ESC European Society of Cardiology

European Heart Journal - Digital Health (2021) **2**, 597–605
doi:10.1093/ehjdh/ztab084

ORIGINAL ARTICLE

Machine learning aids clinical decision-making in patients presenting with angina and non-obstructive coronary artery disease

Ali Ahmad [1†], Michal Shelly-Cohen[1†], Michel T. Corban [1], Dennis H. Murphree Jr [1], Takumi Toya [1,2], Jaskanwal D. Sara[1], Ilke Ozcan [1], Lilach O. Lerman [3], Paul A. Friedman[1], Zachi I. Attia[1], and Amir Lerman [1*]

[1]Department of Cardiovascular Medicine, Mayo Clinic, 200 First Street SW, Rochester, MN 55902, USA; [2]Department of Medicine, Division of Cardiology, National Defense Medical College, Tokorozawa, Saitama, Japan; and [3]Department of Medicine, Division of Nephrology and Hypertension, Mayo Clinic, 200 First Street SW, Rochester, MN 55902, USA

Received 28 June 2021; revised 14 September 2021 online publish-ahead-of-print 14 October 2021

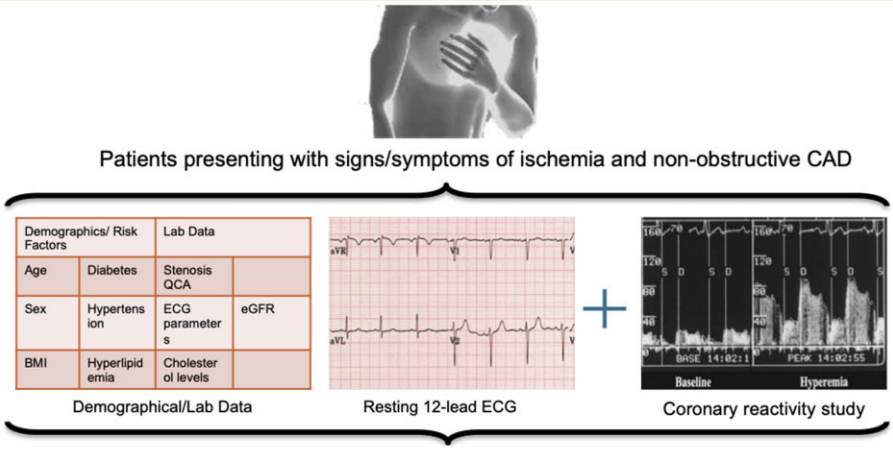
Aims The current gold standard comprehensive assessment of coronary microvascular dysfunction (CMD) is through a limited-access invasive catheterization lab procedure. We aimed to develop a point-of-care tool to assist clinical guidance in patients presenting with chest pain and/or an abnormal cardiac functional stress test and with non-obstructive coronary artery disease (NOCAD).

Methods and results This study included 1893 NOCAD patients (<50% angiographic stenosis) who underwent CMD evaluation as well as an electrocardiogram (ECG) up to 1-year prior. Endothelial-independent CMD was defined by coronary flow reserve (CFR) ≤2.5 in response to intracoronary adenosine. Endothelial-dependent CMD was defined by a maximal percent increase in coronary blood flow (%ΔCBF) ≤50% in response to intracoronary acetylcholine infusion. We trained algorithms to distinguish between the following outcomes: CFR ≤2.5, %ΔCBF ≤50, and the combination of both. Two classes of algorithms were trained, one depending on ECG waveforms as input, and another using tabular clinical data. Mean age was 51 ± 12 years and 66% were females ($n = 1257$). Area under the curve values ranged from 0.49 to 0.67 for all the outcomes. The best performance in our analysis was for the outcome CFR ≤2.5 with clinical variables. Area under the curve and accuracy were 0.67% and 60%. When decreasing the threshold of positivity, sensitivity and negative predictive value increased to 92% and 90%, respectively, while specificity and positive predictive value decreased to 25% and 29%, respectively.

Conclusion An artificial intelligence-enabled algorithm may be able to assist clinical guidance by ruling out CMD in patients presenting with chest pain and/or an abnormal functional stress test. This algorithm needs to be prospectively validated in different cohorts.

* Corresponding author. Tel: +1 507 255 4152, Fax: +1 507 255 7798, Email: lerman.amir@mayo.edu
†The first two authors contributed equally to the study.

Graphical Abstract



Coronary Microvascular Dysfunction

☐ Yes

☐ No

Outcome	Model	AUC (95% CI)	Sensitivity	Specificity	PPV	NPV
CFR ≤ 2.5	Demographics	0.67 (0.62, 0.73)	70	56	35	85
	AI-ECG	0.56 (0.51, 0.63)	43	57	26	75
	Dem + AI-ECG	0.64 (0.57,0.70)	51	59	57	53
ΔCBF (%) ≤ 50	Demographics	0.55 (0.50, 0.60)	51	59	57	53
	AI-ECG	0.51 (0.45,0.57)	42	60	54	48
	Dem + AI-ECG	0.53 (0.53,0.60)	58	62	72	46
ΔCBF (%) ≤ 50 and/or CFR ≤ 2.5	Demographics	0.61 (0.57, 0.66)	51	67	73	45
	AI-ECG	0.51 (0.45,0.57)	54	35	58	32
	Dem + AI-ECG	0.52 (0.54,0.59)	62	60	35	83

* Demographics include Age, sex, body mass index (BMI), smoking, diabetes mellitus, dyslipidemia, hypertension, cholesterol levels (Total cholesterol, low-density lipoprotein cholesterol (LDL), high-density lipoprotein cholesterol (HDL), triglycerides (TGs)), eGFR, and ECG parameters (QRS and QT intervals), and Stenosis by angiography (QCA).

Keywords AI • Machine learning • Coronary microvascular dysfunction • ECG • Coronary circulation

Introduction

Around two-thirds of patients presenting with angina and non-obstructive coronary artery disease (NOCAD) on clinically indicated coronary angiography have coronary microvascular dysfunction (CMD) detected with pharmacologic provocation testing.[1–3] Coronary microvascular dysfunction has been associated with atherosclerosis, myocardial ischaemia, heart failure with preserved ejection fraction, increased mortality, and a higher risk of major adverse cardiovascular events, including myocardial infarction, progressive congestive heart failure, atrial fibrillation, and sudden cardiac death.[4–12]

The gold standard method for the assessment of coronary microvascular and endothelial function involves a comprehensive formal invasive and expensive procedure. Multiple other non-invasive tests to assess the coronary microvasculature have been evaluated (e.g. echocardiography-derived or positron emission tomography-derived coronary flow reserve [CFR]). However, a weak correlation has been observed between non-invasive and invasive assessment of coronary reserve and vasomotion in several studies.[13–15]

Artificial intelligence (AI) is an increasingly recognized powerful tool to help equip clinicians in the decision-making process across multiple domains and subspecialties.[16,17] Artificial intelligence electrocardiogram (ECG) analysis allows clinicians to identify physiological ageing[18,19] and multiple cardiovascular diseases, such as paroxysmal atrial fibrillation, depressed left ventricular (LV) dysfunction, and hypertrophic cardiomyopathy, through a single resting 10-s 12-lead ECG.[20–22] In a previous study by our group, CMD was associated with minor ECG differences (QTc and T-waves),[23,24] which could potentially indicate that different ECG blueprints may be present in patients with CMD. Therefore, the current study was designed to test the hypothesis that AI can assist the clinical decision-making and identify the patients with high and low probability to have CMD and help physicians to decide whether to proceed with invasive diagnostic procedures in patients presenting with signs/symptoms of ischaemia.

Methods

Data sources and study population

This study included consecutive subjects with angina and NOCAD on coronary angiography (<50% stenosis in major vessels) who underwent a clinically indicated invasive coronary reactivity testing (CRT) for the evaluation of CMD, as well as resting 10-s 12-lead ECG up to 1 year before CRT.[3] Patients with acute coronary syndrome presentation and those with a history of myocardial infarction or cerebrovascular accident within the past 6 months, previous percutaneous coronary intervention or coronary artery bypass surgery, use of radiographic contrast agents within 12 h before catheterization, valvular heart disease, advanced chronic kidney disease, cardiomyopathy (LV ejection fraction <45%), active malignancy, local or systemic infectious disease within 4 weeks before catheterization, and inflammatory diseases were excluded. Pregnant patients and those unable to provide written informed consent were also excluded from this study.

Standard 10-s, 12-lead ECG acquired in the supine position at the Mayo Clinic ECG laboratory between 1992 and 2019 were included in this analysis. All ECGs were analysed at a sampling rate of 500 Hz using a GE-Marquette ECG machine (Marquette, WI, USA), ECGs that were originally sampled at 250 Hz were unsampled to 500 Hz using the 'Resample' function from SciPy python package.[25] The raw data were stored using the MUSE data management system. The study was compliant with the Declaration of Helsinki and approved by the Mayo Foundation Institutional Review Board.

Patient information and variable selection

Clinical history, laboratory data, and current medications were collected from a detailed chart review by an investigator blinded to functional angiography results from 1975 patients seen in Mayo Clinic as previously described.[3,26] Data were collected on conventional cardiovascular risk factors including age, hypertension (HTN), diabetes mellitus (DM), hyperlipidaemia (HLD), smoking status, and body mass index (BMI); biochemical parameters including serum total cholesterol (TCHOL) , low-density lipoprotein (LDL) cholesterol, high-density lipoprotein (HDL) cholesterol, and triglycerides (TGs). Smoking was defined as positive for exposure (current or former) or never. The estimated glomerular filtration rate (eGFR) was calculated using the CKD-EPI formula. All blood levels documented had been drawn up to 2 weeks before the index procedure. Dyslipidaemia was defined by a documented history of hyperlipidaemia, treatment with lipid-lowering therapy, an LDL cholesterol level above the target (<130 mg/dL for low-risk patients, <100 mg/dL for moderate-high-risk patients, <70 mg/dL for very high-risk patients, and <55 mg/dL for extreme high-risk patients based on 10-year atherosclerotic cardiovascular disease risk), HDL cholesterol <40 mg/dL in men or <50 mg/dL in women, or TGs >150 mg/dL. Type 2 diabetes mellitus was defined as a documented history of or treatment for type 2 diabetes, or an HbA1c of >6.5, if available. Hypertension was defined as a documented history of the disease or treatment.

Coronary reactivity testing

Patients discontinued vasodilatory medications (calcium channel blockers, beta-blockers, and long-acting nitrates) at least 24 h before the study. They were only allowed to take sublingual nitroglycerine tablets or spray for angina up to 6 h before the catheterization procedure. The Mayo Clinic protocol of CRT has been described previously in detail.[3,26–31] In brief, patients underwent diagnostic coronary angiography using standard clinical protocols. Those with NOCAD (no or mild angiographic stenosis <50% in any major vessel) went on to receive 5000 U of heparin intravenously, after which a Doppler guidewire (FloWire, Philips/Volcano Corp., San Diego, CA, USA) was positioned in the mid-left anterior descending coronary artery (LAD) along with an infusion catheter. First, to assess endothelium-independent vasodilation, intracoronary bolus injections of incremental doses (18–72 μg) of adenosine were administered through the guiding catheter until maximal hyperaemia was achieved. Coronary flow reserve was calculated as the ratio of hyperaemic over baseline blood velocities. Abnormal CFR was subsequently defined as CFR $\leq$2.5 in response to adenosine.[3] Next, coronary microvascular endothelial function was assessed using infusions of increasing concentrations at 1 mL/min of intracoronary acetylcholine (10^{-6}, 10^{-5}, and 10^{-4} mol/L for 3 min each). Doppler measurements of peak velocity were performed after each acetylcholine infusion, followed by repeat coronary angiography. The mid-LAD diameter was measured in the segment 5 mm distal to the tip of the Doppler wire, using a quantitative coronary angiography program (QAngio, Medis Corp, Leiden, Netherlands). Coronary blood flow (CBF) was then calculated using the formula: CBF = $\pi \times$ (peak velocity/2) $\times$ (coronary artery diameter/2)^2, as previously described.[27,28,31,32] The maximal percent change in CBF in response to acetylcholine compared to baseline (%ΔCBF) was then calculated, and abnormal response was defined as %ΔCBF $\leq$50%.[3,31,33]

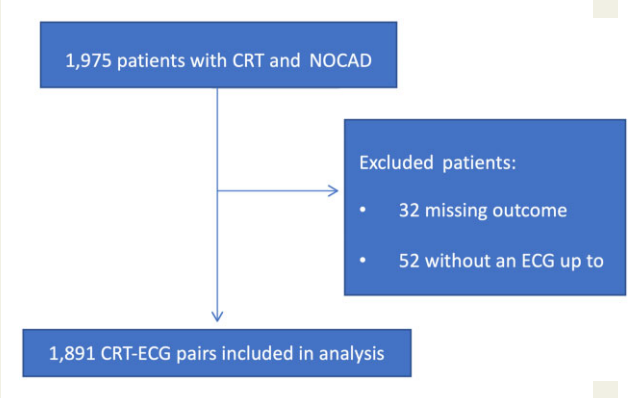


Figure 1 Patients included in the study. CRT, coronary reactivity testing; ECG, electrocardiogram; NOCAD, non-obstructive coronary artery disease

Overview of the artificial intelligence model

Two primary modelling frameworks were developed: a traditional machine learning (ML) framework based on the tabular clinical data and a deep learning (DL) framework using convolutional neural networks (CNN) applied on both the ECG waveform and clinical data.

Machine learning framework

We implemented several predictive algorithms using Python 3.7 with the Scikit-learn and XGBoost packages.[34] These included logistic regression, Gaussian naive Bayes, *K*-nearest neighbours, GradientBoost, XGBoost, and random forest.[34] Model hyperparameters were tuned using grid search in combination with K = three-fold cross-validation to determine the best model.[35] Continuous variables were normalized, and categorical variables were level encoded to multiple binary variables. Patients with missing values were dropped from the analysis.

Deep learning framework

We implemented several CNN using Keras[36] with a Tensorflow (Google; Mountain View, CA, USA) backend and Python 3.7 to train binary classification models. For each outcome, we created two models: one containing only ECG waveforms as input and a second with tabular data in addition to the ECG waveforms. Electrocardiograms with paced rhythms and complete left bundle branch blocks were excluded.

For the training process, each ECG was converted to the matrix of 12×5000, where the first dimension represents the spatial leads and the second a time series of 10 s at 500 Hz. The CNN architecture in this framework was identical to a network published by our group previously for detecting patient age and sex from a single ECG alone.[19]

To maximize the utility of available data we used transfer learning from the aforementioned model, where the weights of the pre-existing network were either updated very slowly or frozen entirely and then integrated into the new model. Although a full explanation is beyond the scope of this article, for CNN these 'weights' refer to tunable parameters that are updated during training as the model learns the data.[37] They are similar to coefficients in linear regression, mathematically combining information from different portions of the image to calculate the predicted output. For this to work well, however, large quantities of data are required, and a common workaround is to use transfer learning.[37] In a transfer learning approach, the weights of a pre-existing network that is known to work well on a similar scientific question are used as the starting point and are 'frozen' such that they are not updated during training. Since the new network starts closer to the optimal state, less data are needed for training, but if the entire network is frozen then the new network cannot learn, so the number of layers frozen from the update is determined by the experiment. In our case, the number of frozen layers did not have a big impact on the performance, with 15 layers being frozen for final models. Models were trained for 30 epochs with a 0.001 learning rate and a batch size of 32. Further experiments were done using different hyperparameter values but did not impact results.

To create the models that contained both tabular and waveform data a similar architecture was used, but before the final fully connected layer the tabular data were concatenated with the features extracted by the convolutional blocks from the ECG waveform.[38] These networks were trained using the same parameters as above.

Threshold tuning

To choose the classification threshold for sensitivity analysis, different approaches were taken for the ML and DL frameworks. For the DL framework, the classification threshold was determined by selecting the point on the validation set receptor operating characteristics curve that maximized Youden's J index. For the ML framework threshold was chosen to yield the best balance between sensitivity and specificity.

Experiments performed

Our primary goal was to evaluate the ability of the ML/DL algorithms to predict the following outcomes: CFR ≤2.5, ΔCBF (%) ≤50, or the combination of both, the latter of which corresponds to CMD.

To do so we first assessed the ability of each tabular data and ECG data separately to discriminate between normal and abnormal. After that, we combined both and trained it all over again to see if this provided an added benefit. During models where ECG waveforms are used, the QRS and QT values were removed from the variables to avoid the redundancy of information.

Statistical analysis

Continuous variables distributed normally were expressed as mean ± standard deviation, and those with a skewed distribution were expressed as the median with interquartile range. Categorical variables were expressed as frequency (percentage). To compare variables between groups, we performed an unpaired *t*-test for normally distributed continuous variables, a Mann–Whitney *U* test for non-normally distributed variables, and a χ^2 test (or Fisher's exact test) for categorical variables. For the structured data analysis, common cardiovascular disease risk factors and biochemical markers, previously shown to be related to CMD, were included in the models. The features used in the tabular data models include the following: age, sex, BMI, smoking, diabetes mellitus, dyslipidaemia, hypertension, cholesterol levels (total cholesterol, LDL cholesterol, HDL cholesterol, TGs), eGFR, and ECG parameters (QRS and QT intervals), and QCA. All statistical analyses were performed using R 3.6.1 or Python 3.7.7.

Results

Of the 1975 patients in the CRT registry, 1893 patients had complete CRT study outcomes (both CFR and ΔCBF available) ECG available within the year preceding the CRT study date (*Figure 1*). The mean age was 51 ± 12 years and 66% were females (n = 1257). The prevalence of patients with CFR ≤2.5, ΔCBF (%) ≤50, and ΔCBF (%) ≤50

Table 1 Baseline characteristics

	All patients (n = 1893)	Missing data
Age (years)	51.3 ± 12.4	0
Female sex (%)	66% (n = 1257)	0
Outcomes		
CFR ≤ 2.5	25% (n = 483)	0
ΔCBF (%) ≤ 50	53% (n = 1004)	0
CMD [ΔCBF (%) ≤ 50 and/or CFR ≤ 2.5]	62% (n = 1181)	0
Comorbidities		
Diabetes (%)	11% (n = 205)	0
Hypertension (%)	43% (n = 813)	0
Hyperlipidaemia (%)	56% (n = 1058)	0
Smoking exposure (%)	46% (n = 869)	0
Labs		
eGFR (m^2/mL/kg)	78 ± 18	0
Leucocytes	6.8 ± 2.1	2%
Neutrophils	4.0 ± 1.7	18%
Neutrophils/leucocytes	0.59 ± 0.1	19%
Total cholesterol	183 (155–212)	5%
High-density lipoprotein cholesterol	51 (42–64)	6%
Low-density lipoprotein cholesterol	102 (78–127)	6%
QCA (%)	0 (0–20)	1%
QRS duration (ms)	90 (84–98)	0
QTc (ms)	423 (409–442)	0

ΔCBF, percent change in coronary blood flow; CFR, coronary flow reserve; CMD, coronary microvascular dysfunction; eGFR, estimated glomerular filtration rate; QCA, percent coronary stenosis.

and/or CFR ≤2.5 was 25%, 53%, and 62%, respectively. Patients' characteristics and CMD outcomes distribution were similar among training (60%), validation (20%), and holdout test (20%) groups across different outcomes and analyses. Baseline characteristics are shown in *Table 1*. The time between ECG and CRT study was 1 (0–4) days.

Artificial intelligence model performance

The holdout test set area under the curve (AUC) sensitivity, specificity, accuracy, positive predictive value (PPV), and negative predictive value (NPV) for different models is described in *Table 2*.

As shown in *Table 2*, the models did not exhibit a great discriminatory ability to detect the outcomes. For the CFR ≤2.5 outcome, AUC values ranged from 0.56 to 0.67, for the ΔCBF (%) ≤50 outcome, values ranged from 0.49 to 0.57. And finally, for the outcome of CMD [ΔCBF (%) ≤ 50 and/or CFR ≤ 2.5], AUC values ranged from 0.48 to 0.61. The performance was not different between the structured (tabular) and unstructured (ECG only) models. Furthermore, combining the two networks, as outlined above, did not improve the performance of the model.

The best performance in our analysis was for the outcome CFR ≤2.5, with an AUC of 0.67 (0.62, 0.73) 95% confidence interval, by the logistic regression ML algorithm using tabular variables, with accuracy, sensitivity, specificity, PPV, and NPV values of 60%, 70%, 56%, 35%, and 85%, respectively (*Figure 2*). Feature importance analysis (*Figure 3*) showed that sex, age, and smoking exposure were the most important variables. The technical calculation was done via the 'Weight' method in the XGBoost package and essentially looks at how often a variable appears in all trees. A theoretical explanation is found in books such as *The Elements of Statistical Learning*,[39] while practical instructions are found in the latest XGBoost instructions edition.[40]

Finally, we observed different threshold options for different clinical needs. When decreasing the threshold of what is considered a positive test arbitrarily from the 'optimal' one (24%) to 15%, sensitivity and NPV increase to 92% and 90%, respectively, while specificity and PPV decreased to 25% and 29%, respectively. When increasing the threshold to 37%, specificity and PPV increases to 92% and 50%, respectively, sensitivity and NPV decrease to 25% and 78%, respectively (*Figure 2*).

Discussion

The current study demonstrated that an AI-enabled algorithm based on demographics and ECG waveforms was not able to detect the difference between patients with and without CMD with high sensitivity and/or specificity. However, it has sufficient power so that with the selection of a different point along the receiver operator characteristic, the algorithm could function with a high NPV, which if prospectively confirmed could eliminate the need for invasive testing in a subset of angina patients by integrated the AI ECG into clinical guidance and decision-making in patients presenting with signs or symptoms of angina and NOCAD with suspicion of CMD. This early proof

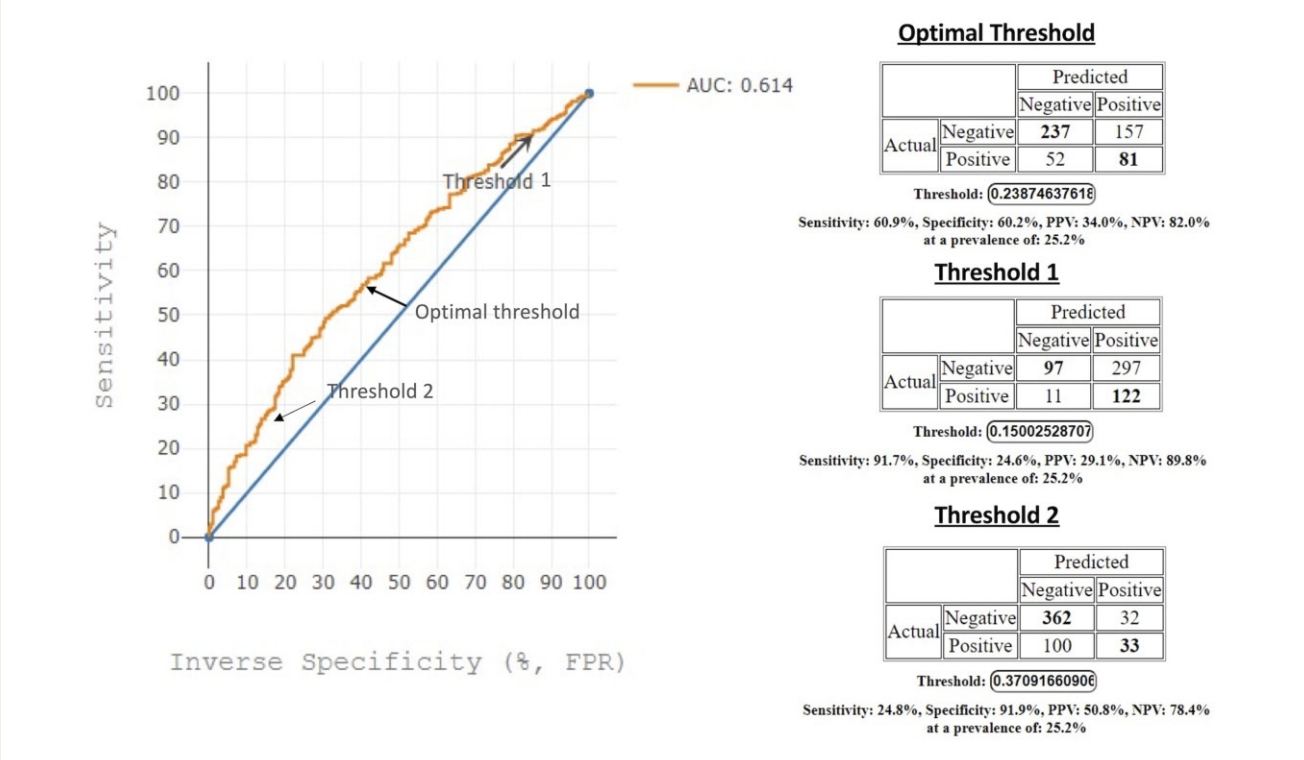


Figure 2 Receiver operating characteristic curve for the coronary flow reserve ≤2.5 outcome with proposed optimal and two other thresholds along with their corresponding confusion matrices. AUC, area under the curve; NPV, negative predictive value; PPV, positive predictive value.

Table 2 Results of structured data analysis: the area under for the receptor operating characteristics curve, accuracy, sensitivity, specify, negative predictive value, and positive predictive value for various machine learning algorithms for different outcomes

Outcome	Model	AUC (95% CI)	Accuracy	Sensitivity	Specificity	PPV	NPV
CFR ≤ 2.5	LogisticRegression	0.67 (0.62–0.73)	60	70	56	35	85
	GaussianNB	0.66 (0.62–0.72)	65	60	66	38	83
	RandomForest	0.65 (0.60–0.70)	61	62	60	35	83
	GradientBoost	0.63 (0.58–0.68)	54	74	48	32	85
	XGBoost	0.61 (0.55–0.66)	59	53	61	31	79
	AI ECG	0.56 (0.51–0.63)	54	43	57	26	75
	AI ECG + tabular data	0.64 (0.57,0.70)	57	51	59	57	53
ΔCBF (%) ≤ 50	LogisticRegression	0.55 (0.50–0.60)	55	57	55	58	54
	GaussianNB	0.57 (0.52–0.62)	56	55	46	52	49
	RandomForest	0.51 (0.46–0.56)	51	59	53	58	55
	GradientBoost	0.54 (0.49–0.59)	56	61	47	56	53
	XGBoost	0.52 (0.47–0.57)	55	42	60	54	48
	AI ECG	0.51 (0.45,0.57)	50	51	67	73	45
	AI ECG + tabular data	0.53 (0.53,0.60)	52	58	62	72	46
ΔCBF (%) ≤ 50 and/or CFR ≤ 2.5	LogisticRegression	0.61 (0.57–0.66)	57	55	59	70	43
	GaussianNB	0.61 (0.56–0.66)	59	49	63	70	42
	RandomForest	0.59 (0.54–0.64)	56	59	54	69	44
	GradientBoost	0.59 (0.54–0.64)	54	54	35	58	32
	XGBoost	0.58 (0.53–0.63)	57	70	56	35	85
	AI ECG	0.51 (0.45–0.57)	47	60	66	38	83
	AI ECG + tabular data	0.52 (0.54–0.59)	52	62	60	35	83

AI, artificial intelligence; AUC, area under the curve; ΔCBF, percent change in coronary blood flow; CFR, coronary flow reserve; CI, confidence interval; ECG, electrocardiogram; NPV, negative predictive value; PPV, positive predictive value.

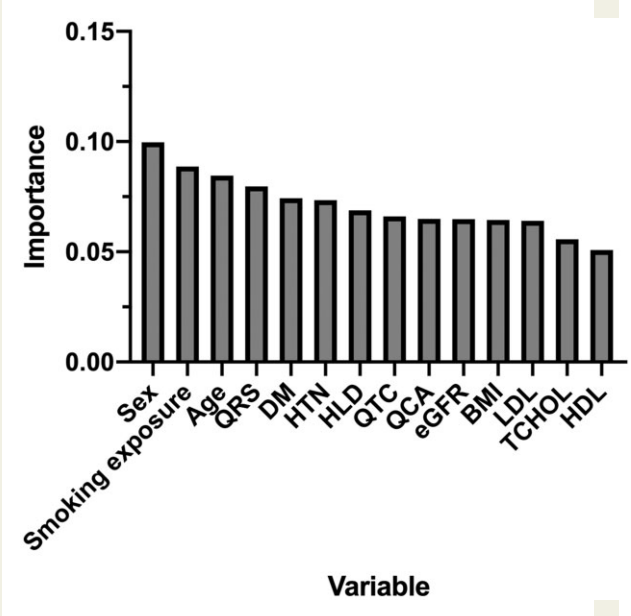


Figure 3 Feature importance for the outcome coronary flow reserve ≤2.5. BMI, body mass index; DM, diabetes mellitus; eGFR, estimated glomerular filtration rate; HDL, high-density lipoprotein; HLD, hyperlipidemia; HTN, hypertension; LDL, low-density lipoprotein; TCHOL, total cholesterol.

of concept analysis demonstrated biological plausibility, with moderate signal strength.

In the current study, we assessed separately multiple comprehensive underlying mechanisms for the diagnosis of CMD, including endothelium-dependent CMD, endothelium-independent CMD, and the combination of both. We integrated the use of both structured (tabular) and unstructured (ECG waveforms) data into the same network. However, our networks were unable to discriminate between normal and abnormal CMD with high power. In the best model predicting CFR ≤2.5, the use of tabular data with the logistic regression algorithm provided an AUC of 0.67. Decreasing the threshold of positivity increased the sensitivity to 92% and NPV to 90%.

The limited modest power of the resting, unprovoked ECG to detect CMD may reflect the biology of CMD. In the catheterization laboratory, pharmacologic manipulations are utilized to assess endothelial and non-endothelial function. Since a resting 12-lead ECG is performed in a supine, relaxed state, typically in the absence of angina, the physiologic changes that may manifest on the ECG may be absent at the time of the recording. Given that CMD is among the earliest stages of coronary atherosclerosis,[41] the inter-episode electrocardiographic abnormalities are likely more subtle or even absent, providing a weaker signal-to-noise ratio, even with the addition of demographical data. However, the fact that a very weak signal was present suggests that a provoked or symptomatic test may have a higher yield. Since smartphone-enabled ECGs are now widely available, further study using platforms with greater temporal data acquisition, potentially during symptoms, may have a higher yield.

The utility of this algorithm stems from the use of widely available demographical data along with a simple, inexpensive, non-invasive, universally available, 10-s test, to permit the identification of patients without CMD. The application of the algorithm may identify the patients who will benefit and require further invasive testing to confirm the diagnosis. An invasive test would also be needed for the identification of the endotype of CMD, endothelium-dependent or -independent, which would usually result in slightly different treatment strategies.

Historically, several other medical screening tests do not have favourable AUC values over 0.7. For example, tests such as B-type natriuretic peptide for heart failure (AUC 0.60–0.70),[42] CHA_2DS_2-VASc Score for stroke risk (0.57–0.72),[43] and even the Papanicolaou smear for cervical cancer (AUC 0.70),[44] all show a modest AUC value. Given that the output produced by the algorithm is continuous; the threshold for a positive result could be altered for various clinical applications. The binary cut-off is usually chosen to balance sensitivity and specificity, but a more sensitive cut-off might be useful in excluding patients who do not need invasive assessment of their microcirculation for the diagnosis of CMD.

Confirmatory diagnosis of CMD varies between centres and requires specialized technique and equipment.[27,45,46] The gold standard constitutes an invasive assessment using pharmacological reagents such as adenosine and acetylcholine. Although generally safe and well-tolerated, adenosine/acetylcholine might be associated with unpleasant side effects in some patients while contraindicated in other patients. Furthermore, non-invasive modalities to diagnose peripheral endothelial dysfunction, as a surrogate for coronary microcirculation, were shown to have a moderate correlation with CMD.[13–15] Hence, the development of an algorithm that uses demographical and ECG data would be of great value.

Impaired endothelial microcirculatory function is considered to be the earliest form of atherosclerosis.[41,47,48] Thus, CMD shares multiple risk factors with atherosclerosis, including age, sex, and common cardiovascular risk factors (hypertension, dyslipidaemia, and diabetes).[47,49–53] Furthermore, structural changes that might accompany CMD, which might include fibrosis, might lead to subtle changes that are detected by an AI algorithm. We previously also noted some minor electrical changes in QTc between patients with and without CMD.[23] These small variations, in addition to the small number of patients, could be the reason our model did not detect a clear pattern of differences between CMD and non-CMD patients. However, another biologically plausible reason is that CMD represents a very early disease process with very minor changes as opposed to other more advanced disease processes (like heart failure and aortic stenosis) detected by ECG. Moreover, CMD might not be a binary disease. This is highlighted in a previous paper showing that CMD indices like CFR and hyperemic microvascular resistance (HMR) provide prognostic values more precisely as continuous than as binary variables.[12] Finally, current methods to detect CMD almost always include a strategy to increase the physiological demand on the heart, therefore, unmasking abnormalities that are only apparent during states of increased demand.

Machine learning and other computational methods enable the scientific community to consider complex datasets with structured and unstructured data rather than preselecting only relevant variables. However, a key limitation for the application of these networks in the current world is validation and explainability. Uncovering the so-

called black box would add to the certainty of physicians to use the models. Although the list of feature importance may explain what the network prioritizes in the structured data model. Investigations are ongoing to uncover how the network looks at unstructured data such as ECG waveforms.

Once an algorithm is trained, it can be applied to any set of demographical data. This would greatly facilitate point-of-care clinical guidance in patients presenting with symptoms of ischaemia but who have NOCAD on angiogram. Furthermore, if ECG data are included in the model, widely available smartphone technology could have a role in implementing the algorithm on the ECG. For example, our group has previously shown the ability to implement these algorithms, on single-lead smartphone-generated waveforms.[54]

Limitations

Our study is best understood in the context of its limitations. In comparison to other applications of CNN by our group,[19–21] the population size is small, which might diminish the discriminatory power of our models. Furthermore, our centre is a tertiary centre therefore referral bias cannot be excluded. Furthermore, other labs such as hsCRP, homocysteine, and NT-proBNP were included in separate models, but this decreased the sample number severely since patients with missing values (hsCRP 43%, homocysteine 51%, and NT-proBNP 62% patients with missing values, respectively) were excluded from the pipeline. This may have led to selection bias. Further development with larger populations would assess if this lack of high discriminatory power is due to CMD being a challenging diagnosis without invasive provocatory testing, if the algorithm lacks enough data, or a mixture of both limitations. Additional validation is also needed to ensure the diagnostic performance of this model in specific, such as in patients with co-existing obstructive coronary artery disease or patients with heart failure. Finally, NPV is dependent on the pretest probability and the prevalence of the disease, therefore, sensitivity is a better assessor of the utility of this algorithm.

Conclusion

Coronary microvascular dysfunction is common in patients with NOCAD presenting with signs/symptoms of ischaemia. An AI-based method may be able to assist in the clinical decision in this population and be used to exclude patients who may not require further invasive testing.

Funding

A.A. is supported by the James Nutter Family and Maria Long Family Fellowship in Cardiovascular Research grant.

Conflict of interest: none declared.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. Jespersen L, Hvelplund A, Abildstrom SZ, Pedersen F, Galatius S, Madsen JK, Jorgensen E, Kelbaek H, Prescott E. Stable angina pectoris with no obstructive coronary artery disease is associated with increased risks of major adverse cardiovascular events. *Eur Heart J* 2012;**33**:734–744.
2. Patel MR, Peterson ED, Dai D, Brennan JM, Redberg RF, Anderson HV, Brindis RG, Douglas PS. Low diagnostic yield of elective coronary angiography. *N Engl J Med* 2010;**362**:886–895.
3. Sara JD, Widmer RJ, Matsuzawa Y, Lennon RJ, Lerman LO, Lerman A. Prevalence of coronary microvascular dysfunction among patients with chest pain and nonobstructive coronary artery disease. *JACC Cardiovasc Interv* 2015;**8**:1445–1453.
4. Britten MB, Zeiher AM, Schachinger V. Microvascular dysfunction in angiographically normal or mildly diseased coronary arteries predicts adverse cardiovascular long-term outcome. *Coron Artery Dis* 2004;**15**:259–264.
5. Herrmann J, Kaski JC, Lerman A. Coronary microvascular dysfunction in the clinical setting: from mystery to reality. *Eur Heart J* 2012;**33**:2771–2782b.
6. Marinescu MA, Loffler AI, Ouellette M, Smith L, Kramer CM, Bourque JM. Coronary microvascular dysfunction, microvascular angina, and treatment strategies. *JACC Cardiovasc Imaging* 2015;**8**:210–220.
7. Marks DS, Gudapati S, Prisant LM, Weir B, diDonato-Gonzalez C, Waller JL, Houghton JL. Mortality in patients with microvascular disease. *J Clin Hypertens (Greenwich)* 2004;**6**:304–309.
8. Pepine CJ, Anderson RD, Sharaf BL, Reis SE, Smith KM, Handberg EM, Johnson BD, Sopko G, Bairey Merz CN. Coronary microvascular reactivity to adenosine predicts adverse outcome in women evaluated for suspected ischemia results from the National Heart, Lung and Blood Institute WISE (Women's Ischemia Syndrome Evaluation) study. *J Am Coll Cardiol* 2010;**55**:2825–2832.
9. Serruys PW, di Mario C, Piek J, Schroeder E, Vrints C, Probst P, de Bruyne B, Hanet C, Fleck E, Haude M, Verna E, Voudris V, Geschwind H, Emanuelsson HKAN, MüHlberger V, Danzi G, Peels HO, Ford AJ, Boersma E. Prognostic value of intracoronary flow velocity and diameter stenosis in assessing the short- and long-term outcomes of coronary balloon angioplasty: the DEBATE Study (Doppler Endpoints Balloon Angioplasty Trial Europe). *Circulation* 1997;**96**:3369–3377.
10. Lerman A, Zeiher AM. Endothelial function: cardiac events. *Circulation* 2005;**111**:363–368.
11. Corban MT, Godo S, Burczak DR, Noseworthy PA, Lewis TT, Lerman BR, Gulati LO, Lerman R. A coronary endothelial dysfunction is associated with increased risk of incident atrial fibrillation. *J Am Heart Assoc* 2020;**9**:e014850.
12. Toya T, Corban MT, Park JY, Ahmad A, Ozcan I, Sebaali F, Sara JDS, Gulati R, Lerman LO, Lerman A. Prognostic impact and clinical outcomes of coronary flow reserve and hyperemic microvascular resistance. *EuroIntervention* 2021;**17**:569–575.
13. Cassar A, Chareonthaitawee P, Rihal CS, Prasad A, Lennon RJ, Lerman LO, Lerman A. Lack of correlation between noninvasive stress tests and invasive coronary vasomotor dysfunction in patients with nonobstructive coronary artery disease. *Circ Cardiovasc Interv* 2009;**2**:237–244.
14. Barton D, Xie F, O'Leary E, Chatzizisis YS, Pavlides G, Porter TR. The relationship of capillary blood flow assessments with real time myocardial perfusion echocardiography to invasively derived microvascular and epicardial assessments. *J Am Soc Echocardiogr* 2019;**32**:1095–1101.
15. Bierig SM, Mikolajczak P, Herrmann SC, Elmore N, Kern M, Labovitz AJ. Comparison of myocardial contrast echocardiography derived myocardial perfusion reserve with invasive determination of coronary flow reserve. *Eur J Echocardiogr* 2009;**10**:250–255.
16. Lopez-Jimenez F, Attia Z, Arruda-Olson AM, Carter R, Chareonthaitawee P, Jouni H, Kapa S, Lerman A, Luong C, Medina-Inojosa JR, Noseworthy PA, Pellikka PA, Redfield MM, Roger VL, Sandhu GS, Senecal C, Friedman PA. Artificial intelligence in cardiology: present and future. *Mayo Clin Proc* 2020;**95**:1015–1039.
17. Ahmad A, Ibrahim Z, Sakr G, El-Bizri A, Masri L, Elhajj IH, El-Hachem N, Isma'eel H. A comparison of artificial intelligence-based algorithms for the identification of patients with depressed right ventricular function from 2-dimentional echocardiography parameters and clinical features. *Cardiovasc Diagn Ther* 2020;**10**:859–868.
18. Toya T, Ahmad A, Attia Z, Cohen-Shelly M, Ozcan I, Noseworthy PA, Lopez-Jimenez F, Kapa S, Lerman LO, Friedman PA, Lerman A. Vascular aging detected by peripheral endothelial dysfunction is associated with ECG-derived physiological aging. *J Am Heart Assoc* 2021;**10**:e018656.
19. Attia ZI, Friedman PA, Noseworthy PA, Lopez-Jimenez F, Ladewig DJ, Satam G, Pellikka PA, Munger TM, Asirvatham SJ, Scott CG, Carter RE, Kapa S. Age and sex estimation using artificial intelligence from standard 12-lead ECGs. *Circ Arrhythm Electrophysiol* 2019;**12**:e007284.
20. Attia ZI, Kapa S, Lopez-Jimenez F, McKie PM, Ladewig DJ, Satam G, Pellikka PA, Enriquez-Sarano M, Noseworthy PA, Munger TM, Asirvatham SJ, Scott CG,

Carter RE, Friedman PA. Screening for cardiac contractile dysfunction using an artificial intelligence-enabled electrocardiogram. *Nat Med* 2019;**25**:70–74.
21. Attia ZI, Noseworthy PA, Lopez-Jimenez F, Asirvatham SJ, Deshmukh AJ, Gersh BJ, Carter RE, Yao X, Rabinstein AA, Erickson BJ, Kapa S, Friedman PA. An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction. *Lancet* 2019;**394**:861–867.
22. Ko WY, Siontis KC, Attia ZI, Carter RE, Kapa S, Ommen SR, Demuth SJ, Ackerman MJ, Gersh BJ, Arruda-Olson AM, Geske JB, Asirvatham SJ, Lopez-Jimenez F, Nishimura RA, Friedman PA, Noseworthy PA. Detection of hypertrophic cardiomyopathy using a convolutional neural network-enabled electrocardiogram. *J Am Coll Cardiol* 2020;**75**:722–733.
23. Sara JD, Lennon RJ, Ackerman MJ, Friedman PA, Noseworthy PA, Lerman A. Coronary microvascular dysfunction is associated with baseline QTc prolongation amongst patients with chest pain and non-obstructive coronary artery disease. *J Electrocardiol* 2016;**49**:87–93.
24. Sara JD, Sugrue A, Kremen V, Qiang B, Sapir Y, Attia ZI, Ackerman MJ, Friedman PA, Lerman A, Noseworthy PA. Electrocardiographic predictors of coronary microvascular dysfunction in patients with non-obstructive coronary artery disease: Utility of a novel T wave analysis program. *Int J Cardiol* 2016;**203**:601–606.
25. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, Burovski E, Peterson P, Weckesser W, Bright J, van der Walt SJ, Brett M, Wilson J, Millman KJ, Mayorov N, Nelson ARJ, Jones E, Kern R, Larson E, Carey CJ, Polat İ, Feng Y, Moore EW, VanderPlas J, Laxalde D, Perktold J, Cimrman R, Henriksen I, Quintero EA, Harris CR, Archibald AM, Ribeiro AH, Pedregosa F, van Mulbregt P; SciPy 1.0 Contributors. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods* 2020;**17**:261–272.
26. Ahmad A, Corban MT, Toya T, Verbrugge FH, Sara JD, Lerman LO, Borlaug BA, Lerman A. Coronary microvascular dysfunction is associated with exertional haemodynamic abnormalities in patients with heart failure with preserved ejection fraction. *Eur J Heart Fail* 2021;**23**:765–772.
27. Hasdai D, Cannan CR, Mathew V, Holmes DR Jr, Lerman A. Evaluation of patients with minimally obstructive coronary artery disease and angina. *Int J Cardiol* 1996;**53**:203–208.
28. Hasdai D, Holmes DR Jr, Higano ST, Burnett JC Jr, Lerman A. Prevalence of coronary blood flow reserve abnormalities among patients with nonobstructive coronary artery disease and chest pain. *Mayo Clin Proc* 1998;**73**:1133–1140.
29. Maniu CV, Higano ST, Lerman A. Assessing coronary endothelial dysfunction. *Circulation* 2002;**106**:e48; discussion e48.
30. Widmer RJ, Samuels B, Samady H, Price MJ, Jeremias A, Anderson RD, Jaffer FA, Escaned J, Davies J, Prasad M, Grines C, Lerman A. The functional assessment of patients with non-obstructive coronary artery disease: expert review from an international microcirculation working group. *EuroIntervention* 2019;**14**:1694–1702.
31. Ahmad A, Corban MT, Toya T, Sara JD, Lerman B, Park JY, Lerman LO, Lerman A. Coronary microvascular endothelial dysfunction in patients with angina and nonobstructive coronary artery disease is associated with elevated serum homocysteine levels. *J Am Heart Assoc* 2020;**9**:e017746.
32. Bell MR, Britson PJ, Chu A, Holmes DR Jr, Bresnahan JF, Schwartz RS. Validation of a new UNIX-based quantitative coronary angiographic system for the measurement of coronary artery lesions. *Catheter Cardiovasc Diagn* 1997;**40**:66–74.
33. Widmer RJ, Flammer AJ, Herrmann J, Rodriguez-Porcel M, Wan J, Cohen P, Lerman LO, Lerman A. Circulating humanin levels are associated with preserved coronary endothelial function. *Am J Physiol Heart Circ Physiol* 2013;**304**:H393–H397.
34. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V. Scikit-learn: machine learning in Python. *J Mach Learn Res* 2011;**12**:2825–2830.
35. Kuhn M, Johnson K. *Applied Predictive Modeling*. New York, NY: Springer; 2013.
36. Chollet F. keras. 2015. https://scholar.google.com/citations?view_op=view_citation&hl=en&user=VfYhf2wAAAAJ&citation_for_view=VfYhf2wAAAAJ:9pM33mqn1YgC.
37. Goodfellow I, Bengio Y, Courville A. *Deep Learning*. Cambridge, MA: MIT Press; 2016.
38. Cohen-Shelly M, Attia ZI, Friedman PA, Ito S, Essayagh BA, Ko W-Y, Murphree DH, Michelena HI, Enriquez-Sarano M, Carter RE, Johnson PW, Noseworthy PA, Lopez-Jimenez F, Oh JK. Electrocardiogram screening for aortic valve stenosis using artificial intelligence. *Eur Heart J* 2021;**42**:2885–2896.
39. Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 20 Springer Series in Statistics*. New York: Springer-Verlag; 2009.
40. Python API Reference. https://xgboost.readthedocs.io/en/latest/python/python_api.html?highlight=date#xgboost.Booster.update (9 January 2021).
41. Gutiérrez E, Flammer AJ, Lerman LO, Elízaga J, Lerman A, Fernández-Avilés F. Endothelial dysfunction over the course of coronary artery disease. *Eur Heart J* 2013;**34**:3175–3181.
42. Bhalla V, Isakson S, Bhalla MA, Lin JP, Clopton P, Gardetto N, Maisel AS. Diagnostic ability of B-type natriuretic peptide and impedance cardiography: testing to identify left ventricular dysfunction in hypertensive patients. *Am J Hypertens* 2005;**18**:73S–81S.
43. Wu JT, Wang SL, Chu YJ, Long DY, Dong JZ, Fan XW, Yang HT, Duan HY, Yan LJ, Qian P. CHADS2 and CHA2DS2-VASc scores predict the risk of ischemic stroke outcome in patients with interatrial block without atrial fibrillation. *J Atheroscler Thromb* 2017;**24**:176–184.
44. Chen Y, Cui Z, Xiao Z, Hu M, Jiang C, Lin Y, Chen Y. PAX1 and SOX1 methylation as an initial screening method for cervical cancer: a meta-analysis of individual studies in Asians. *Ann Transl Med* 2016;**4**:365.
45. Bugiardini R, Bairey Merz CN. Angina with "normal" coronary arteries: a changing philosophy. *JAMA* 2005;**293**:477–484.
46. Ford TJ, Rocchiccioli P, Good R, McEntegart M, Eteiba H, Watkins S, Shaukat A, Lindsay M, Robertson K, Hood S, Yii E, Sidik N, Harvey A, Montezano AC, Beattie E, Haddow L, Oldroyd KG, Touyz RM, Berry C. Systemic microvascular dysfunction in microvascular and vasospastic angina. *Eur Heart J* 2018;**39**:4086–4097.
47. Bonetti PO, Lerman LO, Lerman A. Endothelial dysfunction: a marker of atherosclerotic risk. *Arterioscler Thromb Vasc Biol* 2003;**23**:168–175.
48. Lerman A, Holmes DR Jr, Bell MR, Garratt KN, Nishimura RA, Burnett JC Jr. Endothelin in coronary endothelial dysfunction and early atherosclerosis in humans. *Circulation* 1995;**92**:2426–2431.
49. Corban MT, Lerman LO, Lerman A. Endothelial dysfunction. *Arterioscler Thromb Vasc Biol* 2019;**39**:1272–1274.
50. AlBadri A, Eshtehardi P, Hung OY, Bouchi Y, Khawaja S, Mercado K, Corban MT, Mehta PK, Shaw LJ, Samady H. Coronary microvascular dysfunction is associated with significant plaque burden and diffuse epicardial atherosclerotic disease. *JACC Cardiovasc Interv* 2019;**12**:1519–1520.
51. Corban MT, Lerman LO, Lerman A. Endothelial dysfunction: cardiovascular disease pathophysiology hidden in plain sight. *Am Heart Assoc* 2019;**39**:1272–1274.
52. Godo S, Corban MT, Toya T, Gulati R, Lerman LO, Lerman A. Association of coronary microvascular endothelial dysfunction with vulnerable plaque characteristics in early coronary atherosclerosis. *EuroIntervention* 2020;**16**:387–394.
53. Hebbel RP, Wei P, Milbauer L, Corban MT, Solovey A, Kiley J, Pattee J, Lerman LO, Pan W, Lerman A. Abnormal endothelial gene expression associated with early coronary atherosclerosis. *J Am Heart Assoc* 2020;**9**:e016134.
54. Yasin OZ, Attia Z, Dillon JJ, DeSimone CV, Sapir Y, Dugan J, Somers VK, Ackerman MJ, Asirvatham SJ, Scott CG, Bennet KE, Ladewig DJ, Sadot D, Geva AB, Friedman PA. Noninvasive blood potassium measurement using signal-processed, single-lead ECG acquired from a handheld smartphone. *J Electrocardiol* 2017;**50**:620–625.