The Year in Cardiovascular Medicine 2020: Digital Health and Innovation

Charalambos Antoniades1,2, Folkert W. Asselbergs3,4 and Panos Vardas5,6
1 Acute Vascular Imaging Centre, University of Oxford, UK
2 Division of Cardiovascular Medicine, Radcliffe Department of Medicine, University of Oxford
3 Department of Cardiology, Division Heart & Lungs, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands;
4 Institute of Cardiovascular Science and Institute of Health Informatics, Faculty of Population Health Sciences, University College London, London, United Kingdom;
5 Heart sector, Hygeia hospitals groups, Athens, Greece
6 Medical school, University of Crete, Heraclion, Greece

Acknowledgement: CA is supported by the British Heart Foundation ((FS/16/15/32047 and TG/19/2/34831), the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre and Innovate UK. FA is supported by UCL Hospitals NIHR Biomedical Research Centre. The authors are grateful to Dr Christos Kotanidis for his technical contribution in drafting the graphical abstract.

Word count: 3543, References: 46

Corresponding author:
Professor of Cardiovascular Medicine
Director, Acute Vascular Imaging Centre
Deputy Head, Division of Cardiovascular Medicine
University of Oxford
L6 West Wing, John Radcliffe Hospital, Headley way, Oxford OX37TH, UK
Email: Charalambos.antoniades@cardiov.ox.ac.uk
Tel: +441865228340
Abstract

Advances in digital health and particularly in artificial intelligence (AI) have led us very close to the true implementation of personalised medicine. The year 2020 has brought an exponential increase of studies using various forms of artificial intelligence, from supervised machine learning to unsupervised deep learning, with applications across all domains of Cardiovascular Medicine. AI is now moving from research to implementation, affecting all aspects of clinical Cardiology. The studies bringing AI close to clinical practice span from fast clinical and biochemical data analysis and interpretation of results, to image analysis, ECG interpretation, arrhythmia detection, or even the use of face recognition to diagnose cardiovascular diseases. We review some of the most exciting development if the field of AI in Cardiology, published from fall of 2019 up until now. The studies highlighted in this article give only a small glimpse into this booming field, creating more anticipation for what will come to clinical practice in the coming years.
Preamble

Digital Health and particularly Artificial intelligence (AI), are getting fast ground during the last few years in cardiovascular diagnostics and therapeutics. Indeed, the number of publications using various AI techniques has been increased by >20 fold from 2010 to 2020. Since last year’s ESC congress, the role of artificial intelligence in Cardiovascular Medicine had been highlighted as the next frontier in cardiovascular diagnostics, paving the way to the implementation of personalised strategies in cardiovascular therapeutics.¹ In a similar line, the last American Heart Association 2020 meeting, also had a session entitled “Hype or Hope? Artificial Intelligence and Machine Learning in Imaging, reminding us the importance paid by the major clinical cardiovascular medicine societies in this field. Indeed, issues like algorithm transparency, data open access transparency, were key issues introduced. The concept of using digital innovation and particularly AI and Big Data to optimise treatments in clinical trials and eventually in clinical practice was brought up as a fundamental aspect of digital health of the future.

The introduction of AI in research but also in clinical practice, is mainly driven by the technological advances in the handling and analysis of big data. AI is referred to the ability of a machine to execute tasks characteristic of human intelligence, such as problem-solving or pattern recognition, and it is typically characterised by the element of positive or negative reinforcement as part of the learning process, similar to what typically happens with human learning. Indeed, machine learning refers to the ability of computers to improve their knowledge without being explicitly programmed to do so; so the machines can identify patterns in digital data, and make generalizations, learning from their observations.² Unsupervised deep learning is used to build convolutional Neuronal Networks (CNNs) that recognise features in digital data, not
visible to the human eye. These data can be clinical information, images, ECGs or even standard “selfies” taken using smartphone cameras.

**AI As a tool for arrhythmias prediction and management**

Management of arrhythmias has always been a challenge, especially when we have to deal with subclinical conditions such as paroxysmal atrial fibrillation, who often have stroke as their first presentation. Indeed, including clinical risk factors into a machine learning algorithm was recently found to identify patients at risk for atrial fibrillation in a primary care population of >600k individuals in the DISCOVER registry in the UK. That algorithm could achieve negative predictive value of 96.7% and sensitivity to detect atrial fibrillation of 91.8%. In another landmark study published by the Mayo Clinic last year, it seems now possible that by using a convolutional neural network (CNN) to screen standard 12 lead ECGs for characteristics not visible to the eye of the clinician, we can detect subclinical paroxysmal atrial fibrillation from sinus rhythm ECGs, achieving AUC as high as 0.9. This study was conducted in a population of >180k individuals with >450k ECGs included in the training set, >64k ECGs in the internal validation dataset and >130k in the testing dataset. Algorithms like this could completely transform population screening for atrial fibrillation, and will most likely enable timely administration of anticoagulant treatment to prevent cardioembolic stroke. The astonishing size of this dataset, gives a clear example of how deep learning should be performed, to yield reproducible, practice-changing tools. Algorithms like this, will soon be available on our portable ECGs in the clinic. One of the major problems of deep learning algorithms used for ECG interpretation, is their susceptibility to adversarial examples, leading to consistently wrong classification of the test by detecting false
patterns undetectable to the human eye. An elegant study by Hat et al\(^5\) has provided recently the tools needed to study the impact of these adversarial patterns in automated ECG classification, and provides new opportunities to develop appropriate mitigation measures.

The recent release of large, publically available ECG databases such as the PTB-XL (that includes ~21k records from ~19k patients)\(^6\) or the one from the Shaoxing Hospital Zhejiang University School of Medicine (~10k patients)\(^7\) brings further optimism that by increasing the variability and ethnic diversity of the training and validation datasets, these ECG applications are not far from clinical implementation. Further to the use of AI to detect atrial fibrillation, a recent study built a deep neural network to classify various types of ECGs using >2.3m ECGs from >1.6m patients, demonstrating a remarkable ability of these networks to provide accurate interpretation of these tests.\(^8\) Finally, in the Apple Heart Study\(^9\) the use of smartphones was demonstrated to be a very effective way to detect patients with subclinical paroxysmal atrial fibrillation. That was a large study that included ~420k participants followed up for a median of 117 days through their smartphones. The technology developed by apple, identified 0.5% with potentially irregular pulse (34% of which were proven to have atrial fibrillation confirmed by ECG). Although the exact nature of the technology used in the smartphone is not available, this study demonstrates that large volumes of data can be collected even using standard smartphones or portable devices like apple watches, opening new opportunities for big data research and development of AI algorithms for timely detection of cardiac arrhythmias in asymptomatic individuals.

**AI for the management of heart failure**
Risk stratification plays a key role in designing the therapeutic strategies in heart failure, given that the expected survival inevitably affects the decision for device implantation. Currently, decision of implanting a defibrillator and/or applying resynchronization therapy (CRT) in patients with heart failure, relies on well-defined clinical, electrophysiological and imaging characteristics. However, a recent study published in EHJ earlier this year, came to remind us that the prediction of responsiveness to CRT focused on mid- or long- term outcomes should be a key driver in decision making. Indeed, Tokodi et al used machine learning to help them build a risk score for prediction of mortality following CRT. The score used information from medical history, physical examination, medication records, ECG, echocardiographic and laboratory data commonly obtained as part of routine hospital visits of patients with heart failure, and after it was trained in 1510 patients using a random forest algorithm, it has achieved a remarkable prognostic value for all-cause mortality, with AUC in ROC analysis ranging from 0.77 (in 1 year prediction) to 0.8 (in 5 year prediction). The risk calculator is now available for use (SEMMELWEIS-CRT Score, https://arguscognitive.com/crt, Figure 1). As the authors mention, this score could facilitate the prompt recognition of high-risk patients, guiding deployment of the appropriate prophylactic measures. It could also assist the patients and the families in making advance care decisions, while it could assist clinicians in deciding which patients are most suitable for CRT.

The results of that study were in line with another recent study showing that the use of machine learning to integrate clinical data together with imaging characteristics can provide meaningful information about the future responsiveness of heart failure patients to CRT in a population of >1.1k patients from the MADIT-CRT study.
CRT would have meaningful impact in patient’s prognosis. Before we reach at that stage though, it seems important to understand how to manage the high-risk individuals identified through such algorithms, given that most of the factors included into these models are non-modifiable.\textsuperscript{15} Randomised clinical trials are needed, to evaluate the clinical benefit and cost-effectiveness of applying such algorithms in clinical practice.

The last year brought also new advances in the use of artificial intelligence for the diagnosis of heart failure. Indeed, a CNN was trained based on paired ECGs and transthoracic echocardiograms from ~45k patients and validated in an independent cohort of >52k patients.\textsuperscript{16} The ROC for detection of systolic dysfunction using this AI-enhanced ECG interpretation reached an AUC of 0.93. This impressive result confirms the notion that AI could extract invaluable information even from simple, low-cost tests like ECG, which could even be used as screening tests for detection of subclinical heart failure in the community.

**AI in cardiac imaging**

This year has been extraordinary for medical imaging, as a wide range of AI-powered algorithms have been introduced in clinical care by the hardware vendors and software manufacturers. These algorithms range from image reconstruction, to automated segmentation and improvement of workflows, or even to detection of imaging characteristics not visible to the human eye assisting diagnosis.\textsuperscript{17, 18} 2020 is considered by many as the year of cardiac computed tomography angiography (CTA), as this has just been incorporated into the recent ESC guidelines as a first line investigation for the management of chest pain.\textsuperscript{19} This approach came a few years after a similar recommendation was published in the UK NICE guidelines, but it is still
more advanced compared to the US standard of care.\textsuperscript{20} Given the standardised way by which CT images are captured, the modality is particularly attractive to machine-learning methods to improve segmentation and interpretation. Indeed, in a study by Al’Aref et al\textsuperscript{21} from the CONFIRM registry, a population of >13k patients undergoing coronary calcium score measurements (CCS) was used to examine whether including CCS in a machine learning model together with clinical risk factors, could improve risk stratification. Indeed, adding CCS in a baseline model that included clinical risk factors, resulted in \textasciitilde 9% improvement in the ability to estimate the pretest probability of obstructive coronary artery disease, with remarkable diagnostic accuracy. Particularly in the young patients (<65 years old), the algorithm improved the ability to detect coronary artery disease by \textasciitilde 17%. However, it remains to be proven that machine learning performs better than simple statistical regression models, when risk factors are combined with results from tests like CCS.\textsuperscript{22} Further to the use of machine learning to integrate imaging with other datasets, the practical value of AI lies with the improvement of image analysis workflows.\textsuperscript{18} Automated segmentation of coronary atherosclerotic plaques, coronary calcification or even epicardial fat in CT, makes image interpretation faster, more accurate and eliminates user-dependent variability.\textsuperscript{23}

The true power of AI though comes from its ability to “see the invisible”. The field of radiomics allows extraction of thousands of different pieces of information from images, which provide information on the texture and composition of the tissue visualised (Figure 2). Indeed, further to the analysis of the composition and volume of coronary atherosclerotic plaques, it is now widely accepted that vascular inflammation causes changes in the composition and texture of perivascular fat, which activates lipolysis and increases its hydrophilic content around inflamed vascular structures.\textsuperscript{24} Visualising perivascular fat using standard CTA, allows the calculation of a metric of
these changes, driven by the 3D changes of perivascular fat attenuation. An AI-derived biomarker that captures that biology, the fat attenuation index (FAI), has striking prognostic value\(^{25}\) that goes beyond atherosclerotic plaque characteristics,\(^{26}\) as demonstrated in \(~4,000\) patients from the CRISP-CT outcomes study. In a recent paper published by Oikonomou et al. published in EHJ\(^{27}\) the same principle, i.e. the ability of perivascular fat to change its texture and composition in response to inflammatory signals coming from the vascular wall, was transferred in the field of radiomics. The concept of radiotranscriptomics has been introduced in the cardiovascular dictionary, as by using the gene expression profile of adipose tissue in fat biopsies obtained from 167 patients undergoing cardiac surgery, they created molecular classifiers for inflammation, fibrosis and angiogenesis, all features characterising perivascular fat after prolonged exposure to vascular inflammation. Then they extracted the radiomic features from the CT images of the same adipose tissue, and by using machine learning they built a radiomic signature to detect chronic vascular inflammation (capturing perivascular fibrosis, angiogenesis and inflammation). The new radiotranscriptomic metric generated, the Fat Radiomic Profile (FRP) index (Figure 3), was then tested for its performance in 1575 patients from the SCOTHEART trial, who were followed up for 5 years after their CTA.\(^{26,28}\) Indeed, FRP had a remarkable prognostic value, as those people with abnormal FRP had \(>10\) times higher risk for a fatal or non-fatal cardiac event, with an AUC to detect those who will have the event, of 0.88. When abnormal FRP was combined with the presence of high risk plaque, the patient’s relative risk for cardiac event was \(>43\) times higher than the reference group (Figure 3). As it was discussed in the associated editorial,\(^{29}\) this technology could guide therapeutic interventions; this could be done in the future either in the form of a companion diagnostic to allow targeted deployment of expensive...
treatments, or as an enrichment tool for clinical trials. Other papers published this year\textsuperscript{30} seem to confirm the validity of this approach, while the strategies for Imaging residual inflammatory risk have been presented in a recent state of the art review published in EHJ.\textsuperscript{31} This method needs further validation in non-Caucasian ethnic groups, while it’s translation into a clinically applicable tool is challenging due to the complexity of the analysis which makes it difficult to perform on standard clinical workstations onsite.

From an ultrasound point of view, 2020 has been a year for consolidation of earlier technical developments in how to train neural networks to handle raw images and video loops from echocardiograms to segment and extract useful metrics such as ejection fraction and myocardial strain.\textsuperscript{32} The study by Ouyang is probably the most significant advancement in the field, driven by AI in 2020.\textsuperscript{33} In that study, they took echocardiography analysis from still frame segmentation to a video-based deep learning approach through development of a specific EchoNet-Dynamic algorithm that combines temporal and spatial information within the neural network. Training networks for evaluation of segmentation and quantification, achieves an acceptable accuracy for estimation of ejection fraction on a beat to beat basis that can help in identification of heart failure. The next interesting phase is going to be the application of these approaches to larger scale datasets to improve accuracy of disease prediction, in large databases like the UK Biobank.\textsuperscript{34} This approach could open new horizons in applying deep learning in image interpretation for risk prediction. Indeed, this year the outputs of the UK-Biobank confirmed its potential to drive innovation for years to come. In a study just published\textsuperscript{35}, more that 26k cardiac MRI scans were used in machine learning algorithms to allow detection of >2k interactions between imaging phenotypes and non-imaging phenotypes in the UK-Biobank, providing new insights
into the influence of early-life factors and diabetes on cardiac and aortic structure and function, linking them also with cognitive phenotypes.

**AI and COVID-19**

Above all, 2020 will be remembered as the year when COVID-19 brought the world upside down.\(^{36}\) As our knowledge accumulates about the disease, it becomes clear that COVID-19 is, in the end, a vascular endothelial disease\(^{37, 38}\) The need for rapid integration of large volumes of data collected from around the world to facilitate the urgent development of treatments to combat the disease, brought to the surface the power of AI to give solutions fast and accurately.\(^{39}\) Indeed, fast and accurate data collection has been in the centre of the efforts to combat the disease. European registries like the CAPACITY-COVID,\(^{40}\) are actively collecting data around the disease, working together with international efforts from the International Severe Acute Respiratory and Emerging Infection Consortium (ISARIC) and World Health Organization (WHO). The use of AI to interrogate these datasets is expected to improve our understanding on the incidence and pattern of cardiovascular complications in patients with COVID-19 and evaluate the vulnerability and clinical course of patients with underlying cardiovascular diseases. In addition, AI algorithms have been used to integrate chest CT findings with clinical symptoms, laboratory testing and exposure history to rapidly diagnose COVID-19. In a very recent study\(^{41}\) that included 905 patients tested with (419 of which tested positive for SARS-CoV-2), the AI system achieved an area under the curve of 0.92 to diagnose the disease without the need of a PCR method, having sensitivity comparable to a senior thoracic radiologist.\(^{41}\) The use of computational learning methods to integrate biomarkers of inflammation and myocardial injury (e.g. C-reactive protein (CRP), N-terminus pro B type natriuretic peptide (NT-proBNP), myoglobin (MYO), D-dimer, procalcitonin (PCT),
creatine kinase-myocardial band (CK-MB) and cardiac troponin I (cTnI) in COVID-19, was recently found to predict mortality with AUC 0.94.\textsuperscript{42} These initial models could lead to point-of-care Severity Score systems, and could have major impact in clinical decision making, in the coming months. In the post-COVID-19 period, the expertise gained in applying machine learning to integrate multi-omic and clinical data,\textsuperscript{43} is expected to revolutionise cardiac diagnostics.

**AI: From pattern recognition to analysis of “selfies”!**

AI and particular convolutional neural networks are being accused as “black boxes”, that combine features which are individually meaningless, into algorithms that give meaningful predictions. Indeed, in a landmark study just published in the EHJ,\textsuperscript{44} a Chinese group of scientist has developed a deep convolutional neural network that detects coronary artery disease (with stenosis >50% documented by angiography), by analysing the patient’s facial photos (Figure 4). They included >5k patients in their training dataset and 580 in the test dataset. The algorithm had sensitivity 80% and specificity 54% to detect significant coronary artery disease from the faces of the patients, with an AUC 0.73. Could this be demonstrating genetic predisposition to atherosclerosis? Could it demonstrate secondary effects on the skin and structure of the face due to risk factors or the disease itself? Or is it just the result of training the algorithm in an ethnically homogenous population, that will not survive the test of time?\textsuperscript{45} If the concept behind this study is confirmed, then medical confidentiality may be at risk; walking into a train station or walking through the doors of an insurance company (where CCTV is in operation), may already give away health problems that you would like to keep private (breaching individual confidentiality), or inform you about health issues you are not aware of (saving your life). These issues will definitely spark extensive debates in the coming years.
Challenges of AI application in clinical practice

Further to the great opportunities presented by AI, these technologies generate also significant scepticism. The results generated by most machine learning algorithms often fail to generalise in different populations. Since these algorithms are often in the form of a “black box”, it is hard to understand (and therefore criticise and edit) their content, and this generates unavoidable bias. Such bias could lead to results applicable only to specific populations, specific technical equipment or specific clinical practices included into the training datasets. Many deep learning algorithms are also susceptible to adversarial examples, leading to consistently wrong classification of the measured parameter(s) by detecting false patterns undetectable to the human eye.

The limited generalisability in machine learning (i.e. the poor adaptability of these models to previously unseen data), comes to limit the applicability of these algorithms to clinical practice. This issue is mitigated by applying beyond training and internal validation (which inevitably leads to overestimation of the model’s performance), also independent testing. For the proper generalisability assessment, independent test dataset should represent the population of interest, but in a dataset totally independent of the training dataset (typically from independent institutions and/or geographically distant populations). The training part should be used for dimension reduction, development of the model and for hyperparameter tuning (and can use methods like cross-validation, random sampling or nested cross-validation). To prevent bias in performance evaluation, the model should be locked before the independent testing. The lack of transparency on the true links between the training and the independent validation dataset often makes it hard to evaluate the quality of the published literature. Transparent reporting can be ensured by following specific
principles,\textsuperscript{46} while open data sharing would allow independent reproducibility tests, securing high standards in publishing in the field.

**Conclusions**

There is no doubt that 2020 has been an extraordinary year, dominated by the COVID-19 pandemic. Under these difficult circumstances for humanity, and with most areas of Cardiovascular research compromised due to national lockdowns, the data science endured. The ability of artificial intelligence to extract and analyse large volumes of data remotely, allowed this field of Cardiovascular Medicine to continue its evolution, and we have seen major discoveries transforming many aspects of clinical care. From workflow improvements to automated image segmentation, accurate cardiovascular risk prediction or event facial recognition to screen for cardiac diseases, artificial intelligence is now major part of Cardiovascular Medicine. The studies highlighted in this article give only a small glimpse into this booming field, creating more anticipation for what will come to clinical practice in the coming years.

The European Society of Cardiology has early recognised the importance of the fast evolving field of digital health technologies and has prioritized it as a strategic domain of cardiovascular medicine. The European Heart Journal family is at the forefront of the international effort to set high standards in publishing AI studies, actively promoting translation of AI technologies into clinical applications. A new section on digital health has recently been included in the EHJ, aiming to cultivate the culture of digitization in the full spectrum of cardiovascular medicine. In addition, a new journal (EHJ Digital health) has been added into the EHJ family. Finally, the European Union has recently launched an effort to regulate the use of AI algorithms as medical devices, especially
for risk prediction. AI algorithms will need to receive CE mark as medical devices from May next year. This approach is being adopted by both, FDA and EMA, and will have direct implications on the clinical implementation of newly developed AI cardiovascular risk calculators that will be included in the clinical guidelines in the future.
The 12 most important predictors of all-cause mortality as assessed by the SEMMELWEIS-CRT score. The importance of each feature was quantified by calculating the decrease in the model's performance (area under the receiver operating characteristic curve) after permuting its values (permutation feature importance method). The higher its value, the more important the feature is. As the values of feature importance were spread over a wide range (more orders of magnitude), base-10 logarithmic transformation was performed to facilitate plotting. CRT, cardiac resynchronization therapy; LVEF, left ventricular ejection fraction; NYHA, New York Heart Failure Association functional class (From: Tokodi et al Eur Heart J 2020).
Artificial intelligence can be used to combine different types of information, from clinical and laboratory data, to imaging or any other type of information, to assist clinical diagnosis and decision making. (from Oikonomou et al Cardiovasc Res 2020).
Figure 3

(A) Coronary inflammation first drives changes in peri-coronary adipocyte size, while at a later stage it leads to perivascular fibrosis and angiogenesis. (B) These changes can be visualised in standard coronary CT angiography (CCTA) by a method called Fat Attenuation Indexing. By using a radiotranscriptomic approach, oikonomou et al (Eur Heart J 2019) have built an imaging signature that captures these changes (Fat Radiomic Profile or FRP). (C) That signature has striking prognostic value over and above risk factors including coronary calcium score (CCS), as it was validated in the SCOTHEART population. MACE: Majod Adverse Cardiac Events; FAI: Fat attenuation index; BMI: body mass index; CAD: coronary artery disease (from Bartelt et al; Eur Heart J 2019).
Areas in the face with information useful in face pattern recognition, involved in prediction of coronary artery disease. In tests occluding facial regions (A), mAUC was defined as the decrease in algorithm performance after occluding a specific facial region. In tests occluding regions of 11x11 pixels (B), the green regions were highlighted by the algorithm as important for detecting CAD. In the dose–response relationship test (C), the positive facial areas were judged based on the change in algorithm performance after occlusion. Having 7-9 positive areas was related with presence of coronary artery disease in 84% of the cases, while there was >1 vessel disease in ~42% of the cases (C) (From: Lin S et al, Eur Heart J 2020). This technology can be used for screening in the community (D) (from Kotanidis C et al; Eur Heart J 2020).
References


