and digital health are changing health care delivery, impacting, for example, medical diagnoses, reducing duplication in testing, and empowering patients in their health care decisions. The rise of AI and digital health has now established a proof of concept in certain medical fields such as radiology, psychiatry, pathology, and ophthalmology with their application overwhelmingly transforming and revolutionizing medical practice.

Cardiovascular medicine has a vast impact on health care as heart and vascular disease greatly impact global health. On the contrary, the use of AI in the field of cardiology is an emerging practice that still requires further development to actualize its full application potential. AI, digital health, and their applications can complement clinical practice by using efficient predictive models and algorithms developed through initiatives of bioinformatics to detect and analyze large-scale heterogeneous datasets, including but not limited to cardiovascular imaging, environmental, and clinical big data to leverage disease-risk prediction and diagnosis to foster personalized medicine.

In a context of cardiology, AI and digital health intersection, congenital heart defects (CHDs) are the most common congenital anomalies observed worldwide, and remain the leading cause of infant morbidity and mortality (Gilboa et al., 2010; Sun et al., 2017; Yeh et al., 2013). The estimated global incidence rate of CHD is $\sim 1\%$ per live birth (Qu et al., 2016; Triedman and Newburger, 2016). CHDs often result in long-term ill health and require long-term expert medical management, which significantly impact individuals, families, and society at large. In resource-limited settings such as sub-Saharan Africa (SSA), the social and clinical context of CHD is further challenged by access to new technologies and devising applications that are meaningful to local communities.

Many studies being conducted currently on CHDs are examining epidemiological data and clinical characteristics that will ascertain the burden and improve diagnostic power culminating into better management and corrective surgeries. These studies focus on the distribution patterns and risk factors that potentially lead to CHD births because of individual exposures. Due to the complications of CHD, it becomes a daunting task to seek corrective measures once diagnosis is made due to several factors, including socioeconomic variations, state-of-the-art facilities, and religion (Carlsson et al., 2016; Knowles et al., 2016).

The million-dollar question then becomes, ''Is it possible to use available tools and/or emerging diagnostic approaches and technologies such as AI and digital health to predict groups or individuals at considerable risk of CHD birth?''

Screening for high-risk CHD birth before and during pregnancy is clinically important as it will play a crucial role in early intervention and management. An example of the power of prediction tools is evident in the simple use of premarital sickle screening to inform reproductive and offspring management decisions in several African countries where the burden of sickle cell is significantly high (Abioye-Kuteyi et al., 2009; Gallo et al., 2010). Models with good predictive potential for identifying women whose offspring is prone to CHD will help manage the burden of this birth defect, especially in resource-limited settings.

We examine and share here the lessons learned in current attempts to implement AI and digital health in CHD for precision risk prediction and diagnosis in resource-limited settings.

AI Learning Architecture and Disease Prediction Algorithms

The clinical introduction of Internet of medical things, data-rich technologies such as next-generation sequencing (NGS), and mobile smart watches, which generate, collect, and store data on vital signs, will soon require that clinicians including cardiologists utilize and operationalize data from many fields of medicine to effectively diagnose CHD and its associated phenotype (Dimitrov, 2016; Lu et al., 2016; Ozdemir and Hekim, 2018; Reeder and David, 2016) and recommend appropriate management and policies.

The field of precision medicine is a medical discipline individually tailored with emphasis on disease diagnosis, management, and personalized therapy. Prediction of disease risk is a focus of precision medicine that is gradually gaining momentum and requires innovations for transformation in this postgenomic era. CHDs risk prediction in precision medicine using AI, a data-driven technique, will enhance cardiology practice and improve outcomes of CHD. The practice of cardiology is being overwhelmed with various data types that require complex analysis and interpretation while cardiologists are expected to competently perform.

The practical application of AI in cardiology lies in incorporating and utilizing data from a multidisciplinary group of clinicians in improving diagnosis and personalized management as in the case of clinical decision support systems. AI can be used to analyze and interpret electrocardiographs (ECGs) and echocardiographic (ECHO) images to elucidate underlying dynamic patterns relevant to addressing the defect (Alsharqi et al., 2018; Ouyang et al., 1997; Saini et al., 2013). There are so much underutilized data in the health care industry in resource-limited countries, which do not augur well for robust health care delivery in the current AI and digital health age. This article looks at the AI in CHDs and cardiology and the challenges faced in resource-limited settings and opportunities it presents.

The rising cost of health care is one of the most important factors worldwide in making disease prediction, a significant area of research. The development of big data in biomedical and clinical communities is providing important data types that can lead to early disease recognition, risk prediction, and accurate diagnosis. The datasets generated in these communities need to be queried to explore what is known, trend analysis and predict what is about to happen. Probing models that can decrypt the data and pick features that can lead to disease prediction will play a pivotal role in inching closer to personalized medicine (Fröhlich et al., 2018; Mesko, 2017).

AI comprising neural networks, machine learning (ML), natural language processing, and deep learning uses algorithms and developed disease-specific risk score predictive models to learn from various complex large-scale data types and make relevant predictions to complement clinical decisions, thus facilitating a more detailed understanding of the disease. Disease prediction has been made possible with the aid of AI and digital health techniques such as support vector machine (SVM) and artificial neural network (ANN) in several conditions (Mantzaris et al., 2008; Weng et al., 2016).

The computational intelligence technique Probabilistic Neural Networks (PNNs) was used in osteoporosis risk prediction and was shown to appropriately predict osteoporosis risk (Mantzaris et al., 2008). In this study, one of the unique highlights was that another technique that was used, multilayer perceptrons, was observed to have underperformed compared with PNNs. This means that the development of algorithms for disease prediction and prognosis requires robust ability to connect input and output data at different layers using precise vector mapping. Coronary artery disease (CAD) is a major cause of death and disability in both developed and developing countries, and has been a public health concern for the past few decades (Goyal et al., 2017; Sanchis-Gomar et al., 2016).

Eight ANN models comprising Levenberg-Marquardt, quasi-Newton (Broyden, Fletcher, Goldfarb, and Shannon—BFGS), quasi-Newton (one step secant), conjugate gradients of scaled, Polak-Ribiére, Fletcher-Reeves and Powell-Beale and back propagation were used to predict CAD with considerably high sensitivity, specificity, and accuracy (Çolak et al., 2008). A specific type of ANN known as Convolutional Neural Network (CNN) uses perceptrons (a ML unit algorithm) for supervised learning to analyze data and has also been widely used in predicting accurate diagnosis. CNNs have been applied to image and natural language processing and cognitive tasks in disease predictions. Deep CNN algorithm was used to predict and diagnose periodontally compromised teeth (Lee et al., 2018).

This study used a combination of pretrained deep CNN architecture and self-trained networks to optimize CNN algorithm and weights on periapical radiographic images for accurate prediction and diagnosis.

Certain disease conditions such as CAD, aneurysm, and congenital heart diseases require tests such as magnetic resonance imaging (MRI), coronary and cardiac computed tomography angiogram (CCTA), echocardiogram (ultrasound) and X-ray to be able to make effective diagnosis and prognosis. Due to differences in medical training and experience, the interpretation of these tests varies, which may lead to misdiagnosis and error in patient care. AI-based models and algorithms have been very useful in image processing that has led to improved diagnosis and prognosis of these conditions.

Alzheimer's disease (AD), an irreversible brain degenerative disease, affects people ~65 years old and above. Although the progression of AD can be delayed, there is no effective cure. One of the ways of diagnosing AD at the early stages for patient care and management is using a functional molecular imaging modality known as fluorodeoxyglucose positrons emission tomography, which is very good to understanding the anatomical and neural changes of the brain in AD. One of the challenges with image analysis is the high level of noise and often low resolution, which makes resolving and diagnosing or predicting very subtle. A voxel-based volumetric image analysis algorithm was performed to classify three groups, which made it possible to extract information about location of metabolic changes induced by AD that directly relies statistical features and brain regions of interest (Vigneron et al., 2016).

A ML model that resulted in an SVM with radial-basis function kernel was used to successfully predict AD conversion individuals with mild cognitive impairment and was shown to have high cross-validation performances (Grassi et al., 2018). A long-short-term memory recurrent neural network, which relies on an enhanced "many-to-one" neural network architecture to support shift of time steps, was adopted to predict patient's AD progression with efficient accuracy (Wang et al., 2018). In most of the predictive models that have been highlighted above, it is observed that most of the algorithms involve neural networks. The outcome of the prediction is highly dependent on the dataset and developed algorithm, which influences the robustness of the outcome and reliability.

The datasets that are used in the predictive models are from electronic health records (EHRs) and clinical records. The deployment of AI systems in predictive models requires "trained" data that are mostly generated from clinical activities, including screening, diagnosis, and treatment. The clinical data used include data on demographics, electronic data from medical devices such as MRI, echocardiographs (ECHO), computed tomography (CT) scans, clinicians' notes, laboratory results, physical examinations, and medical images (Jiang et al., 2017; Raposo, 2015). The motivation for trained data is for learning between similar subject groups, associations between subject features and outcomes of interest (Jiang et al., 2017). The dataset is critical in the accuracy of prediction and diagnosis, which also brings to the fore data-driven methods for prediction and diagnosis with our focus on CHDs and cardiology.

Challenges with "Big-Data" in Precision Cardiology

AI and digital health predictive models in resource-limited settings

There are massive amounts of data generated in the hospitals, including clinical information, genomic data, and data from EHRs. Advancement in big data is continuously playing a key role in health care management, which enables evaluation of big datasets for disease management, therapeutic dosing and prediction (Topol, 2019). Despite the vast amount of data generated in health care, majority of the data remain unmined primarily due to the challenge of storing, managing, and the complex nature of the datasets (multidimensional and nonlinear relationship between data variables). Utilization of these datasets especially in rare conditions such as CHD using AI predictive models will be useful for identifying individuals at risk of having CHD children, having complications after a successful heart surgery and management (Kuo et al., 2018; Min et al., 2019; Moonesinghe et al., 2013; Olive and Owens, 2018).

CHDs are the most severe common birth defects affecting ~1% live births with uncertain etiology (Abdulkadir and Abdulkadir, 2016; Hoffman, 2013). The devastating effects of CHD are seen in structural and great vessel disorders that are seen at birth, which often require major surgery to correct the condition. CHD diagnosis is often performed with electrocardiogram (ECG), electrocardiography (ECHO), and cardiologist's experience. Vast amount of data is generated on CHD parents, the affected proband, and other siblings in both developed and developing countries (Williams and McCrindle, 2002; Witter et al., 2013).

These data include laboratory data on other diagnosis, maternal and paternal age, prenatal and postnatal data on mothers such as gestational diabetes, pre-eclampsia, administered medication, medical history, smoking and alcohol consumption and genetic data (Abqari et al., 2016; Bahtiyar et al., 2007; Czeizel et al., 2015). Medical informaticians have proposed the use of AI models such as machine and deep learning to predict accurate patient response to medication, predisposition to certain conditions, and diagnosis (Clifton et al., 2015). This review provides an overview of the applications of AI in disease prediction, the challenges and opportunities in resource-limited settings, and lessons learned. The interoperability and usage of data generated from data sources in resource-limited settings are discussed in the next sections.

EHRs and eHealth in the era of AI and digital health

EHRs have become a key bedrock in the health care management system as it presents easy access to patient data. This is a digital record of a patient's paper chart that provides detailed information on a patient's past and present medical history, diagnosis, laboratory test, immunization information and medications. Despite the amount of information provided in EHRs, there is more to EHRs than clinical data collected as it is inclusive of broader view of a patient's care. In countries with limited resources, EHR is an underutilized data source although it has great prospects in clinical practice, especially when used as part of predictive algorithms (Wu et al., 2018). Sub-Saharan African countries are lagging behind in the availability and use of EHRs, although there has been an appreciable increase in utilization and availability of EHRs over the last decade (Akanbi et al., 2012).

Ghana, a sub-Saharan country, for example launched an e-strategy to incorporate electronic records as part of the health care system to improve health care delivery in the country (Afarikumah, 2014). Such an initiative will generate big data, which can be utilized for the robust disease prediction. EHRs offer promising opportunities for application in disease-risk prediction such as CHD and accurate diagnosis. EHRs from women visiting a health facility in the United States were used to predict breast cancer phenotypes and their ''harmfulness'' using developed regularized predictive models a year in advance (Wu et al., 2018).

The prediction model that demonstrated significantly higher predictive performance was Lasso Logistic Regression (LassoLR), which used data extracted from EHRs on demographics, diagnosis, medications, laboratory analysis, age at diagnosis, and treatment procedures. There are reported studies on predictive models using EHRs on cardiovascular-related conditions such as heart failures and hypertension in developed jurisdictions or countries (Table 1).

AI algorithms have been successfully used in predicting cardiovascular conditions from electronic heath records with very few records that can give a prediction accuracy of $\sim$70% (Table 1). This puts the utilization of EHRs for prediction a key factor in precision cardiology, especially in resource-limited countries where such records are generated in terabytes of data.

AI models such as ML and deep learning algorithms for disease-risk prediction using EHRs have been successfully deployed in key clinical queries over the past decade in advanced countries such as the United States, Germany, and China. These predictive models explore new features or nonlinear relationships in EHR data using predictor variables such as age, gender, diagnosis, comorbidities, and laboratory results than observations. These predictive variables are continuously generated and available in developing countries, and can be incorporated into predictive algorithms.

However, one of the challenges of developing EHRs in resource-limited environments is capacity to advance stable computerized systems to keep such records or data. Health records in developing countries (resource-limited ones) are mostly kept in paper formats, which over time get some of the pages being lost or torn into pieces due to mishandling. Transitioning from paper-based record keeping to electronic records will require massive investments in resources, which most developing countries lack (Dennehy et al., 2011; Jawhari et al., 2016). Despite these challenges, several sub-Saharan African countries such as Ghana, Nigeria, and South Africa knowing the opportunities associated with keeping robust EHRs are actively deploying measures to cause significant investment in e-health systems (Afarikumah, 2014; Ford et al., 2018; Katurura and Cilliers, 2018).

SVM and random forest (RF) algorithms integrated with logistic regression models are some of the current AI enterprises used to predict risks and outcomes of heart failures and other cardiovascular-related conditions. Challenges of EHR data include missing variables, misclassification, errors in measurement, disorganization among others. Utilizing multiple algorithms and picking the single best algorithm or using a weighted approach help lessen the challenge of quality issues with EHRs.

High-Throughput Molecular Profiling Data: "Omics" in Resource-Limited Settings?

Advancement in laboratory technology is at the forefront of providing robust reliable biological data useful for clinical application. Big data generation is being pioneered through ''omics'' methods such as genomics, proteomics, metabolomics, metabonomics, and transcriptomics. The list of ''omics'' methods is continuously increasing even as the goal toward precision or personalized medicine is being pursued (Bope et al., 2019). How are these data generated in resource-limited settings and how will integrating AI lead to risk prediction for CHD? The molecular basis of CHD has become a critical theme for medical research and discourse among cardiologists and medical scientists for the past few decades (Garg, 2012; Huang et al., 2010; Sander et al., 2006; Wang et al., 2013). Understanding the mechanisms of CHD has primarily focused on biomarker identification development for precision phenotyping and diagnosis.

However, due to the massive amount of data generated and segregation in the data analysis the goal of reliable biomarker identification for CHD phenotypes is becoming elusive. The genesis of AI methods such as ML and deep learning neural network algorithms will provide the means to integrate such big data for reliable analytical outcome (Topol, 2019; Trakadis et al., 2018).

Table 1. Predictive Models Using Electronic Health Records in Cardiovascular-Related Diseases

Country	*Records obtained (patient numbers)*	*Predictive algorithm(s) models*	*Prediction target outcome*	*Data period used for prediction (years)*	*Predictor variables*	*Reported prediction accuracy*	*References*
Slovenia	755,966	LASSO	Cardiovascular diseases polypharmacy prediction	2006–2016	Age, patient id, gender, geographical information, drug identifier, doctor information	0.898–0.901	Kocbek et al. (2018)
United Kingdom	200,000	Not indicated	Cardiovascular risk prediction	≥2005	Age, gender, SBP, laboratory results, ethnicity, IMD score, family history of angina, smoking status, BMI, history of diabetes, chronic kidney disease, rheumatoid arthritis, history of atrial fibrillation, and history of LV hypertrophy	0.938–0.940	Stevens et al. (2018)
United States	11,510	DUNs	Heart failure readmission risk	2014–2015	Age, gender, ethnicity, marital status, education, race, comorbidities, hospital utilization information, diagnoses, procedures, laboratories, medications, unstructured data elements: physician notes and discharge summaries	0.764	Golas et al. (2018)
United States	171,510	ML algorithm	Cardiology survival risk	1998–2017	Age, gender, smoking status, height, weight, heart rate, blood pressure, laboratory results	0.791–0.893	Samad et al. (2018)
United Kingdom	41,373	Multivariate mixed-effects linear models	Cardiovascular diseases risk	1997–2016	Age, gender, ethnicity, marital status, education, race, comorbidities, practice-level data, diagnoses and symptoms, specialist referrals, laboratory testing, disease monitoring, prescribing, and death	0.783	Paige et al. (2018)
United States	45,579	l_1-Regularized sSVM classifier	Cardiovascular conditions	2005–2010	Age, gender, and race, weight, height, BMI, medical history, diagnoses, procedures, office visits, history of drug prescription	0.7667–0.7806	Brisimi et al. (2018)
China	296	SVM, Weka	Depression and hypertension predictive markers	2011–2016	Age, gender, race, marital status, blood tests, vital signs, expenses, laboratory results	0.735	Song et al. (2018)
United States	15,209	L1-regularized LR and RF models	Early detection of heart failure	2003–2010	Age, gender, ethnicity, smoking, alcohol use, diagnoses, medications, laboratories, hospitalization, two-dimensional Echo, transesophageal Echo, stress Echo, pulse, blood pressure height, weight, temperature	0.791	Ng et al. (2016)
United States	119,749	CPXR(Log) algorithm	Heart failure risk	1993–2013	Age, gender, race, ethnicity, BP, BMI, laboratory results, medications, comorbidities	0.809–0.914	Taslimitehrani et al. (2016)
United States	16,971	NLP, IMRS	High-risk heart failure	2013–2015	Age, gender, laboratory, pharmacy, clinical, echocardiogram, diuretic use, b-type natriuretic peptide level, ejection fraction, eligibility for any of the centers for Medicare and Medicaid Services or Joint Commission HF core measures, discharged with a primary diagnosis of HF	0.9745	Evans et al. (2016)
United States	5044	COX proportional regression model, RF, LR, SVR, DT, AB	Heart failure survival risk prediction	1993–2013	Age, gender, race, ethnicity and survival status, laboratory results, medications, comorbidities, BMI, calcium channel blocker	0.68–0.80	Panahiazar et al. (2015)

AB, ada boost; BMI, body mass index; BP, blood pressure; DT, decision tree; DUNs, deep unified networks; Echo, echocardiographic; IMRS, Intermountain Risk Score; LASSO, L1-norm model; LR, logistic regression; LV, left ventricular; ML, machine learning; NLP, natural language processing; RF, random forest; SBP, systolic blood pressure; sSVM, sparse Support Vector Machine; SVM, support vector machine; SVR, support vector regression; Weka, Waikato Environment for Knowledge Analysis.

Cardiogenomics Data and AI

There has been a tremendous improvement in using genomic information to determine genomics regulation and interaction with the environment in the control of complex biochemical function in health and disease. Generation of complex genomic data/information up to the transcriptomics level has been made possible using NGS technologies such as whole exome sequencing (WES), whole genome sequencing (WGS), and RNAseq/Ampliseq. The influx of low-cost genomic technologies has made the application of genomics in CHD research very affordable. Due to the complexity of genomic data and analysis, predictive risk profiling for disease conditions has become a big challenge.

AI methods have enabled the potential use of genomic data and analysis in risk prediction in health and diseases. For example, ML algorithms were used to integrate genetic, epigenetic, and phenotypic data from the Framingham Heart Study to build and test a RF classification for coronary heart diseases at an accuracy, sensitivity, and specificity of 78%, 0.75, and 0.80, respectively (Dogan et al., 2018). Besides identifying at-risk patients, AI integration of genomic data can assist in precision therapy, primarily it presents a multidimensional scope of the disease (Fröhlich et al., 2018; Madhukar and Elemento, 2018).

Several genome-wide and phenome-wide association studies have shown that cardiovascular disease pathophysiology occurs from complex interactions from the genome (Lucas et al., 2012; Musameh et al., 2015). Risk score assessment is achieved for a potential cardiovascular condition incidence using classified risk scores of factors that potentially predispose one to the condition (Giampaoli et al., 2006; Studziński et al., 2017). In CHD risk prediction, genomic data will be key in calculating risk scores for predisposition toward acquisition, phenotyping and survival postsurgery, a much needed intervention in resource-limited settings such as sub-Saharan African countries.

NGS-based technologies can produce big genomic datasets that can be used in combination with AI for CHD disease prediction. Targeted NGS has provided data on three pathogenic gene variants (*MYH6*, *NOTCH1,* and *TBX5*), which potentially explains the defects observed in six Belgian families (Jia et al., 2015) and data from WES on nine kindred with familial CHD showing *GATA4* mutation in the transactivation domain p.G115W associated with familial atrial septal defects (ASDs) and splice donor site mutation in *MYH11* (*c.4599+1delG*) identified in familial patent ductus arteriosus (PDA) (LaHaye et al., 2016). Such studies expanded to families will generate more data, which could be analyzed to determine mode of inheritance and risk prediction in unborn babies.

AI risk prediction will be beneficial to familial congenital heart conditions for planning and management. A study conducted using WES in combination with a CHD-related gene filter discovered a novel variant in *TBX20* c.526G>A associated with familial ASD where it was identified to cosegregate in all affected members of the family (Liu et al., 2014). The success story of ML algorithms utilizing WES data is reported in predicting at-risk individuals for conditions such as schizophrenia (SCZ). Gradient trees with regularization (eXtreme Gradient Boosting implementation) yielding an accuracy of >85% predict SCZ occurrence in $\sim$600 asymptomatic unaffected individuals confirmed with SCZ cases (Trakadis et al., 2018).

AI methods such as supervised ML algorithms can be used in WES data by recognizing patterns of variants observed in the different genes to determine which set of genes can potentially predict risk for an individual. WES data from CHD research can be subjected to such AI algorithms and used to predict risk for individuals who can potentially have babies born with CHD.

There is paucity of data on WES/WGS studies in resource-limited settings. Limited resources such as NGS facilities and research funds motivate samples to be sent to collaborators in advanced economies where data specific to such resource-limited countries get lumped in the entire studies preventing specific analysis to such settings. Although CHDs are a burden in resource-limited settings such as SSA, there is lack of investment in database setup for clinical, epidemiological, and genomic data. With the growing influence of AI in medicine and its global implementation in helping make important life-dependent health decisions, resource-limited countries can gradually implement AI in CHD risk prediction as an effort to plan and manage the condition.

Cardioproteomics Data and AI

Proteomics is contributing significantly in our understanding of cardiovascular diseases. The improvement in analytics and computational algorithm added another critical component in proteomics, which is the ability to address cardiac physiology question beyond protein species discovery. The increasing number of datasets generated by mass spectrometry has increased the intersection between proteomics and big data science (Lam et al., 2016). Proteomics data provide the dynamics to disease phenotype and could be used for understanding of the mechanism, which leads to the development of cardiovascular disease (Barallobre-Barreiro et al., 2013).

The use of AI in the field of ''omics'' especially utilizing proteomics data is at early stage. However, despite the challenge in terms of algorithm development and data management, the use of AI in proteomics will enhance our knowledge and understanding on cardiovascular diseases. There are reported studies on predictive models using proteomics on cardiovascular conditions (Table 2).

Readily available biomarkers for CHD and phenotype diagnosis are a perennial challenge that is facing the clinical community. However, the development of high-throughput techniques such as proteomics is earnestly promoting biomarker discoveries and with the added benefit of AI algorithms, the level of accuracy and specificity of CHD prediction and diagnoses in medical ecosystem will be apt. In a recent study to gain insights into the potential mechanisms of cardiovascular development using ventricular septal defect (VSD) patients and healthy controls using peptidomic analysis of amniotic fluid, it was discovered that $\sim$35 signature peptides located with key functional domains of their precursor proteins could serve as biomarkers for VSD (Li et al., 2016).

Although the analysis used peptide profiles obtained from mass spectrometry, a bioinformatics ingenuity pathway analysis program was used in obtaining the 35 peptides and ''predicting'' the role of the peptides in the cardiovascular system morphogenesis and cardiogenesis. LASSO has been extensively utilized coupled with other AI algorithms in analyzing proteomic data and successfully used in revealing differential signaling and insulin-like growth factor pathways (Erdem et al., 2016; Huang et al., 2013). There is ample

Table 2. Predictive Models Using Proteomics in Cardiovascular Related Diseases

Year	*Country*	*Records obtained (Patient No.)*	*Predictive algorithm models*	*Prediction target outcome*	*Data period used for prediction (year)*	*Predictor variables*	*Reported prediction accuracy*	*References*
2018	Sweden	92	Lasso penalized Cox proportional hazards regression	Plasma kidney injury molecule-1 (KIM-10 as a risk marker for cardiovascular mortality and coronary artery calcification)	Discovery cohort MIMICK, 2003–2004; replication cohort, SKS, 2012–2014	Age, sex, blood	0.78	Feldreich et al. (2019)
2016	United States	938	LASSO; 9-protein risk score	Cardiovascular prediction protein-based risk score for patient with coronary heart disease	2000–2002	Age, sex, total cholesterol, HDL-C, diabetes, SBP, and current smoking status	0.74	Ganz et al. (2016)
2017	United Kingdom	98	LASSO	Prediction of term pre-eclampsia		Age, parity (nulliparous, multiparous), race (black, white, Asian, mixed), BMI, MAO mom, previous PE (multipara-PE, multipara-no PE), FH PE-mother	0.987	Bahado-Singh et al. (2017)
2015	France	198	SVM; sparse partial least-square discriminant analysis; LASSO	Predict the potential value of plasma proteomic profiling for risk stratification in heart failure	1998–2010		0.66–0.68	Lemesle et al. (2015)
2018	Sweden	1211	Multivariable Cox regression; gradient-boosted ML; LASSO regularize Cox regression	Prediction of major cardiovascular events in type 2 diabetes	2005–2008	Event/total *N*, follow-up and years, % women, age and year, BMI, HbA, eGFR, SBP, total cholesterol, LDL-C, HDL-C, current smoker, history of cardiovascular disease, antihypertensive medication, statin in use	0.68–0.75	Nowak et al. (2018)
2018		196	Machine learning	Target proteomics to predict the presence of high-risk plague		Age, year, male, BMI, risk factor (DM type II, hypertension, hyperlipidemia, current smoker, family history), type of chest pain (typical angina, atypical angina, nonspecific chest discomfort), laboratory tests (TC, HDL-C, triglycerides, hs-troponin, NT-proBNP, creatinine, eGFR <60 mL/min, CRP ≥2.5 mg/L), medication use (statin, acetylsalicylic acid, betablocker, ACE-inhibitor/ARB), other (SBP, DBP, Framingham risk score)	0.79	Bom et al. (2019)
2015	Europe	42	LASSO	Predictive of incident cardiovascular in type 2 diabetes people		Male, age, diabetes duration, BMI, height, SBP, DBP, $HbA_{1c,}$ triacylglycerols, HDL-C, LDL-C, eGFR, smoking status, insulin therapy, antihypertensive therapy, lipid lowering therapy, aspirin therapy	0.66–072	Looker et al. (2015)

DBP, diastolic blood pressure; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol.

indication that LASSO and other AI algorithms can be used to analyze proteomics data for prediction of cardiovascular-related diseases (Table 2), which presents AI as the future (if not the current) tool for accurate disease prediction in the medical system, especially for conditions with relatively high mortality such as CHDs.

Cardiac Imaging, Signal Detection, and AI

Noninvasive imaging technologies play a key role in diagnosis, prognosis, and management of patients with CHDs. Imaging techniques such as echocardiography, cardiac MRI, nuclear cardiology, and CT have essentially formed the bedrock of cardiology with the interpretation and analysis of data subject to the expertise of the clinician with a certain ratio of subjectivity. Due to different training and exposure of clinicians, cardiac imaging diagnoses have led to variable interpretation and management strategies for patients most especially with complicated conditions such as CHDs.

The quality of diagnosis in the various cardiac imaging techniques is also strongly dependent if the clinician has enough knowledge about the normal growth and development of the heart, different types of CHDs, and some level of knowledge on principles of ultrasound physics. There have been incidences of diagnostic errors using cardiac imaging such as echocardiography (Moradian, 2012) and MRI (Raggi et al., 1996) in CHDs in pediatric patients and a large left atrial thrombus patient. One major challenge with interpretation could also be resolution of images, which brings about the variable diagnosis.

AI presents a reliable and accurate alternative for the management and diagnosis of complex heart conditions with resultant images, which drastically reduce the potential risk of human errors (Alsharqi et al., 2018; Itchhaporia et al., 1996). If the key features in cardiology images are used to train an algorithm, it will be easy for the algorithm to detect and report accurately the diagnosis, and there will be consistency in the reportage. With the benefit of AI toolsets, which have grown to include robust models such as Bayesian classifiers, decision trees, neural networks, SVM, clustering algorithms, and components analysis, accuracy of image interpretation in cardiology can be improved to make diagnosis, prognosis, and management of CHDs apt (Table 3).

There have been efforts recently to practically inculcate AI toolsets such as ML in cardiac imaging, which has brought significant progress in autointerpretation and autoquantification in echocardiography using RF, SVM, and trained ensemble classifiers (Narula et al., 2016; Tsang et al., 2016) and cardiac MRI using linear SVM classifier (Afshin et al., 2014).

ECG measures cardiac electrical activity from the body surface, which serves a crucial step in diagnosing, understanding, and predicting cardiovascular disorders. The role of ECG in cardiology has to do with diagnosis and prognosis pre- and postsurgery as heart function can be monitored from ECG recordings. Interpreting ECG signals is a very subjective topic as looking for key features that help understand the function of the heart also requires a high degree of subjectivity. Modern signal processing and AI techniques will greatly increase the accuracy and diagnostic power of ECG exponentially, and take out a significant amount of ambiguity in interpretation.

AI algorithms that will tease out certain features of ECG such as P-, P′-, Q-, R-, R′-, S-, S′-, T-, and T′-wave duration, amplitude, area, and intrinsicoid middle and end of ST-segment amplitudes, amplitude at the point of 60 msec from J point, STJ amplitude, total QRS area, balance, deflection balance, intrinsicoid, leads (aVL, aVr, aVF, I, II, III, V1, V2, V3, V4, V5, V6) will bring a massive improvement with cardiovascular diagnosis in the usage of ECG (Ambale-Venkatesh et al., 2017; Lyon et al., 2018).

There is vast amount of data from echocardiography, cardiac CT scans, and ECG being generated in SSA used in diagnosing and managing cardiac conditions, which have been underutilized in terms of predictive abilities. Such data have been kept in storage devices without proper analysis for precision cardiology in resource-limited countries. AI is the imminent development for accurate cardiac image interpretation, and embracing it in resource-limited countries will sharpen diagnosis and management of complex heart malformation conditions such as CHDs.

AI Models for Precision Cardiology

The amount of data generated by ''omics,'' EHR, imaging, and signal detection technologies is considerably huge and contain unprecedented information. Data obtained could provide a basis for risk factor analysis, therapeutic interventions, and prognosis of diseases. For example, it is possible to predict the risk of myocardial infarction using available records of patients and their family, and find patterns associated with heart diseases. The analysis and process of this information require the implementation of big data technologies. The purpose of generating huge volume of data is to extract meaningful information to improve diagnosis and patient care.

Clinical data derived from different sources such as genomic, metabolomics, proteomic, imaging, and EHRs contain useful information for a clinician. The physician faces significant challenges of combining all this information for better diagnosis and prediction of CHD and cardiology conditions. In the literature, most of the studies focus on epidemiological data and/or clinical characteristics to predict adverse pregnancy outcomes (Goyal et al., 2015; Jung et al., 2016; Kim et al., 2015; Li et al., 2017; Seravalli et al., 2014).

Data-driven methods to manage and model big data such as data mining and ML are good assets to assist the clinician for diagnosis and prognosis of cardiology conditions such as CHD. Data mining has played an important role in smart medical systems and digital health. Large medical dataset from different sources can be managed by data mining to find hidden information or pattern (Chaurasia and Pal, 2009; El Houby, 2018; Lee et al., 2009; Purusothaman and Krishnakumari, 2015). Evidence of AI in precision cardiology is observed in the usage of data mining algorithms to predict heart diseases and interpret echocardiograms, which have proven useful for diagnosis and prognosis.

An automated ML method algorithm with convoluted neural networks was developed by Zhang et al. (2018) to interpret transthoracic echocardiograms using >10,000 echocardiograms. The subjective human factor will be minimized with trained datasets that to interpret data was observed in a report by Davies and colleagues where they found a change of mind in $\sim 10\%$ of clinical experts in percutaneous coronary intervention compared with a 100% consistency in their ML model (Davies, 2018a, 2018b).

Table 3. Predictive Models Using Imaging and Signal Detection in Cardiovascular-Related Diseases

Year	*Country*	*Records obtained (Patient No.)*	*Predictive algorithm model*	*Predictive target outcome*	*Technique used*	*Predictor variables*	*Diagnosis*	*Reported prediction*	*References*
1997	Japan	132	ANN	ECG interpretation of QS complex	Automated ECG	Q-wave (or QS complex) or small R progression in leads V7 and V2	Anterior wall myocardial infarction	90–93	Ouyang et al. (1997)
2005	United States	241	ANN	Reliable screening device for diagnosis of heart murmurs in pediatrics	Phonocardiograms	Heart sound recordings	Innocent murmurs, VSD	90–93	Bhatikar et al. (2005)
2016	United States	159	Adaptive analytics algorithm	Automated quantification of three-dimensional transthoracic echocardiography-derived LA and LV volumes and LVEF	Echocardiography; cardiac magnetic resonance; electrocardiogram	LA, LV volumes, LVEF volumes	—	—	Tsang et al. (2016)
2018	United States	223	Mel-Frequency Cepstral Coefficient and Hidden Markov Model	Develop an innovative mobile Health (mHealth) service platform	Cardiac auscultation linked to smart phones	Normal heart sound heart murmurs, cardiac signal	Aortic, mitral, tricuspid and pulmonary regurgitation, aortic stenosis, ASD, VSD, PDA, mitral stenosis, mitral valve prolapse	92.68	Thiyagaraja et al. (2018)

ANN, artificial neural network; ASD, atrial septal defects; ECG, electrocardiographs; LA, left atrial; LVEF, left ventricular ejection fraction; VSD, ventricular septal defect.

An implementation of the AI pipeline, which integrate omics data, EHR, and imaging for congenital cardiac defect, will improve clinician diagnosis (Fig. 1).

Current Landscape and Future Perspectives

Outlook in SSA

SSA is saddled with a lot of communicable diseases such as HIV/AIDS and tuberculosis and recently a lot of noncommunicable diseases. The incidence of CHDs in sub-Saharan African countries is no different from what happens in the world with 1% of live births being born with CHDs. Accurate diagnosis, surgical correction, and effective treatment of CHDs are standard practices in advanced economies. In SSA, major challenges in CHD management include clinical management such as surgical corrections, affordability, and postsurgery management of patients.

Financial challenges continue to plague families and individuals with CHDs in SSA. AI and digital health implementation will improve risk prediction in clinical settings in SSA. The benefit of knowing at-risk individuals in advance for CHD and other cardiology conditions is to help in clinical and personal decisions. For resource-limited settings, this will be critical for personalized and preventive medicine. A look at the AI disease-risk prediction landscape shows advanced application of AI in health care in developed countries.

Screening and risk prediction have been implemented in chromosomal anomalies such as Down's syndrome, which gives parents the choice of continuing with the pregnancy or terminating it. Without requiring very complex information, EHRs can still provide enough predictive power for conditions such as CHD in clinical settings. For countries with more advanced biomedical settings, addition of genomic, proteomic data and imaging records coupled with AI algorithms will provide a more robust predictive tool for CHD prevention and management. SSA can take advantage of the AI-digital health revolution to bring a vast improvement in the health care system.

Key lessons: innovative policy-oriented framework in health

Sub-Saharan African countries are making effort to improve the management of CHD. The CHD burden in the medical ecosystem in SSA cannot be underestimated with significant interventions required to improve in areas such as preventive approaches, corrective surgery, and management. A lot of data are being generated in health facilities in SSA but remain underutilized due to lack of proper policy-oriented framework for such data in the health care sector. The advent of AI and digital health calls for an all-inclusive public health approach to assess, predict risk and improve diagnostic capacity for CHD in public health.

In SSA where specialist physicians are even a problem, medical AI technology will be more beneficial as most of the work can be performed by the AI tech leaving less work for specialist physicians to work in terms of patient management and treatment of CHD patients. AI-related methods have been successfully implemented in rural areas in countries such as India and China with less trained physicians and health workers available. The overall consistency of the AI diagnostic system and physicians was 94% for ~ 1000 patients in India, while an all-in-one diagnostic station, which run 11 tests, including blood pressure, ECG, and routine urine and blood analyses, has been used in rural community health care settings in China.

Policy-oriented frame work for health care and CHD management in SSA will help achieve the goal of implementing AI and digital health in CHD and cardiology management. We hereby prescribe some policies aimed at achieving this goal:

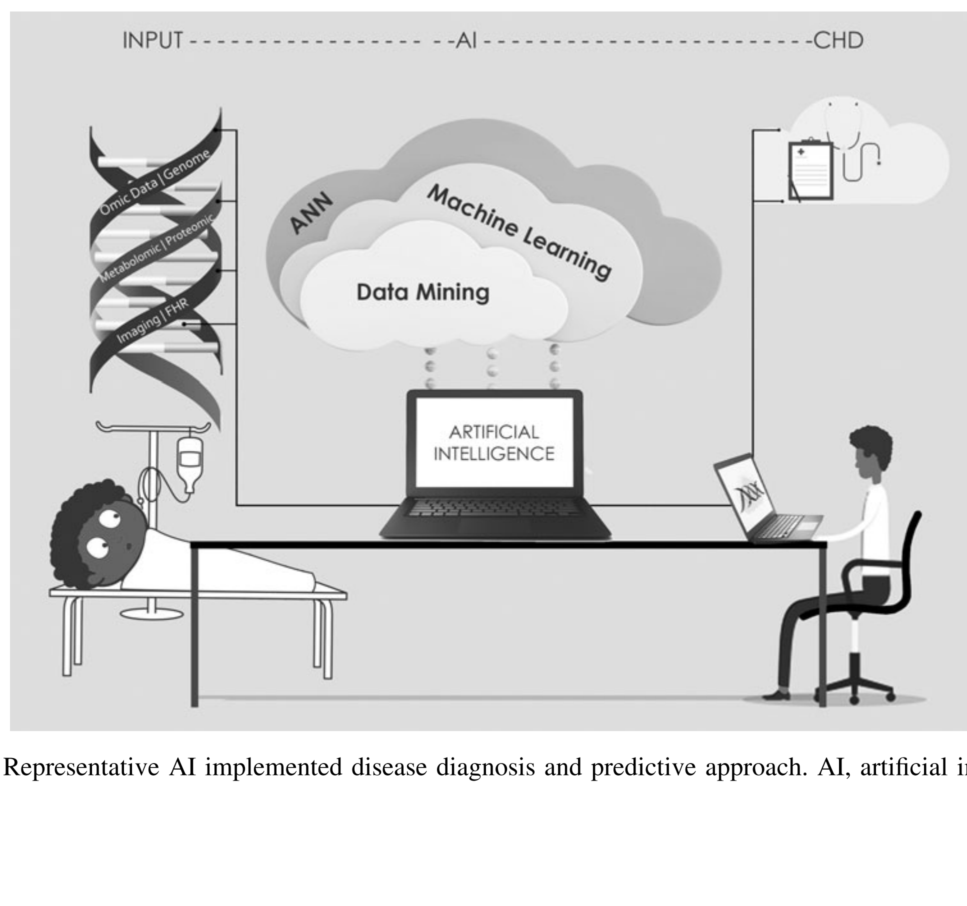

FIG. 1. Representative AI implemented disease diagnosis and predictive approach. AI, artificial intelligence.

Top 10 Lessons Learned

The lessons for resource-limited settings such as sub-Saharan African countries are summarized:

(1) Implementation of eHealth as part of health care in SSA. eHealth will provide a range of systems and services that include clinical information from EHRs, integrated information network, and disease registries with nonclinical systems. There is paucity of information on CHD in SSA and AI-eHealth will flag such incidence to improve management.
(2) Digitization of all EHRs generated in health facilities transferable to national servers with appropriate backup. Such data once generated will be useful in AI-oriented risk prediction and diagnosis of conditions such as CHD. Common sources of EHR in CHD include comorbidities, ECG, echocardiography, CCT (cardiac computer tomography), laboratory results, family medical history, etc. EHR resources provide several phenotypic information undeniably more than any lone registry.
(3) Revamping laboratory services in health care facilities to include different omics level analyses as part of health care to improve diagnostic capacity for CHD and improve familial risk prediction.
(4) Creating a common platform for medical research institutions and health care facilities to translate CHD research, AI, and policies to the medical ecosystem in health care facilities.
(5) Capacity building training clinicians and health practitioner to use the AI-eHealth system and EHR.
(6) A centralized AI-eHealth country data center, coordination center and development of legal policy framework, which will ensure data processing, management, storage, confidentiality, and appropriate data handling.
(7) Further research and investments are required in AI-eHealth in resource-limited settings to accelerate early identification of risk factors/susceptibility to CHD and improve diagnosis.
(8) A conscious implementation of AI-health informatics for echocardiogram, ECG, and CCT interpretation to enhance cardiology patient care and management.
(9) AI-oriented health care industry will augment the capabilities of clinical and health professionals and give them reliable tools with predictable outcome and improvement in quality of service.
(10) The current cost of health care in SSA will be greatly reduced with the implementation of AI and cognitive computing in the health care industry.

These top 10 lessons on AI and digital health summarized in this expert review are relevant broadly beyond CHD in cardiology and medical innovations. As with AI itself that calls for systems approaches to data capture, analysis, and interpretation, both developed and developing countries can usefully learn from their respective experiences as digital health continues to evolve worldwide.

Acknowledgment

The authors thank Mr. Mzodidi Tutuka for assistance in illustrating Figure 1.

Author Disclosure Statement

The authors declare they have no conflicting financial interests.

Funding Information

N.E.T. is supported by the Faculty of Health Sciences, University of Cape Town through the Harry Crossley Foundation. F.E.A. is funded by DELGEME (grant 107740/Z/15/Z).

References

Abdulkadir M, and Abdulkadir Z. (2016). A systematic review of trends and patterns of congenital heart disease in children in Nigeria from 1964–2015. Afr Health Sci 16, 367–377.

Abioye-Kuteyi EA, Oyegbade O, Bello I, and Osakwe C. (2009). Sickle cell knowledge, premarital screening and marital decisions among local government workers in Ile-Ife, Nigeria. Afr J Prim Health Care Fam Med 1, 022.

Abqari S, Gupta A, Shahab T, et al. (2016). Profile and risk factors for congenital heart defects: A study in a tertiary care hospital. Ann Pediatr Cardiol 9, 216–221.

Afarikumah E. (2014). Electronic health in Ghana: Current status and future prospects. Online J Public Health Inform 5, 230.

Afshin M, Ayed IB, Punithakumar K, et al. (2014). Regional assessment of cardiac left ventricular myocardial function via MRI statistical features. IEEE Trans Med Imaging 33, 481–494.

Akanbi MO, Ocheke AN, Agaba PA, et al. (2012). Use of electronic health records in Sub-Saharan Africa: Progress and challenges. J Med Trop 14, 1–6.

Alsharqi M, Woodward WJ, Mumith JA, et al. (2018). Artificial intelligence and echocardiography. Echo Res Pract 5, R115–R125.

Ambale-Venkatesh B, Yang X, Wu Colin O, et al. (2017). Cardiovascular event prediction by machine learning. Circ Res 121, 1092–1101.

Bahado-Singh R, Poon, LC, Yilmaz A, et al. (2017). Integrated proteomic and metabolomic prediction of term preeclampsia. Sci Rep 7, 16189.

Bahtiyar MO, Dulay AT, Weeks BP, Friedman AH, and Copel JA. (2007). Prevalence of congenital heart defects in monochorionic/diamniotic twin gestations: A systematic literature review. J Ultrasound Med 26, 1491–1498.

Barallobre-Barreiro J, Chung, YL, and Mayr M. (2013). Proteomics and metabolomics for mechanistic insights and biomarker discovery in cardiovascular disease. Rev Esp Cardiol (Engl Ed) 66, 657–661.

Bhatikar SR, Degroff C, and Mahajan RL. (2005). A classifier based on the artificial neural network approach for cardiologic auscultation in pediatrics. Artif Intell Med 33, 251–260.

Bom MJ, Levin E, Driessen RS, et al. (2019). Predictive value of targeted proteomics for coronary plaque morphology in patients with suspected coronary artery disease. EBioMedicine 39, 109–117.

Bope CD, Chimusa ER, Nembawave V, et al. (2019). Dissecting in silico mutation prediction of variants in African genomics: Challenges and perspectives. Front Genet 10, 601.

Brisimi TS, Chen R, Mela T, et al. (2018). Federated learning of predictive models from federated electronic health records. Int J Med Inform 112, 59–67.

Carlsson T, Marttala UM, Mattsson E, and Ringnér A. (2016). Experiences and preferences of care among Swedish immigrants following a prenatal diagnosis of congenital heart defect in the fetus: A qualitative interview study. BMC Pregnancy Childbirth 16, 130.

Chaurasia V, and Pal S. (2009). Early prediction of heart diseases using data mining techniques. Caribbean J Sci Technol 1, 208–217.

Clifton DA, Niehaus KE, Charlton P, and Colopy GW. (2015). Health informatics via machine learning for the clinical management of patients. Yearb Med Inform 10, 38–43.

Çolak MC, Çolak C, Kocatürk H, Saðýroðlu S, and Barutçu I. (2008). Predicting coronary artery disease using different artificial neural network models. Anatol J Cardiol 8, 249–254.

Czeizel AE, Vereczkey A, and Szabo I. (2015). Folic acid in pregnant women associated with reduced prevalence of severe congenital heart defects in their children: A national population-based case-control study. Eur J Obstet Gynecol Reprod Biol 193, 34–39.

Davies J. (2018a). Cerebria-1: Machine learning vs expert human opinion to determine physiologically optimized coronary revascularization strategies. Presented at the Trans Catheter Therapeutics, San Diego, CA.

Davies J. (2018b). Machine-learning algorithm improves iFR measurement quality. Presented at the Trans Catheter Therapeutics, San Diego, CA.

Dennehy P, White MP, Hamilton A, et al. (2011). A partnership model for implementing electronic health records in resource-limited primary care settings: Experiences from two nurse-managed health centers. J Am Med Inform Assoc 18, 820–826.

Dimitrov DV. (2016). Medical internet of things and big data in healthcare. Healthc Inform Res 22, 156–163.

Dogan MV, Grumbach IM, Michaelson JJ, and Philibert RA. (2018). Integrated genetic and epigenetic prediction of coronary heart disease in the Framingham heart study. PLoS One 13, e0190549.

El Houby EMF. (2018). A survey on applying machine learning techniques for management of diseases. J Appl Biomed 16, 165–174.

Erdem C, Nagle AM, Casa AJ, et al. (2016). Proteomic screening and Lasso regression reveal differential signaling in insulin and insulin-like growth factor I (IGF1) pathways. Mol Cell Proteomics 15, 3045–3057.

Evans RS, Benuzillo J, Horne BD, et al. (2016). Automated identification and predictive tools to help identify high-risk heart failure patients: Pilot evaluation. J Am Med Inform Assoc 23, 872–878.

Feldreich T, Nowak C, Fall T, et al. (2019). Circulating proteins as predictors of cardiovascular mortality in end-stage renal disease. J Nephrol 32, 111–119.

Ford CM, Darlow K, Massie A, and Gorman DR. (2018). Using electronic maternity records to estimate female genital mutilation in Lothian from 2010 to 2013. Eur J Public Health 28, 657–661.

Fröhlich H, Balling R, Beerenwinkel N, et al. (2018). From hype to reality: Data science enabling personalized medicine. BMC Med 16, 150.

Gallo AM, Wilkie D, Suarez M, et al. (2010). Reproductive decisions in people with sickle cell disease or sickle cell trait. West J Nurs Res 32, 1073–1090.

Ganz P, Heidecker B, Hveem K, et al. (2016). Development and validation of a protein-based risk score for cardiovascular outcomes among patients with stable coronary heart disease. JAMA 315, 2532–2541.

Garg V. (2012). Molecular basis of cardiac development and congenital heart disease. In: *Translational Cardiology: Molecular Basis of Cardiac Metabolism, Cardiac Remodeling, Translational Therapies and Imaging Techniques*. Patterson C, and Willis MS, eds. Totowa, NJ: Humana Press, 317–339.

Giampaoli S, Palmieri L, Donfrancesco C, et al. (2006). Assessment of the absolute global cardiovascular risk: Comparison between the risk chart and the individual score of the CUORE project [in Italian]. G Ital Cardiol (Rome) 7, 359–364.

Gilboa SM, Salemi JL, Nembhard WN, Fixler DE, and Correa A. (2010). Mortality resulting from congenital heart disease among children and adults in the United States, 1999 to 2006. Circulation 122, 2254–2263.

Golas SB, Shibahara T, Agboola S, et al. (2018). A machine learning model to predict the risk of 30-day readmissions in patients with heart failure: A retrospective analysis of electronic medical records data. BMC Med Inform Decis Mak 18, 44.

Goyal A, Kahlon P, Jain D, et al. (2017). Trend in prevalence of coronary artery disease and risk factors over two decades in rural Punjab. Heart Asia 9, e010938.

Goyal NK, Hall ES, Greenberg JM, and Kelly EA. (2015). Risk prediction for adverse pregnancy outcomes in a Medicaid population. J Womens Health (Larchmt) 24, 681–688.

Grassi M, Perna G, Caldirola D, et al. (2018). A clinically-translatable machine learning algorithm for the prediction of Alzheimer's disease conversion in individuals with mild and premild cognitive impairment. J Alzheimers Dis 61, 1555–1573.

Hoffman J. (2013). The global burden of congenital heart disease. Cardiovasc J Afr 24, 141–145.

Huang JB, Liu YL, Sun PW, et al. (2010). Molecular mechanisms of congenital heart disease. Cardiovasc Pathol 19, e183–e193.

Huang T, Gong H, Yang C, and He Z. (2013). ProteinLasso: A Lasso regression approach to protein inference problem in shotgun proteomics. Comput Biol Chem 43, 46–54.

Itchhaporia D, Snow PB, Almassy RJ, and Oetgen WJ. (1996). Artificial neural networks: Current status in cardiovascular medicine. J Am Coll Cardiol 28, 515–521.

Jawhari B, Ludwick D, Keenan L, Zakus D, and Hayward R. (2016). Benefits and challenges of emr implementations in low resource settings: A state-of-the-art review. BMC Med Inform Decis Mak 16, 116.

Jia Y, Louw JJ, Breckpot J, et al. (2015). The diagnostic value of next generation sequencing in familial nonsyndromic congenital heart defects. Am J Med Genet A 167, 1822–1829.

Jiang F, Jiang Y, Zhi H, et al. (2017). Artificial intelligence in healthcare: Past, present and future. Stroke Vasc Neurol 2, 230–243.

Jung EY, Park JW, Ryu A, et al. (2016). Prediction of impending preterm delivery based on sonographic cervical length and different cytokine levels in cervicovaginal fluid in preterm labor. J Obstet Gynaecol Res 42, 158–165.

Katurura MC, and Cilliers L. (2018). Electronic health record system in the public health care sector of South Africa: A systematic literature review. Afr J Prim Health Care Fam Med 10, e1–e8.

Kim J, Lee J, and Lee Y. (2015). Data-mining-based coronary heart disease risk prediction model using fuzzy logic and decision tree. Healthc Inform Res 21, 167–174.

Knowles RL, Ridout D, Crowe S, et al. (2016). Ethnic and socioeconomic variation in incidence of congenital heart defects. Arch Dis Child 102, 496–502.

Kocbek S, Kocbek P, Stozer A, et al. (2018). Building interpretable models for polypharmacy prediction in older chronic patients based on drug prescription records. PeerJ 6, e5765.

Kuo P-J, Wu S-C, Chien P-C, et al. (2018). Artificial neural network approach to predict surgical site infection after freeflap reconstruction in patients receiving surgery for head and neck cancer. Oncotarget 9, 13768–13782.

LaHaye S, Corsmeier D, Basu M, et al. (2016). Utilization of whole exome sequencing to identify causative mutations in familial congenital heart disease. Circ Cardiovasc Genet 9, 320–329.

Lam MPY, Lau E, Ng DCM, Wang D, and Ping P. (2016). Cardiovascular proteomics in the era of big data: Experimental and computational advances. Clin Proteomics 13, 23.

Lee AJT, Liu Y-H, Tsai H-M, Lin H-H, and Wu H-W. (2009). Mining frequent patterns in image databases with 9D-SPA representation. J Syst Softw 82, 603–618.

Lee J-H, Kim D-H, Jeong S-N, and Choi S-H. (2018). Diagnosis and prediction of periodontally compromised teeth using a deep learning-based convolutional neural network algorithm. J Periodontal Implant Sci 48, 114–123.

Lemesle G, Maury F, Beseme O, et al. (2015). Multimarker proteomic profiling for the prediction of cardiovascular mortality in patients with chronic heart failure. PLoS One 10, e0119265.

Li H, Luo M, Zheng J, et al. (2017). An artificial neural network prediction model of congenital heart disease based on risk factors: A hospital-based case-control study. Medicine (Baltimore) 96, e6090.

Li X, Wu LJ, Gu M, et al. (2016). Peptidomic analysis of amniotic fluid for identification of putative bioactive peptides in ventricular septal defect. Cell Physiol Biochem 38, 1999–2014.

Liu JJ, Fan LL, Chen JL, Tan ZP, and Yang YF. (2014). A novel variant in TBX20 (p.D176n) identified by whole-exome sequencing in combination with a congenital heart disease related gene filter is associated with familial atrial septal defect. J Zhejiang Univ Sci B 15, 830–837.

Looker HC, Colombo M, Agakov F, et al. (2015). Protein biomarkers for the prediction of cardiovascular disease in type 2 diabetes. Diabetologia 58, 1363–1371.

Lu T-C, Fu C-M, Ma MH-M, Fang C-C, and Turner AM. (2016). Healthcare applications of smart watches: A systematic review. Appl Clin Inform 7, 850–869.

Lucas G, Lluís-Ganella C, Subirana I, et al. (2012). Hypothesis-based analysis of gene-gene interactions and risk of myocardial infarction. PLoS One 7, e41730.

Lyon A, Mincholé A, Martínez JP, Laguna P, and Rodriguez B. (2018). Computational techniques for ECG analysis and interpretation in light of their contribution to medical advances. J R Soc Interface 15, 20170821.

Madhukar NS, and Elemento O. (2018). Bioinformatics approaches to predict drug responses from genomic sequencing. Methods Mol Biol 1711, 277–296.

Mantzaris DH, Anastassopoulos GC, and Lymberopoulos DK. (2008). Medical disease prediction using artificial neural networks. Presented at the 2008 8th IEEE International Conference on BioInformatics and BioEngineering, 1–6. Athens, Greece.

Mesko B. (2017). The role of artificial intelligence in precision medicine. Exp Rev Precis Med Drug Dev 2, 239–241.

Min X, Yu B, and Wang F. (2019). Predictive modeling of the hospital readmission risk from patients' claims data using machine learning: A case study on COPD. Sci Rep 9, 2362.

Moonesinghe SR, Mythen MG, Das P, Rowan KM, and Grocott MPW. (2013). Risk stratification tools for predicting morbidity and mortality in adult patients undergoing major surgery: Qualitative systematic review. Anesthesiology 119, 959–981.

Moradian M. (2012). Diagnostic errors in echocardiography: Review of five interesting pediatric cases. J Tehran Heart Cent 7, 33–36.

Musameh MD, Wang WYS, Nelson CP, et al. (2015). Analysis of gene-gene interactions among common variants in candidate cardiovascular genes in coronary artery disease. PLoS One 10, e0117684.

Narula S, Shameer K, Salem Omar AM, Dudley JT, and Sengupta PP. (2016). Machine-learning algorithms to automate morphological and functional assessments in 2D echocardiography. J Am Coll Cardiol 68, 2287–2295.

Ng K, Steinhubl SR, Defilippi C, Dey S, and Stewart WF. (2016). Early detection of heart failure using electronic health records: Practical implications for time before diagnosis, data diversity, data quantity, and data density. Circ Cardiovasc Qual Outcomes 9, 649–658.

Nowak C, Carlsson AC, Ostgren CJ, et al. (2018). Multiplex proteomics for prediction of major cardiovascular events in type 2 diabetes. Diabetologia 61, 1748–1757.

Ogburn WF. (1922). *Social Change with Respect to Culture and Original Nature*. New York, NY: Dell.

Olive MK, and Owens GE. (2018). Current monitoring and innovative predictive modeling to improve care in the pediatric cardiac intensive care unit. Transl Pediatr 7, 120–128.

Ouyang N, Ikeda M, and Yamauchi K. (1997). Use of an artificial neural network to analyse an ECG with QS complex in V1–V2 leads. Med Biol Eng Comput 35, 556–560.

Ozdemir V, and Hekim N. (2018). Birth of industry 5.0: Making sense of big data with artificial intelligence, ''the internet of things'' and next-generation technology policy. OMICS 22, 65–76.

Paige E, Barrett J, Stevens D, et al. (2018). Landmark models for optimizing the use of repeated measurements of risk factors in electronic health records to predict future disease risk. Am J Epidemiol 187, 1530–1538.

Panahiazar M, Taslimitehrani V, Pereira N, and Pathak J. (2015). Using EHRS and machine learning for heart failure survival analysis. Stud Health Technol Inform 216, 40–44.

Purusothaman G, and Krishnakumari P. (2015). A survey of data mining techniques on risk prediction: Heart disease. Indian J Sci Technol 8, 1–5.

Qu Y, Liu X, Zhuang J, et al. (2016). Incidence of congenital heart disease: The 9-year experience of the Guangdong registry of congenital heart disease, China. PLoS One 11, e0159257.

Raggi P, Daniels M, Shanoudy H, and Jarmukli NF. (1996). MRI misinterpretation of spontaneous echo-contrast as a large left atrial thrombus. Int J Card Imaging 12, 85–88.

Raposo VL. (2015). Electronic health records: Is it a risk worth taking in healthcare delivery? GMS Health Technol Assess 11, Doc02.

Reeder B, and David A. (2016). Health at hand: A systematic review of smart watch uses for health and wellness. J Biomed Inform 63, 269–276.

Saini I, Singh D, and Khosla A. (2013). QRS detection using K-nearest neighbor algorithm (KNN) and evaluation on standard ECG databases. J Adv Res 4, 331–344.

Samad MD, Ulloa A, Wehner GJ, et al. (2018). Predicting survival from large echocardiography and electronic health record datasets: Optimization with machine learning. JACC Cardiovasc Imaging 12, 681–689.

Sanchis-Gomar F, Perez-Quilis C, Leischik R, and Lucia A. (2016). Epidemiology of coronary heart disease and acute coronary syndrome. Ann Transl Med 4, 256.

Sander TL, Klinkner DB, Tomita-Mitchell A, and Mitchell ME. (2006). Molecular and cellular basis of congenital heart disease. Pediatr Clin North Am 53, 989–1009, x.

Seravalli V, Block-Abraham DM, Turan OM, et al. (2014). Second-trimester prediction of delivery of a small-for-gestational-age neonate: Integrating sequential Doppler information, fetal biometry, and maternal characteristics. Prenat Diagn 34, 1037–1043.
Song X, Zhang Z, Zhang R, et al. (2018). Predictive markers of depression in hypertension. Medicine (Baltimore) 97, e11768.
Stevens SL, Mcmanus RJ, and Stevens RJ. (2018). The utility of long-term blood pressure variability for cardiovascular risk prediction in primary care. J Hypertens 37, 522–529.
Studziński K, Tomasik T, Krzyszton J, Jóźwiak J, and Windak A. (2017). Effect of using cardiovascular risk scoring in routine risk assessment in primary prevention of cardiovascular disease: Protocol for an overview of systematic reviews. BMJ Open 7, e014206.
Sun P-F, Ding G-C, Zhang M-Y, et al. (2017). Prevalence of congenital heart disease among infants from 2012 to 2014 in Langfang, China. Chin Med J 130, 1069–1073.
Taslimitehrani V, Dong G, Pereira NL, Panahiazar M, and Pathak J. (2016). Developing EHR-driven heart failure risk prediction models using CPXR(log) with the probabilistic loss function. J Biomed Inform 60, 260–269.
Thiyagaraja SR, Dantu R, Shrestha PL, et al. (2018). A novel heart-mobile interface for detection and classification of heart sounds. Biomed Signal Process Control 45, 313–324.
Topol EJ. (2019). High-performance medicine: The convergence of human and artificial intelligence. Nat Med 25, 44–56.
Trakadis YJ, Sardaar S, Chen A, Fulginiti V, and Krishnan A. (2018). Machine learning in schizophrenia genomics, a case-control study using 5,090 exomes. Am J Med Genet B Neuropsychiatr Genet 180, 103–112.
Triedman JK, and Newburger JW. (2016). Trends in congenital heart disease. Circulation 133, 716–2733.
Tsang W, Salgo IS, Medvedofsky D, et al. (2016). Transthoracic 3D echocardiographic left heart chamber quantification using an automated adaptive analytics algorithm. JACC Cardiovasc Imaging 9, 769–782.
Vigneron V, Kodewitz A, Tome AM, Lelandais S, and Lang E. (2016). Alzheimer's disease brain areas: The machine learning support for blind localization. Curr Alzheimer Res 13, 498–508.
Wang E, Sun S, Qiao B, et al. (2013). Identification of functional mutations in GATA4 in patients with congenital heart disease. PLoS One 8, e62138.
Wang T, Qiu RG, and Yu M. (2018). Predictive modeling of the progression of Alzheimer's disease with recurrent neural networks. Sci Rep 8, 9161.
Weng C-H, Huang TC-K, and Han R-P. (2016). Disease prediction with different types of neural network classifiers. Telemat Inform 33, 277–292.
Williams WG, and Mccrindle BW. (2002). Practical experience with databases for congenital heart disease: A registry versus an academic database. Semin Thorac Cardiovasc Surg Pediatr Card Surg Annu 5, 132–142.
Witter TE, Brimmell RR, and Qureshi SA. (2013). Learning from Bristol—National database for congenital heart disease in India? Ann Pediatr Cardiol 6, 3–5.
Wu Y, Fan J, Peissig P, et al. (2018). Quantifying predictive capability of electronic health records for the most harmful breast cancer. Proc SPIE Int Soc Opt Engineering 10577, 105770J.
Yeh S-J, Chen H-C, Lu C-W, et al. (2013). Prevalence, mortality, and the disease burden of pediatric congenital heart disease in Taiwan. Pediatr Neonatol 54, 113–118.
Zhang J, Gajjala S, Agrawal P, et al. (2018). Fully automated echocardiogram interpretation in clinical practice. Circulation 138, 1623–1635.

Address correspondence to:
Nicholas Ekow Thomford, PhD
Division of Human Genetics
Department of Pathology
Faculty of Health Sciences
University of Cape Town
Anzio Road, Observatory 7925
Cape Town, South Africa

E-mail: nicholas.thomford@uct.ac.za
nthomford@ucc.edu.gh

Christian Domilongo Bope, PhD
Division of Human Genetics
Department of Pathology
Faculty of Health Sciences
University of Cape Town
Anzio Road, Observatory 7925
Cape Town, South Africa

E-mail: christian.bope@uct.ac.za
christian.bope@unikin.ac.cd

Abbreviations Used

AD = Alzheimer's disease
AI = artificial intelligence
ANN = artificial neural network
ASD = atrial septal defects
CAD = coronary artery disease
CHD = congenital heart defects
CNN = Convolutional Neural Network
CT = computed tomography
DM = diabetes mellitus
ECGs = electrocardiographs
ECHO = Echocardiographic
eGFR = estimated glomerular filtration rate
EHRs = electronic health records
FH = family history
HF = heart failure
MAO = monoamine oxidase
MIMICK = mapping of inflammatory markers in chronic kidney disease study
ML = machine learning
MRI = magnetic resonance imaging
NGS = next-generation sequencing
PDA = patent ductus arteriosus
PE = pre-eclampsia
PNNs = Probabilistic Neural Networks
RF = random forest
SCZ = schizophrenia
SKS = Salford Kidney Study
SSA = sub-Saharan Africa
SVM = support vector machine
TC = total cholesterol
VSD = ventricular septal defect
WES = whole exome sequencing