ESC European Society of Cardiology

ORIGINAL ARTICLE

Prediction of coronary artery calcium scoring from surface electrocardiogram in atherosclerotic cardiovascular disease: a pilot study

Peter D. Farjo [1†], Naveena Yanamala[1,2†], Nobuyuki Kagiyama[1,3,4], Heenaben B. Patel[1], Grace Casaclang-Verzosa[1], Negin Nezarat[5], Matthew J. Budoff[5], and Partho P. Sengupta[1]*

[1]Division of Cardiology, Department of Medicine, West Virginia University Heart and Vascular Institute, 1 Medical Center Drive, Morgantown, WV 26506, USA; [2]Institute for Software Research, School of Computer Science, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA 15213, USA; [3]Department of Digital Health and Telemedicine R&D, Juntendo University, 211 Hongo, Bunkyo City, Tokyo 113-8421, Japan; [4]Department of Cardiovascular Biology and Medicine, Juntendo University, 211 Hongo, Bunkyo City, Tokyo 113-8421, Japan; and [5]Department of Medicine, Lundquist Institute, Harbor-UCLA Medical Center, 1124 West Carson St, Torrance, CA 90502, USA

Aims	Coronary artery calcium (CAC) scoring is an established tool for cardiovascular risk stratification. However, the lack of widespread availability and concerns about radiation exposure have limited the universal clinical utilization of CAC. In this study, we sought to explore whether machine learning (ML) approaches can aid cardiovascular risk stratification by predicting guideline recommended CAC score categories from clinical features and surface electrocardiograms.
Methods and results	In this substudy of a prospective, multicentre trial, a total of 534 subjects referred for CAC scores and electrocardiographic data were split into 80% training and 20% testing sets. Two binary outcome ML logistic regression models were developed for prediction of CAC scores equal to 0 and ≥400. Both CAC = 0 and CAC ≥400 models yielded values for the area under the curve, sensitivity, specificity, and accuracy of 84%, 92%, 70%, and 75%, and 87%, 91%, 75%, and 81%, respectively. We further tested the CAC ≥400 model to risk stratify a cohort of 87 subjects referred for invasive coronary angiography. Using an intermediate or higher pretest probability (≥15%) to predict CAC ≥400, the model predicted the presence of significant coronary artery stenosis ($P = 0.025$), the need for revascularization ($P < 0.001$), notably bypass surgery ($P = 0.021$), and major adverse cardiovascular events ($P = 0.023$) during a median follow-up period of 2 years.
Conclusion	ML techniques can extract information from electrocardiographic data and clinical variables to predict CAC score categories and similarly risk-stratify patients with suspected coronary artery disease.
Keywords	Calcium score • Machine learning • ECG • Prediction

* Corresponding author. Tel: 304-598-4478, Email: patho.sengupta@hsc.wvu.edu
† The first two authors contributed equally to the study.

Introduction

Computed tomography (CT) derived coronary artery calcium (CAC) scoring is a validated measure that correlates well with subclinical coronary atherosclerotic burden.[1,2] Both the European Society of Cardiology (ESC) and American College of Cardiology/American Heart Association (ACC/AHA) have provided clinical recommendations for using CAC to guide cardiovascular risk assessment and metabolic disease management.[1,3] Typically reported as an Agatston score, the range of CAC quantification is clinically significant with scores of 0 exhibiting almost 100% negative predictive value for significant coronary artery disease, while scores above 400, considered severely elevated, predict higher rates of acute coronary syndrome at about 2 years.[1,4–6] The growing body of literature behind CAC is demonstrating other applications including periprocedural risk assessment, augmentation of functional stress testing, and prediction of atrial fibrillation.[7–9] Nonetheless, with limitations including radiation exposure, lack of widespread availability, and concern for misinterpretation, CAC has yet to be universally embraced.[10–12]

The use of innovative machine learning (ML) techniques could allow for extraction of the predictive value of CAC using office-based clinical features and tests like surface electrocardiogram (ECG). We and others have previously reported the successful application of ECG-based feature analysis for predicting information like left ventricular systolic and diastolic dysfunction available from cardiac imaging tests like echocardiography.[13–15] In the present investigation, we explored the application of ML techniques with clinical features and surface ECG indices to predict clinically used CAC threshold scores (0 and ≥400 Agatston score) from a multi-institutional, prospective patient cohort from a previous trial who underwent surface ECG and CT angiography (CTA). In a separate cohort of patients undergoing invasive coronary angiography, we further prospectively tested whether the model that predicts severe CAC (≥400 Agatston units) can similarly risk-stratify the patients for the severity of coronary artery disease and related adverse clinical outcomes.

Methods

In this *post hoc* substudy, we utilized data from a multicentre, prospective study designed for the development of an ML model from surface ECG for predicting the presence of diastolic dysfunction as measured using echocardiography.[13] From the four North-American sites that enrolled patients in the study, two sites [Icahn School of Medicine at Mount Sinai, New York, and University of California, Los Angeles (UCLA), Los Angeles] included participants who were referred for ambulatory coronary CTA. Subjects were screened prior to enrolment for inclusion criteria of current sinus rhythm and age 18 years or greater and exclusion criteria of pregnancy, chest deformities, pacemaker placement, unwilling, or inability to provide informed consent, and enrolment in another clinical study. Subjects underwent a 12-lead signal-processed surface ECG (spECG) and coronary CTA in the same visit. A total of 534 had complete and adequate CTA and spECG data and were utilized. This sample of subjects was split into 80% training and 20% testing cohorts.

We included a separate cohort from a third site (West Virginia University, Morgantown, WV, USA) that enrolled consecutive patients undergoing invasive coronary angiography as a part of the protocol testing the ML model to predict diastolic dysfunction related to coronary artery disease. They underwent surface ECG within 48 hours before their

procedure. A total of 87 participants had optimal ECGs and available invasive coronary angiography data. Subjects were grouped based on their prediction model output and followed for 24 months to evaluate the primary composite outcome of major adverse cardiovascular events (MACE) including myocardial infarction, unstable angina, stroke, cardiovascular hospitalization, and all-cause mortality. Secondary clinical outcomes included the individual components of the primary outcome, any coronary stenosis defined as any ≥25% stenotic lesion, significant stenosis defined as a ≥50% lesion in the left main coronary artery or ≥70% lesion elsewhere, and revascularization.[16] Specific definitions of the outcomes can be found in the Supplementary material online, *Table S1*. The original study protocol was approved by the institutional review boards for the home institution and all participating facilities, and all participants provided written informed consent prior to enrolment.

Clinical, demographic, and ECG characteristics for all subjects were collected and analysed. Clinical data including body mass index (BMI), heart rate, systolic blood pressure (SBP), and diastolic blood pressure (DBP) were recorded at the time of enrolment. Comorbidities were based on physician-documented history or reported medical therapy for associated diagnoses such as hypertension (HTN), hyperlipidaemia (HLD), and diabetes mellitus (DM). Obesity was defined as a BMI greater than 30 kg/m².

Signal-processed surface ECG

A 12-lead surface ECG was performed on all subjects on the day of their CT scan or within 48 h of their invasive coronary angiogram. Signal processing was performed using continuous wavelet transform technology (MyoVista hsECG Informatics, HeartSciences, Southlake, TX, USA). This technique utilizes validated mathematics similar to Fourier transform to display the spectral components of the ECG signal throughout the cardiac cycle.[17,18] In traditional ECG, these components are averaged forming the well-known output morphologies. With continuous wavelet transform, the signals are converted into normalized energy distributions and decompressed in a time-frequency plane allowing for component analysis of local, transient, and intermittent patterns.[19] Signal amplitude characterizing myocardial energy is converted to a colour scale from 0 to 255 with blue indicating the lowest energy and red the highest energy. It is then plotted in a time-frequency plane with frequency on the *y*-axis and time on the *x*-axis demonstrating the cardiac cycle creating the MyoVista Color Waveform. Predetermined indices are then extracted including relative values at certain points in the cardiac cycle along with omnibus measures.

Machine learning modelling

We developed two logistic regression (LR) models to predict the binary outcome of CAC scoring using spECG features and 16 basic clinical features: age, gender, BMI, heart rate, SBP and DBP, and the presence of comorbidities including cerebrovascular disease, coronary artery disease (CAD), peripheral artery disease, DM, HTN, HLD, obesity, chronic lung disease, tobacco use, and chronic kidney disease. After feature selection using the Boruta algorithm, the first model predicting a CAC score of zero (CAC = 0 model) used clinical features alone to predict the probability of having a CAC score of 0 or not. The second model predicting a CAC score ≥400 or not (CAC ≥400 model) included both clinical and spECG variables for prediction.

LR is a widely used statistical model that allows for multivariate analysis and utilizes a logistic function to model a binary outcome, resulting in the generation of coefficients which can be used to predict the probability of having a CAC score of 0 ('yes' or '1') or greater ('no' or '0') and greater or equal to 400 ('yes' or '1') or not ('no' or '0'). The dataset employed in this study for developing the predictive models has: (i) high imbalance

with unequal distribution of classes (81% vs. 19%) within the dataset, and (ii) high-dimensional data with a large number (>1000) of features in the dataset. LR analysis makes several key assumptions including no multicollinearity and normality relationship between independent variables and the log-odds. As an initial step, data were preprocessed to remove columns with zero variance. The dataset was then randomly split into training (80%) and test (20%) sets. To permit the training of ML algorithms without incurring an overfitting problem, we performed feature selection with a random forest-based approach using the Boruta algorithm in the R statistical environment to capture features that are critical in predicting an outcome variable.[20] Using the features retained after applying the Boruta algorithm, multivariate L2-regularized LR model was built to predict the CAC = 0 model, and an L1-regularized LR model to predict the logarithm of the odds of CAC ≥ 400 model. Importantly, L1-regularized LR, a commonly used regularized version of LR, has been shown to outperform classic LR in modelling imbalanced and high-dimensional data.[21–24] Similarly, L2-regularized LR deals with preventing multicollinearity by coefficient shrinkage to control the trade-off between fitting the training data well and keeping the parameters small to avoid overfitting.[25,26] The developed models in each case were subsequently assessed using the unseen internal validation cohort. The study flow chart is displayed in *Figure 1*. Sensitivity and specificity were calculated as true positive/(true positive + false negative) and true negative/(true negative + false positive), respectively. Receiver operating characteristic (ROC) curves were drawn by plotting sensitivity and 1-specificity to multiple thresholds. An area under the curve (AUC) was calculated as area under the ROC curves and tested for statistical significance. In models obtained from LR, log-odds predictions were converted to probabilities using the inverse logit formula, $e^Y/(1+e^Y)$. Finally, the developed CAC ≥ 400 LR model was set at a threshold probability of 15% or above to have a severe CAC score to predict MACE. This figure was derived from 2019 ESC guidelines labelling symptomatic patients with a pretest probability of ≥15% intermediate- to high risk for obstructive coronary artery disease.[27] The clinical invasive angiography (IA) cohort was then stratified with this ≥15% probability model and followed for outcomes. The CAD Consortium clinical score, another validated pretest probability scoring systems, was used to stratify the invasive cohort and compare with our model.[27–29]

Statistical analysis

Clinical and demographic characteristics of both cohorts along with outcomes of the clinical IA cohort were analysed. For continuous variables, in addition to mean and standard deviation, Kolmogorov–Smirnov tests were performed to assure normality, and Student's *t*-tests were
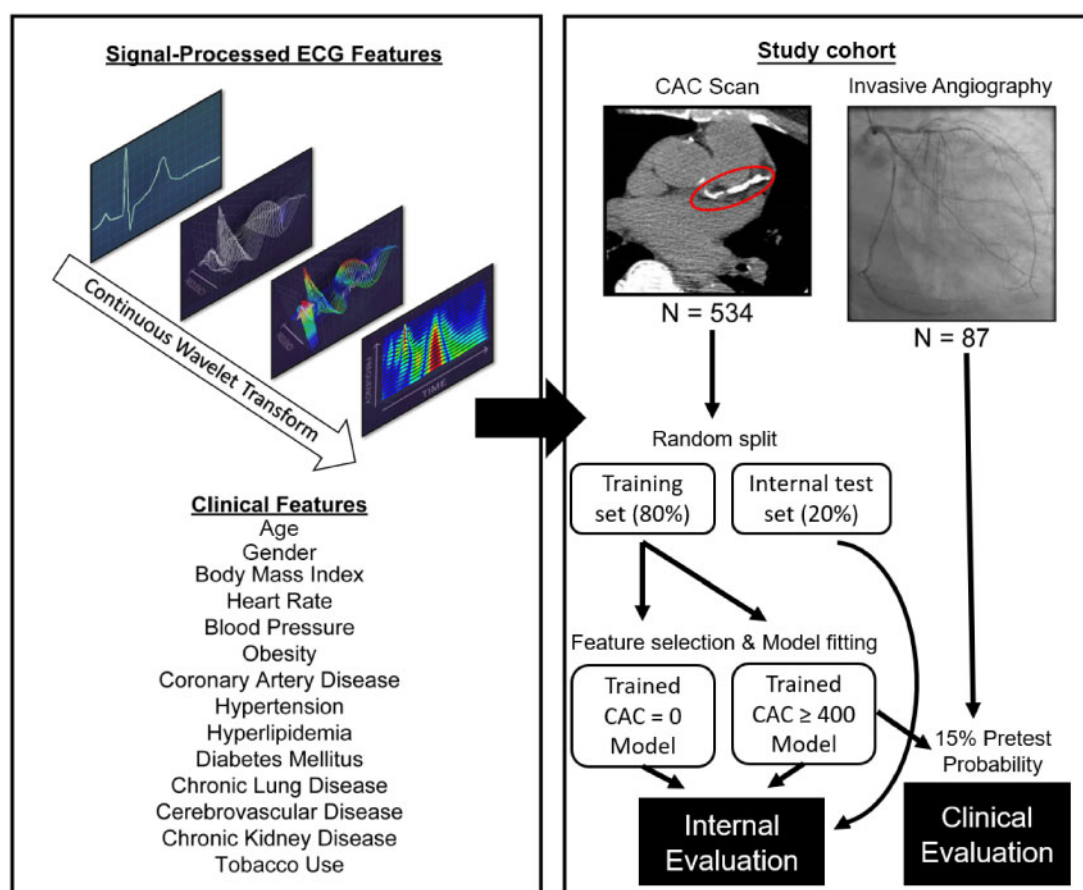
Figure 1 Study flow chart is displayed with features including signal processed ECG and clinical features in the left box. The right box displays the study model development cohort with 534 subjects undergoing CAC scan enrolled. They were split into training and testing sets to develop the CAC = 0 and CAC ≥ 400 models. A separate cohort of 87 subjects undergoing invasive angiography were stratified using the CAC ≥ 400 model at a 15% probability and followed for clinical MACE outcomes for 2 years.

conducted to investigate for significant differences in distributions. For categorical variables, in addition to percentage, χ^2 or Fisher's exact tests were conducted. Specifically for outcomes, time-to-event data were evaluated with the use of Kaplan–Meier estimates and log-rank testing. All missing data were considered missing completely at random and the default pairwise deletion method was utilized.

For *post hoc* sample size justification, we used previously recommended guidelines.[30,31] Let *p* be the smallest of the proportions of negative or positive cases in the population and *k* the number of independent variables, then the minimum number of cases to include is: $N = 10 \ k/p$. For example: In the case of CAC = 0 model, we have 16 covariates to include in the model and the proportion of positive cases in the population is 0.46 (46%). The minimum number of cases required is $N = 10 \times 16/0.46 = 349$ – lower than the sample size (n = 428 – training dataset) considered in our study. Furthermore, the results from any LR model with the number of observations per independent variable ranging from at least five to nine have been reported to be reliable, especially if results are statistically significant.[30] In this study, the ratio of number of independent variables to observations in the case of CAC = 0 and CAC ≥ 400 models were about 1:27 and 1:14, respectively—higher than the

recommended guidelines to ensure that the likelihood of risk of overfitting the developed CAC models is low.

All ML modelling and statistical analyses were performed using JMP version 14.0 (JMP, SAS Institute, Cary, NC, USA), Medcalc for Windows, version 19.4.1 (MedCalc Software, Ostend, Belgium) and RStudio version 3.1.3 (Vienna, Austria). Statistical significance was concluded if alpha was ≤0.05 for all tests.

Results

Non-invasive CAC cohort

A total of 552 subjects (New York City: 194 and Los Angeles: 358) were initially enrolled in the study, and 534 subjects (New York City: 183 and Los Angeles: 351) were included in model development. The 18 excluded subjects were all due to lack of CAC score reporting. The remaining cohort was randomly split into training and testing sets containing 428 and 106 subjects respectively. Clinical characteristics of the entire cohort and subsets are shown in *Table 1*. Based on

Table 1   Baseline characteristics of coronary artery calcium cohort

	CT CAC cohort (n = 534)	Training set (n = 428)	Testing set (n = 106)	*P*-value
Age (years)	58.3 ± 11.2	58.4 ± 11.1	57.7 ± 11.6	0.514
Female sex	228 (42.7)	184 (43.0)	44 (41.5)	0.783
Body mass index (kg/m^2)	29.0 ± 6.3	29.1 ± 6.2	28.6 ± 6.6	0.470
Heart rate	62.0 ± 9.8	62.2 ± 9.9	61.2 ± 9.2	0.355
Blood pressure (mmHg)				
Systolic	129.4 ± 19.0	130.5 ± 18.8	125.3 ± 19.4	0.012
Diastolic	78.2 ± 11.4	78.6 ± 11.2	76.6 ± 12.1	0.109
Symptomatic	139 (26.0)	113 (26.4)	26 (24.5)	0.694
CAC score	250.5 ± 548.2	246.3 ± 538.9	267.8 ± 586.7	0.718
0	196 (36.7)	152 (35.5)	44 (41.5)	0.252
>0 to <400	238 (44.6)	199 (46.5)	39 (36.8)	0.072
≥400	100 (18.7)	77 (18.0)	23 (21.7)	0.381
Any stenosis	35 (6.6)	25 (5.8)	10 (9.43)	0.181
Smoker				0.550
Current	38 (7.1)	33 (7.7)	5 (4.7)	
Former	148 (27.7)	117 (27.3)	31 (29.3)	
Never	348 (65.2)	278 (65.0)	70 (66.0)	
Obesity, BMI ≥30 kg/m^2	189 (35.4)	156 (36.5)	33 (31.1)	0.306
Hypertension	279 (52.2)	228 (53.3)	51 (48.1)	0.341
Hyperlipidaemia	367 (70.6)	294 (70.8)	73 (69.5)	0.791
Diabetes mellitus	147 (28.0)	117 (27.9)	30 (28.6)	0.884
Coronary artery disease	47 (8.8)	37 (8.6)	10 (9.4)	0.797
Previous PCI	6 (1.2)	5 (1.2)	1 (1.0)	0.839
Previous CABG	1 (0.2)	1 (0.2)	0 (0.0)	0.618
Peripheral artery disease	0 (0.0)	0 (0.0)	0 (0.0)	
Prior stroke/TIA	8 (1.5)	4 (0.9)	4 (3.8)	0.033
Chronic kidney disease	13 (3.8)	11 (4.0)	2 (2.7)	0.591
Chronic lung disease	28 (8.2)	23 (8.6)	5 (6.9)	0.654

Values are counts (%) or mean ± standard deviation.
Symptomatic is defined as chest pain or ≥NYHA class II symptoms for indication of the test.
Any stenosis is defined as any noted coronary artery disease seen on associated CT angiography.
BMI, body mass index; CA, coronary angiogram, CAC, coronary artery calcium; CABG, coronary artery bypass graft; PCI, percutaneous coronary intervention; TIA, transient ischaemic attack.

the variable importance scores constituted to the final CAC = 0 model, only nine out of 16 clinical variables initially collected were included in the final model. These included age, gender, SBP, DBP, heart rate, HLD, CAD, chronic lung disease, and BMI. ECG features were not included in the model since they did not improve the prediction performance over these routine clinical factors. For the CAC ≥ 400 model, a total of 31 features constituted the final model with about 90% of the effect from spECG features. However, the overall ranking of features was highest for the three clinical features utilized including CAD, age, and gender. The specific features and importance scores are reported in *Figure 2*.

For model development, multivariable L2-regularized and L1-regularized LR were optimized for the CAC = 0 and CAC ≥ 400 models. The area under the ROC curve for the developed LR CAC = 0 and CAC ≥ 400 models are shown in *Figure 2*. Specifically, the CAC = 0 model demonstrated an AUC of 0.84 [sensitivity 92%, specificity 70%, accuracy 75%, 95% confidence interval (CI) 0.75–0.90, $P < 0.001$] and the CAC ≥ 400 LR model demonstrated an AUC of 0.87 (sensitivity 91%, specificity 75%, accuracy 71%, 95% CI

0.79–0.93, $P < 0.001$).The predictive performances of both the models on the training and internal test datasets are shown in *Table 2*.

Invasive coronary angiography cohort

A total of 89 subjects were enrolled in the invasive coronary angiography cohort and 87 subjects were included in the study. Two enrollees were not included as one did not undergo their procedure, and the other only underwent a right heart catheterization. We utilized the CAC ≥ 400 model to predict the invasive coronary angiographic evidence of CAD. As shown in the Supplementary material online, *Figure S1*, the model performed well in predicting any angiographic evidence of coronary artery disease with an AUC of 0.73 (95% CI 0.62–0.82, $P < 0.001$). In order to further risk stratify the patients, we used a ≥15% pretest probability as threshold cut-off which also coincides with 2019 ESC pretest probability guidelines[24] for defining intermediate or high-risk range of pretest probability. This model stratified 59 subjects to the intermediate- to high-probability cohort and 28 subjects to the low-probability cohort. Clinical characteristics
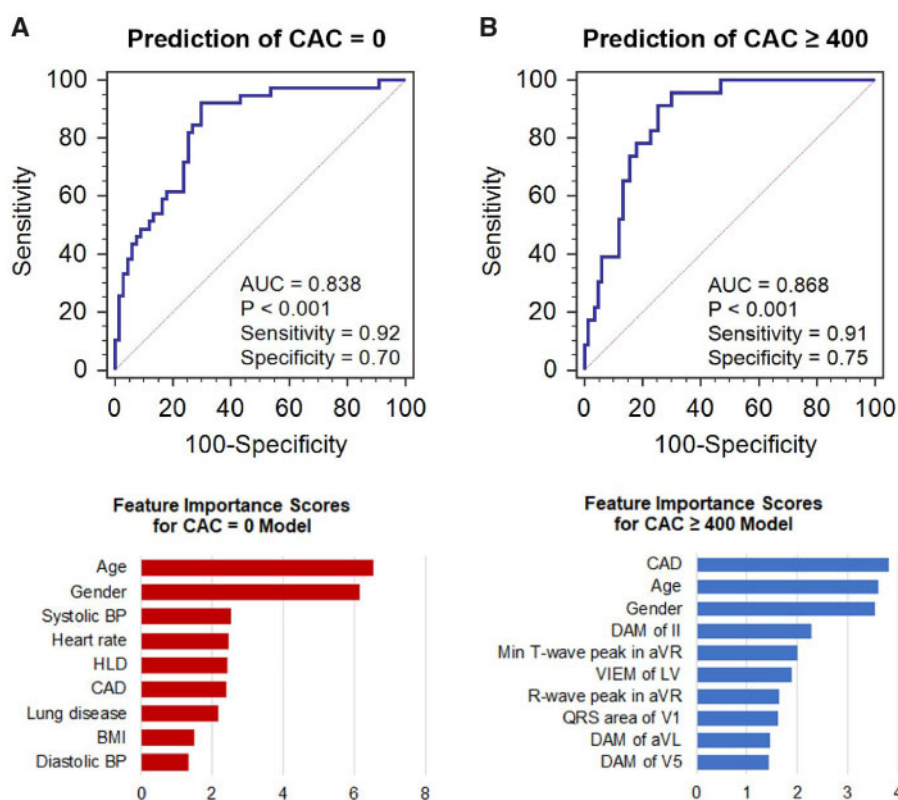
Figure 2 The area under-receiver-operating curve (AUC) of the L2-regularized multivariate logistic regression model predicting the logarithm of the odds of CAC equaling zero (A), and L1-regularized model predicting the logarithm of the odds of CAC ≥ 400 (B). *Y*-axis shows the TP rate (sensitivity) and *X*-axis showing the FP rate (1—specificity). An area under the curve (AUC) > 0.5 indicates better predictive values. The provided sensitivities and specificities are from the optimal cut-off point closest to the top left of the graph. Below each graph are feature importance scores for each model. The CAC ≥ 400 model only displays the top ten features. Signal processed ECG variables are described in the Supplementary material online, *Table S2*.

Table 2 Performance of CAC = 0 and CAC ≥ 400 prediction models on the training and test cohort

Models	Dataset	AUC	Accuracy	F1	Precision	Recall
CAC = 0	Training	0.87	0.79	0.79	0.79	0.79
	Test	0.84	0.75	0.73	0.74	0.75
CAC ≥ 400	Training	0.87	0.86	0.84	0.85	0.86
	Test	0.87	0.81	0.8	0.79	0.81

The number of instances whose classification is estimated correctly are given by the attribute 'Accuracy'. The total number of true positive scores is given by the attribute 'Precision' score. 'Recall' is the total number of true positive instances among all the positive instances. F1 score is the weighted harmonic mean of precision and recall. The 95% confidence intervals for the CAC = 0 Test AUC was 0.75–0.90 and for the CAC ≥ 400 Test AUC was 0.79–0.93.

of the cohort and stratified subsets are shown in *Table 3*. The intermediate- to high-probability cohort were significantly older, mostly male, and had higher rates of prior CAD. Both validated scoring systems displayed significantly higher risk for this cohort as well.

Clinical outcomes are reported in *Table 4*. Participants in the higher-probability cohort had significantly more MACE (*P* = 0.023) with a hazard ratio of 2.09 (95% CI 1.05–4.14) and chance of significant stenosis (*P* = 0.025) with an odds ratio of 2.85 (95% CI 1.12–7.24). The Kaplan–Meier analysis shown in *Figure 3* demonstrates the curves separating over time with the higher-probability cohort having worse outcomes (*P* = 0.035). The ML model performance was better than the CAD Consortium clinical score which did not significantly stratify the cohort as seen in Kaplan–Meier curve in *Figure 3B*. The ML model also performed well in predicting stenosis and need for revascularization with the initial intervention. Any coronary stenosis

Table 3 Baseline characteristics of clinical invasive angiography cohort

	Total (*n* = 87)	Low probability (*n* = 28)	Intermediate to high probability (*n* = 59)	*P*-value
Age (years)	60.3 ± 9.7	55.9 ± 9.1	62.3 ± 9.3	0.003
Female sex	33 (37.9)	22 (78.6)	11 (18.6)	<0.001
Minority race	3 (3.4)	1 (3.6)	2 (3.4)	0.965
Body mass index (kg/m^2)	32.4 ± 7.1	33.9 ± 8.9	31.6 ± 5.9	0.158
Heart rate	70.2 ± 13.4	72.5 ± 12.9	69.1 ± 13.6	0.275
Blood pressure (mmHg)				
Systolic	129.8 ± 18.7	125.0 ± 16.1	132.1 ± 19.5	0.095
Diastolic	75.4 ± 12.7	72.3 ± 12.0	76.9 ± 12.9	0.113
Test indication				0.647
Acute coronary syndrome	45 (51.7)	16 (57.1)	29 (49.2)	
Stable ischaemic heart disease	38 (43.7)	11 (39.3)	27 (45.8)	
Ischaemic cardiomyopathy	2 (2.3)	1 (3.6)	1 (1.7)	
Preoperative evaluation	2 (2.3)	0 (0.0)	2 (3.4)	
CAD consortium clinical probability	30.5 ± 21.0	16.3 ± 10.3	37.3 ± 21.4	<0.001
Smoker				0.373
Current	28 (32.2)	10 (35.7)	18 (30.5)	
Former	24 (27.6)	5 (17.9)	19 (32.2)	
Never	35 (40.2)	13 (46.4)	22 (37.3)	
Obesity, BMI ≥ 30 kg/m^2	51 (58.6)	19 (67.9)	32 (54.2)	0.228
Hypertension	85 (97.7)	27 (96.4)	58 (98.3)	0.585
Hyperlipidaemia	85 (97.7)	27 (96.4)	58 (98.3)	0.585
Diabetes mellitus	39 (44.8)	11 (39.3)	28 (47.5)	0.474
Coronary artery disease	78 (89.7)	20 (71.4)	58 (98.3)	<0.001
Previous PCI	34 (39.1)	9 (32.1)	25 (42.3)	0.361
Previous CABG	1 (1.1)	1 (3.6)	0 (0.0)	0.144
Peripheral artery disease	11 (12.6)	1 (3.6)	10 (17.0)	0.079
Prior stroke/TIA	10 (11.5)	3 (10.7)	7 (11.9)	0.875
Chronic kidney disease	8 (9.2)	3 (10.7)	5 (8.5)	0.736
Chronic lung disease	22 (25.3)	7 (25.0)	15 (25.4)	0.966
Family history of CAD	53 (60.9)	19 (67.9)	34 (57.6)	0.361

Values are counts (%) or mean ± standard deviation.
Minority race is defined as any non-White race as reported by the investigators and family history indicated premature CAD history.
BMI, body mass index; CA, coronary angiogram; CAC, coronary artery calcium; CABG, coronary artery bypass graft; CAD, coronary artery disease; PCI, percutaneous coronary intervention; TIA, transient ischaemic attack.

Table 4 Clinical outcomes of invasive angiography cohort

	Total (n = 87)	Low probability (n = 28)	Intermediate to high probability (n = 59)	P-value
Follow-up time (months)	21.6 ± 4.4	22.0 ± 3.7	21.4 ± 4.8	0.530
Angiogram indication				0.647
Acute coronary syndrome	45 (51.7)	16 (57.1)	29 (49.2)	
Stable ischaemic heart disease	38 (43.7)	11 (36.3)	27 (45.8)	
Preoperative evaluation	2 (2.3)	0 (0.0)	2 (3.4)	
Ischaemic cardiomyopathy	2 (2.3)	1 (3.6)	1 (1.7)	
Any stenosis	60 (69.0)	13 (46.4)	47 (79.7)	0.002
Significant stenosis	55 (63.2)	13 (46.4)	42 (71.2)	0.025
Significantly stenotic coronary artery				
Left main	2 (2.3)	1 (3.6)	1 (1.7)	0.585
Left anterior descending	36 (41.4)	6 (21.4)	30 (50.8)	0.009
Left circumflex	26 (29.9)	7 (25.0)	19 (32.2)	0.493
Right	26 (29.9)	5 (17.9)	21 (35.6)	0.091
Number of significant stenoses				
Single vessel	28 (32.2)	7 (25.0)	21 (35.6)	0.323
Multivessel	26 (29.9)	5 (17.9)	21 (35.6)	0.091
Revascularization	53 (60.9)	10 (35.7)	43 (72.9)	<0.001
PCI	43 (49.4)	10 (35.7)	33 (55.9)	0.078
CABG	10 (11.5)	0 (0.0)	10 (16.9)	0.021
MACE	37 (42.5)	7 (25.0)	30 (50.8)	0.023
Unstable angina	23 (26.4)	4 (14.3)	19 (32.2)	0.077
Myocardial infarction	8 (9.2)	3 (10.7)	5 (8.5)	0.736
Stroke	8 (9.2)	1 (3.6)	7 (11.9)	0.211
CV hospitalization	30 (34.5)	8 (28.6)	22 (37.3)	0.424
All-cause mortality	4 (4.6)	1 (3.6)	3 (5.1)	0.753

Values are counts (%) or mean ± standard deviation.
Any stenosis was reported with any ≥25% lesion and significant stenosis if it was ≥50% in the left main coronary artery or ≥70% elsewhere. Diagonal and obtuse marginal branches were considered left anterior descending and left circumflex coronary arteries respectively. Multivessel stenosis equals two or more significant stenoses or left main stenosis.
Revascularization accounts for intervention performed during the initial catheterization procedure, due to information from the initial procedure, or upon follow-up.
CABG, coronary artery bypass graft; CV, cardiovascular; PCI, percutaneous coronary intervention.

(P = 0.002), severe stenosis (P = 0.025), and revascularization (P < 0.001) were all significantly higher in the intermediate- to high-probability group. Findings of significant multivessel disease including left main disease trended towards a higher rate in the higher probability group (P = 0.091). Furthermore, the model displayed 100% negative predictive value stratifying participants for the outcome of severe multivessel coronary artery disease requiring coronary artery bypass graft (CABG) (P = 0.021).

Discussion

Risk stratifying patients with subclinical or symptomatic CAD continues to be a clinical dilemma for primary care physicians as well as specialists. A recent study found that established ESC guidelines were greatly overestimating pretest probability for obstructive coronary artery disease leading to low-yield diagnostic testing.[27,32] In this *post hoc* substudy of a multicentre, prospective trial, we developed an ML model utilizing common clinical data along with non-invasive, signal-

processed ECG to successfully stratify low- to intermediate-risk patients. We then tested this model with higher-risk subjects demonstrating predictive features for hard, clinical outcomes. Our models were based on CAC which is a growing modality for risk stratification in low- to intermediate-risk patients. The first model attempted to predict subjects with CAC scores of zero, allowing for reliable negative predictive value and showed an AUC of 0.838. Interestingly, clinical features alone were sufficiently robust to develop the model, and the addition of ECG variables did not show incremental value. This suggests that at an early stage of the disease the cardiac electrical wavefront remains relatively spared, and risk factor-based assessments alone can indicate development of early atherosclerosis. However, with later stages of the disease, the development of CAC ≥400 is associated with significant changes in cardiac electrical properties that can be used for non-invasive prediction of subclinical CAD. The model demonstrated a robust AUC of 0.868 (P < 0.001) and sensitivity of 91.3%. A high sensitivity is imperative so patients with obstructive coronary artery disease are not left undiagnosed. Moreover, the model predicting CAC scores ≥400 also predicted a
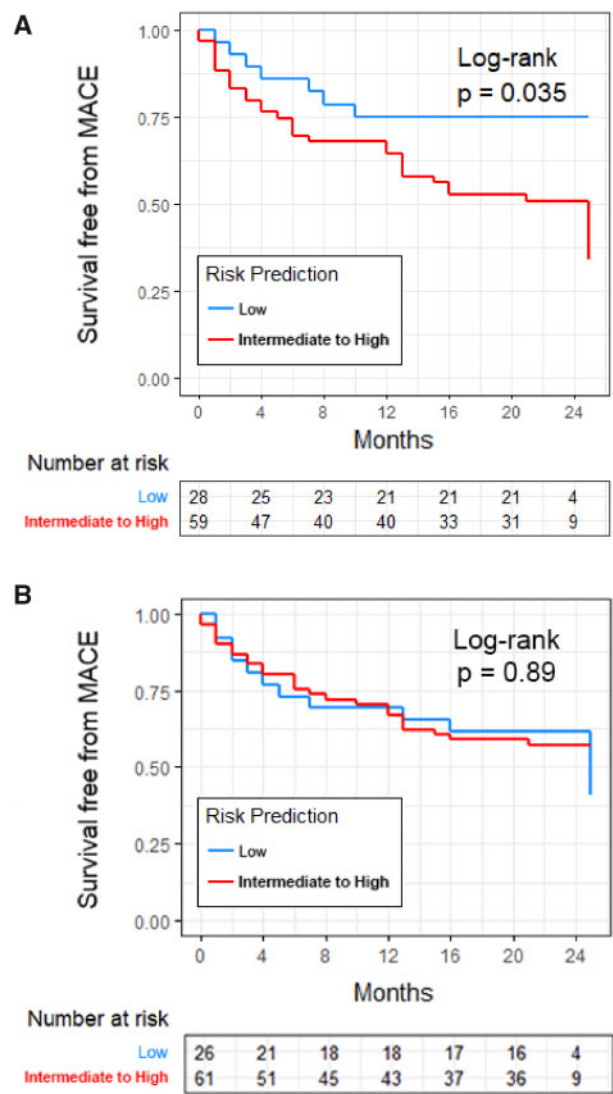
Figure 3 The Kaplan–Meier curves of the primary outcome of composite MACE including unstable angina, myocardial infarction, CV hospitalization, stroke, and all-cause mortality. The (*A*) CAC ≥ 400 model significantly stratified the invasive catheterization population in terms of MACE using a guideline directed cut-off of 15% pretest probability. The outcomes rates were compared to the validated CAD Consortium clinic score (*B*) which did not significantly stratify the groups based on MACE.

higher burden of CAD and the need for revascularization. Notably, it demonstrated a 100% negative predictive rate for the requirement of CABG. This model successfully predicted MACE which is consistent with the known prognostic value of CAC scores ≥400.[5,33,34]

A key element of our design was to use easily attainable variables along with quick, low-risk modalities. High-level classifiers included age, gender, SBP, CAD, and multiple spECG parameters. The spECG mimics the process of a conventional 12-lead ECG and can potentially be performed in the office setting. The continuous waveform transformation allowed for extraction of hundreds of variables. A study in 2018, demonstrated that spECG could predict abnormal myocardial relaxation and diastolic dysfunction with an AUC of 0.91 (95% CI 0.86–0.95).[15] Conversely, we trained our model to predict the

outcomes from CAC because of its growing popularity and clinical utility. Though CAC is a validated risk stratification tool, it is appropriate for only certain populations and comes with limitations. It is recommended for low to intermediate cardiovascular risk patients and has limited utility in only low-risk and higher-risk populations.[35,36] Furthermore, a CAC scan exposes the patient to 1–2 mSv of radiation. Decreasing this exposure with a lower-energy photon can lead to increased image noise, decreased calcium attenuation, and falsely elevated CAC scores.[37,38] Other limitations of CAC scoring and specifically the Agatston score are the need for manual measurements causing interobserver variability and proven inter-scan variability, respectively.[39,40] The use of our model prior to referral for CAC scoring may assist clinicians in the appropriateness of testing preventing these scanning risks.

With the prediction of any coronary stenosis and clinical outcomes, our model potentially has utility for risk stratification for all comers. The 15% probability threshold used reflects symptomatic intermediate- to high-risk patients who benefit the most from non-invasive testing.[27] These patients in our model, if stratified into the low-probability CAC group, had a significantly reduced hazard ratio of 0.48 (95% CI 0.24–0.95) for MACE and only 35% odds of having significant stenosis compared to the higher probability group. Additionally, this model performed better than the CAD Consortium clinical score which is well validated.[28] Even with these promising results, the event rate in the low probability cohort was 25% with almost half of the subjects having significant stenosis leaving room for further optimization in the future.

There are important limitations to this study. Because the training and testing sets were gained as a *post hoc* study of another trial, they contained a relatively small number of subjects to develop the models. Although our *post hoc* analysis justified the sample size, further confirmation with larger samples and the use of other ML techniques should be explored. Additionally, participants included to develop the models were referred for CTA as opposed to a specific CAC scan. Enrolled participants were taken consecutively regardless of indication seeking a generalizable sample. Nonetheless, this sample likely differs, especially concerning symptomatology, to that of patients referred for specific CAC scans. Ethnicity may also have been an important factor. Although we included subjects from different regions including Los Angeles, New York City, and West Virginia, CAC score outcomes vary based on different racial and ethnic groups.[41] The sample from Los Angeles did not report minority status, so this variable was not used to develop the model. The invasive cohort from West Virginia was ethnically homogenous with only three minority participants decreasing generalizability from this test. Considering the event rates in the low probability group along with the sample limitations, more work with larger, more heterogeneous populations is needed to further optimize the model. Finally, we utilized guideline-based CAC score categories with binary outcome predictions. Future work would require focusing on direct regression of the absolute CAC values in larger sample sizes potentially with the use of more robust techniques like convoluted neural networks applied directly to the ECG waveforms.

In conclusion, our ML model utilizing ECG and simple clinical characteristics showed the ability to predict severe CAC scores in symptomatic, low and intermediate pretest probability patients. Moreover, even in symptomatic patients with high burdens of CAD, our model predicted clinical outcomes including MACE and significant coronary stenosis in patients undergoing IA. These preliminary data suggest the potential value of ML techniques in extracting information from electrocardiographic data and clinical variables to predict CAC score categories and the need for further trials with larger populations to further improve the model prediction and clinical utilization.

Supplementary material

Supplementary material is available at *European Heart Journal – Digital Health* online.

Data availability

The data underlying this article cannot be shared publicly due to the use of protected health information. The data will be shared on reasonable request to the corresponding author.

Conflict of interest: P.P.S. has served as a consultant to HeartSciences, Ultromics, and Kencor Health. N.K. is affiliated with and receives salary from a department funded by Philips Healthcare; Asahi KASEI Corporation; Inter Reha Co., Ltd; and Toho Holdings Co., Ltd based on collaborative research agreements. All other authors have reported that they have no relationships relevant to the contents of this paper to disclose.

References

1. Piepoli MF, Hoes AW, Agewall S, Albus C, Brotons C, Catapano AL, Cooney MT, Corrà U, Cosyns B, Deaton C, Graham I, Hall MS, Hobbs FDR, Løchen ML, Löllgen H, Marques-Vidal P, Perk J, Prescott E, Redon J, Richter DJ, Sattar N, Smulders Y, Tiberi M, Bart van der Worp H, van Dis I, Verschuren WMM. 2016 European Guidelines on cardiovascular disease prevention in clinical practice: the Sixth Joint Task Force of the European Society of Cardiology and Other Societies on Cardiovascular Disease Prevention in Clinical Practice (constituted by representatives of 10 societies and by invited experts) Developed with the special contribution of the European Association for Cardiovascular Prevention & Rehabilitation (EACPR). *Atherosclerosis* 2016;**252**:207–274.
2. Tinana A, Mintz GS, Weissman NJ. Volumetric intravascular ultrasound quantification of the amount of atherosclerosis and calcium in nonstenotic arterial segments. *Am J Cardiol* 2002;**89**:757–760.
3. Arnett DK, Blumenthal RS, Albert MA, Buroker AB, Goldberger ZD, Hahn EJ, Himmelfarb CD, Khera A, Lloyd-Jones D, McEvoy JW, Michos ED, Miedema MD, Muñoz D, Smith SC Jr, Virani SS, Williams KA Sr, Yeboah J, Ziaeian B. 2019 ACC/AHA Guideline on the primary prevention of cardiovascular disease: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *J Am Coll Cardiol* 2019;**74**:e177–e232.
4. Haberl R, Becker A, Leber A, Knez A, Becker C, Lang C, Brüning R, Reiser M, Steinbeck G. Correlation of coronary calcification and angiographically documented stenoses in patients with suspected coronary artery disease: results of 1,764 patients. *J Am Coll Cardiol* 2001;**37**:451–457.
5. Budoff MJ, Mayrhofer T, Ferencik M, Bittner D, Lee KL, Lu MT, Coles A, Jang J, Krishnam M, Douglas PS, Hoffmann U; PROMISE Investigators. Prognostic value of coronary artery calcium in the PROMISE study (Prospective Multicenter Imaging Study for Evaluation of Chest Pain). *Circulation* 2017;**136**:1993–2005.
6. Silber S. Comparison of spiral and electron beam tomography in the evaluation of coronary calcification in asymptomatic persons. *Int J Cardiol* 2002;**82**:297–298; author reply 9.
7. Havel M, Koranda P, Kincl V, Quinn L, Kaminek M. Additional value of the coronary artery calcium score in patients for whom myocardial perfusion imaging is challenging. *Kardiol Pol* 2019;**77**:458–464.
8. Kang MG, Kang Y, Jang HG, Kim K, Koh JS, Park JR, Hwang SJ, Hwang JY, Bae JS, Ahn JH, Jang JY, Park Y, Jeong YH, Kwak CH, Park HW. Coronary artery calcium score in predicting periprocedural myocardial infarction in patients undergoing an elective percutaneous coronary intervention. *Coron Artery Dis* 2018;**29**:589–596.
9. Vinter N, Christesen AMS, Mortensen LS, Urbonaviciene G, Lindholt J, Johnsen SP, Frost L. Coronary artery calcium score and the long-term risk of atrial fibrillation in patients undergoing non-contrast cardiac computed tomography for

suspected coronary artery disease: a Danish registry-based cohort study. *Eur Heart J Cardiovasc Imaging* 2018;**19**:926–932.

10. Mahesh M, Zimmerman SL, Fishman EK. Radiation dose shift in relative proportion: the case of coronary artery calcium studies. *J Am Coll Radiol* 2014;**11**:634–635.

11. Parikh P, Shah N, Ahmed H, Schoenhagen P, Fares M. Coronary artery calcium scoring: Its practicality and clinical utility in primary care. *Cleve Clin J Med* 2018;**85**:707–716.

12. Rothberg MB. Coronary artery calcium scoring: a valuable tool in primary care. *Cleve Clin J Med* 2018;**85**:717–719.

13. Kagiyama N, Piccirilli M, Yanamala N, Shrestha S, Farjo PD, Casaclang-Verzosa G, Tarhuni WM, Nezarat N, Budoff MJ, Narula J, Sengupta PP. Machine learning assessment of left ventricular diastolic function based on electrocardiographic features. *J Am Coll Cardiol* 2020;**76**:930–941.

14. 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.

15. Sengupta PP, Kulkarni H, Narula J. Prediction of abnormal myocardial relaxation from signal processed surface ECG. *J Am Coll Cardiol* 2018;**71**:1650–1660.

16. Levine GN, Bates ER, Blankenship JC, Bailey SR, Bittl JA, Cercek B, Chambers CE, Ellis SG, Guyton RA, Hollenberg SM, Khot UN, Lange RA, Mauri L, Mehran R, Moussa ID, Mukherjee D, Nallamothu BK, Ting HH; American College of Cardiology Foundation; American Heart Association Task Force on Practice Guidelines; Society for Cardiovascular Angiography and Interventions. 2011 ACCF/AHA/SCAI Guideline for percutaneous coronary intervention. A report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines and the Society for Cardiovascular Angiography and Interventions. *J Am Coll Cardiol* 2011;**58**:e44–e122.

17. Crowe JA, Gibson NM, Woolfson MS, Somekh MG. Wavelet transform as a potential tool for ECG analysis and compression. *J Biomed Eng* 1992;**14**:268–272.

18. Minhas FU, Arif M. Robust electrocardiogram (ECG) beat classification using discrete wavelet transform. *Physiol Meas* 2008;**29**:555–570.

19. Addison PS. Wavelet transforms and the ECG: a review. *Physiol Meas* 2005;**26**:R155–R199.

20. Kursa MB, Rudnicki WR. Feature selection with the Boruta package. *J Stat Softw* 2010;**36**:1–13.

21. Maalouf M, Trafalis TB. Robust weighted kernel logistic regression in imbalanced and rare events data. *Comput Stat Data Anal* 2011;**55**:168–183.

22. Maalouf M, Siddiqi M. Weighted logistic regression for large-scale imbalanced and rare events data. *Knowl-Based Syst* 2014;**59**:142–148.

23. Ravikumar P, Wainwright MJ, Lafferty JD. High-dimensional ising model selection using 1-regularized logistic regression. *Ann Stat* 2010;**38**:1287–1319.

24. Muchlinski D, Siroky D, He J, Kocher M. Comparing random forest with logistic regression for predicting class-imbalanced civil war onset data. *Polit Anal* 2016;**24**:87–103.

25. Ng AY. Feature selection, L1 vs L2 regularization, and rotational invariance. In: *Proceedings of the 21st International Conference on Machine Learning.* 2004. July 4-8, 2004 Bannf, Alberta, Canada.

26. Ying X. An overview of overfitting and its solutions. *J Phys Conf Ser* 2019;**1168**:022022.

27. Knuuti J, Wijns W, Saraste A, Capodanno D, Barbato E, Funck-Brentano C, Prescott E, Storey RF, Deaton C, Cuisset T, Agewall S, Dickstein K, Edvardsen T, Escaned J, Gersh BJ, Svitil P, Gilard M, Hasdai D, Hatala R, Mahfoud F, Masip J, Muneretto C, Valgimigli M, Achenbach S, Bax JJ; ESC Scientific Document Group. 2019 ESC Guidelines for the diagnosis and management of chronic coronary syndromes: the Task Force for the diagnosis and management of chronic coronary syndromes of the European Society of Cardiology (ESC). *Eur Heart J* 2019;**41**:407–477.

28. Genders TS, Steyerberg EW, Hunink MG, Nieman K, Galema TW, Mollet NR, de Feyter PJ, Krestin GP, Alkadhi H, Leschka S, Desbiolles L, Meijs MF, Cramer MJ, Knuuti J, Kajander S, Bogaert J, Goetschalckx K, Cademartiri F, Maffei E, Martini C, Seitun S, Aldrovandi A, Wildermuth S, Stinn B, Fornaro J, Feuchtner G, De Zordo T, Auer T, Plank F, Friedrich G, Pugliese F, Petersen SE, Davies LC, Schoepf UJ, Rowe GW, van Mieghem CA, van Driessche L, Sinitsyn V, Gopalan D, Nikolaou K, Bamberg F, Cury RC, Battle J, Maurovich-Horvat P, Bartykowszki A, Merkely B, Becker D, Hadamitzky M, Hausleiter J, Dewey M, Zimmermann E, Laule M. Prediction model to estimate presence of coronary artery disease: retrospective pooled analysis of existing cohorts. *BMJ* 2012;**344**:e3485.

29. Juarez-Orozco LE, Saraste A, Capodanno D, Prescott E, Ballo H, Bax JJ, Wijns W, Knuuti J. Impact of a decreasing pre-test probability on the performance of diagnostic tests for coronary artery disease. *Eur Heart J Cardiovasc Imaging* 2019;**20**:1198–1207.

30. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol* 1996;**49**:1373–1379.

31. Vittinghoff E, McCulloch CE. Relaxing the rule of ten events per variable in logistic and Cox regression. *Am J Epidemiol* 2007;**165**:710–718.

32. Foldyna B, Udelson JE, Kárády J, Banerji D, Lu MT, Mayrhofer T, Bittner DO, Meyersohn NM, Emami H, Genders TSS, Fordyce CB, Ferencik M, Douglas PS, Hoffmann U. Pretest probability for patients with suspected obstructive coronary artery disease: re-evaluating Diamond–Forrester for the contemporary era and clinical implications: insights from the PROMISE trial. *Eur Heart J Cardiovasc Imaging* 2018;**20**:574–581.

33. Yamamoto H, Kitagawa T, Kunita E, Utsunomiya H, Senoo A, Nakamoto Y, Kihara Y. Impact of the coronary artery calcium score on mid- to long-term cardiovascular mortality and morbidity measured with coronary computed tomography angiography. *Circ J* 2018;**82**:2342–2349.

34. Rijlaarsdam-Hermsen D, Lo-Kioeng-Shioe MS, Kuijpers D, van Domburg RT, Deckers JW, van Dijkman PRM. Prognostic value of the coronary artery calcium score in suspected coronary artery disease: a study of 644 symptomatic patients. *Neth Heart J* 2020;**28**:44–50.

35. Kim WJ, Kwon CH, Han S, Lee WS, Kang JW, Ahn JM, Lee JY, Park DW, Kang SJ, Lee SW, Kim YH, Lee CW, Park SW, Park SJ. Role of coronary artery calcium scoring in detection of coronary artery disease according to Framingham Risk score in populations with low to intermediate risks. *J Korean Med Sci* 2016;**31**:902–908.

36. Hecht H, Blaha MJ, Berman DS, Nasir K, Budoff M, Leipsic J, Blankstein R, Narula J, Rumberger J, Shaw LJ. Clinical indications for coronary artery calcium scoring in asymptomatic patients: expert consensus statement from the Society of Cardiovascular Computed Tomography. *J Cardiovasc Comput Tomogr* 2017;**11**:157–168.

37. Deprez FC, Vlassenbroek A, Ghaye B, Raaijmakers R, Coche E. Controversies about effects of low-kilovoltage MDCT acquisition on Agatston calcium scoring. *J Cardiovasc Comput Tomogr* 2013;**7**:58–61.

38. Blaha MJ, Mortensen MB, Kianoush S, Tota-Maharaj R, Cainzos-Achirica M. Coronary artery calcium scoring: is it time for a change in methodology? *JACC Cardiovasc Imaging* 2017;**10**:923–937.

39. Wang W, Wang H, Chen Q, Zhou Z, Wang R, Wang H, Zhang N, Chen Y, Sun Z, Xu L. Coronary artery calcium score quantification using a deep-learning algorithm. *Clin Radiol* 2020;**75**:237.e11–e16.

40. Alluri K, Joshi PH, Henry TS, Blumenthal RS, Nasir K, Blaha MJ. Scoring of coronary artery calcium scans: history, assumptions, current limitations, and future directions. *Atherosclerosis* 2015;**239**:109–117.

41. Lee JH, Ó Hartaigh B, Han D, Park HE, Choi SY, Sung J, Chang HJ. Reassessing the usefulness of coronary artery calcium score among varying racial and ethnic groups by geographic locations: relevance of the Korea Initiatives on Coronary Artery Calcification Registry. *J Cardiovasc Ultrasound* 2015;**23**:195–203.

Authors
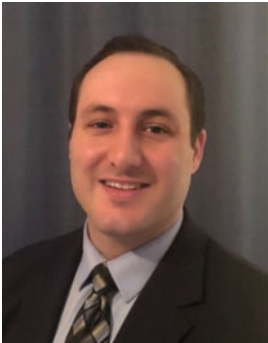
Biography: Dr Peter D. Farjo is a third-year general cardiology fellow at West Virginia University School of Medicine. He also completed his medical school and internal medicine training at West Virginia University. His research interests include machine learning analyses in the setting of arrhythmia. Dr Farjo is pursuing a subspecialty fellowship in clinical cardiac electrophysiology.

Biography: Dr. Naveena Yanamala is the Principal Data Scientist in the Heart & Vascular Institute at West Virginia University Medicine. In addition, she currently holds appointments at the Carnegie Mellon University and the International Institute for Information Technology. She has 13+ years of experience in conducting effective interdisciplinary research at the intersection of biology, health, and computation. She received her PhD in Integrative Systems Biology in 2009 from University of Pittsburgh and her MS in Information Technology from International Institute of Information Technology in Hyderabad, India.