

Machine learning of clinical variables and coronary artery calcium scoring for the prediction of obstructive coronary artery disease on coronary computed tomography angiography: analysis from the CONFIRM registry

Subhi J. Al'Aref ^{1†}, Gabriel Maliakal^{1†}, Gurpreet Singh ¹, Alexander R. van Rosendaal¹, Xiaoyue Ma², Zhuoran Xu ¹, Omar Al Hussein Alawamlh¹, Benjamin Lee ¹, Mohit Pandey¹, Stephan Achenbach³, Mouaz H. Al-Mallah ⁴, Daniele Andreini⁵, Jeroen J. Bax ⁶, Daniel S. Berman⁷, Matthew J. Budoff⁸, Filippo Cademartiri⁹, Tracy Q. Callister¹⁰, Hyuk-Jae Chang¹¹, Kavitha Chinnaiyan¹², Benjamin J.W. Chow¹³, Ricardo C. Cury¹⁴, Augustin DeLago¹⁵, Gudrun Feuchtner ¹⁶, Martin Hadamitzky¹⁷, Joerg Hausleiter¹⁸, Philipp A. Kaufmann¹⁹, Yong-Jin Kim²⁰, Jonathon A. Leipsic²¹, Erica Maffei²², Hugo Marques²³, Pedro de Araújo Gonçalves²³, Gianluca Pontone⁵, Gilbert L. Raff¹², Ronen Rubinshtein²⁴, Todd C. Villines²⁵, Heidi Gransar⁷, Yao Lu², Erica C. Jones¹, Jessica M. Peña¹, Fay Y. Lin ¹, James K. Min^{1*}, and Leslee J. Shaw¹

¹Dalio Institute of Cardiovascular Imaging, Weill Cornell Medicine and NewYork-Presbyterian Hospital, New York, NY, USA; ²Department of Healthcare Policy and Research, New York-Presbyterian Hospital and the Weill Cornell Medical College, New York, NY, USA; ³Department of Cardiology, Friedrich-Alexander-University Erlangen-Nuremberg, Germany; ⁴Houston Methodist DeBakey Heart & Vascular Center, Houston Methodist Hospital, TX, USA; ⁵Centro Cardiologico Monzino, IRCCS Milan, Italy; ⁶Department of Cardiology, Leiden University Medical Center, Leiden, The Netherlands; ⁷Department of Imaging and Medicine, Cedars Sinai Medical Center, Los Angeles, CA, USA; ⁸Department of Medicine, Los Angeles Biomedical Research Institute, Torrance, CA, USA; ⁹Cardiovascular Imaging Center, SDN IRCCS, Naples, Italy; ¹⁰Tennessee Heart and Vascular Institute, Hendersonville, TN, USA; ¹¹Division of Cardiology, Severance Cardiovascular Hospital and Severance Biomedical Science Institute, Yonsei University College of Medicine, Yonsei University Health System, Seoul, South Korea; ¹²Department of Cardiology, William Beaumont Hospital, Royal Oak, MI, USA; ¹³Department of Medicine and Radiology, University of Ottawa, ON, Canada; ¹⁴Department of Radiology, Miami Cardiac and Vascular Institute, Miami, FL, USA; ¹⁵Capitol Cardiology Associates, Albany, NY, USA; ¹⁶Department of Radiology, Medical University of Innsbruck, Innsbruck, Austria; ¹⁷Department of Radiology and Nuclear Medicine, German Heart Center Munich, Munich, Germany; ¹⁸Medizinische Klinik I der Ludwig-Maximilians-Universität München, Munich, Germany; ¹⁹Department of Nuclear Medicine, University Hospital, Zurich, Switzerland and University of Zurich, Switzerland; ²⁰Seoul National University Hospital, Seoul, South Korea; ²¹Department of Medicine and Radiology, University of British Columbia, Vancouver, BC, Canada; ²²Department of Radiology, Area Vasta 1/ASUR Marche, Urbino, Italy; ²³UNICA, Unit of Cardiovascular Imaging, Hospital da Luz, Lisboa, Portugal; ²⁴Department of Cardiology at the Lady Davis Carmel Medical Center, The Ruth and Bruce Rappaport School of Medicine, Technion-Israel Institute of Technology, Haifa, Israel; and ²⁵Division of Cardiovascular Medicine, Department of Medicine, University of Virginia Health System, Charlottesville, VA, USA

Received 9 January 2019; revised 19 March 2019; editorial decision 23 July 2019; accepted 20 August 2019; online publish-ahead-of-print 12 September 2019

See page 368 for the editorial comment on this article (doi: 10.1093/eurheartj/ehz708)

Aims

Symptom-based pretest probability scores that estimate the likelihood of obstructive coronary artery disease (CAD) in stable chest pain have moderate accuracy. We sought to develop a machine learning (ML) model,

* Corresponding author. Tel: +1 646 962 6266, Email: jkm2001@med.cornell.edu

† The first two authors contributed equally to the study.

Published on behalf of the European Society of Cardiology. All rights reserved. © The Author(s) 2019. For permissions, please email: journals.permissions@oup.com.

utilizing clinical factors and the coronary artery calcium score (CACS), to predict the presence of obstructive CAD on coronary computed tomography angiography (CCTA).

Methods and results

The study screened 35 281 participants enrolled in the CONFIRM registry, who underwent ≥ 64 detector row CCTA evaluation because of either suspected or previously established CAD. A boosted ensemble algorithm (XGBoost) was used, with data split into a training set (80%) on which 10-fold cross-validation was done and a test set (20%). Performance was assessed of the (1) ML model (using 25 clinical and demographic features), (2) ML + CACS, (3) CAD consortium clinical score, (4) CAD consortium clinical score + CACS, and (5) updated Diamond-Forrester (UDF) score. The study population comprised of 13 054 patients, of whom 2380 (18.2%) had obstructive CAD ($\geq 50\%$ stenosis). Machine learning with CACS produced the best performance [area under the curve (AUC) of 0.881] compared with ML alone (AUC of 0.773), CAD consortium clinical score (AUC of 0.734), and with CACS (AUC of 0.866) and UDF (AUC of 0.682), $P < 0.05$ for all comparisons. CACS, age, and gender were the highest ranking features.

Conclusion

A ML model incorporating clinical features in addition to CACS can accurately estimate the pretest likelihood of obstructive CAD on CCTA. In clinical practice, the utilization of such an approach could improve risk stratification and help guide downstream management.

Keywords

Coronary artery disease • Coronary artery calcium score • Machine learning • Coronary computed tomography angiography

Introduction

Coronary computed tomography angiography (CCTA) has emerged as an accurate method for the non-invasive evaluation of coronary artery disease (CAD).^{1,2} Numerous studies have shown that the absence of CAD on CCTA conveys very low risk for incident cardiovascular events while there is a graded relationship between increasing CAD burden and cardiovascular risk.^{3,4} Practice guidelines for the management of stable chest pain from the European Society of Cardiology (ESC), American Heart Association (AHA), and American College of Cardiology (ACC) are congruent in their recommendations for the use of CCTA as a primary or secondary diagnostic option in symptomatic individuals deemed to be at an intermediate pretest likelihood of having obstructive CAD.^{5,6} However, in day-to-day clinical practice, a significant number of individuals undergoing CCTA have minimal or no CAD.^{7,8} As a direct consequence of the expanding use of CCTA, there is a growing interest within the medical community regarding ways to optimize patient selection with the goal of improving diagnostic yield and cost-effectiveness of CCTA utilization within the context of clinical practice.⁹

A recent approach has sought to improve risk stratification measures in order to streamline patient selection for CCTA performance. Numerous population-derived risk scores have been developed in order to estimate the pretest likelihood of having CAD, such as the updated Diamond and Forrester (UDF) score and the CAD consortium clinical score.^{10–12} The PROMISE and SCOT-HEART trials were performed within a contemporary cohort of stable chest pain individuals, and yet within the PROMISE trial only 10.7% of participants, who were enrolled in the CCTA arm, had obstructive CAD despite a mean pretest likelihood of $53.3 \pm 21.4\%$.¹³ SCOT-HEART, on the other hand, showed that 25% of participants undergoing CCTA had obstructive CAD.¹⁴ Therefore, there continues to be a need for better pretest assessment tools in order to improve patient selection for

CCTA or other diagnostic tests. Coronary artery calcium (CAC), providing a specific marker of coronary atherosclerosis, has been shown to provide incremental predictive power over clinical pretest probability (PTP) assessments regarding the extent and severity of angiographically significant CAD in symptomatic patients.¹⁵ In this study, we sought to develop a machine learning (ML) model, utilizing readily available clinical factors to predict patients likely to have obstructive CAD on CCTA and to evaluate its effectiveness alone and in combination with the coronary artery calcium score (CACS) using a contemporary, international and multi-ethnic cohort of individuals undergoing CCTA evaluation for CAD detection.

Methods

Study population

The COronary CT Angiography Evaluation For Clinical Outcomes: An International Multicenter (CONFIRM) registry is a prospective, observational registry that enrolled patients in 12 medical centres across six countries (Canada, Germany, Italy, Korea, Switzerland, and the United States).¹⁶ The CONFIRM registry was designed to include demographic, clinical and imaging parameters for patients who underwent ≥ 64 -detector row CCTA evaluation because of either suspected or previously established CAD. Of the total cohort, 13 054 patients (≥ 30 years of age) were identified for whom clinical data and the CACS were available. Patients with prior CAD or revascularization (percutaneous or surgical) were excluded. Patient consent or a waiver of informed consent was obtained at each site in keeping with site-specific regulations and all sites had approval of respective institutional review boards. The study complied with the Declaration of Helsinki.

Prior to CCTA acquisition, demographic and clinical information was prospectively collected for all patients included in the CONFIRM registry. Specifically, the presence of significant cardiovascular risk factors such as diabetes mellitus, hypertension, hyperlipidaemia, smoking status, family

history of premature CAD, and baseline cholesterol values were documented.

Coronary computed tomography angiography and coronary artery calcium scanning

Coronary computed tomography angiography image acquisition and processing, as well as coronary artery calcium scanning, were performed in accordance with the guidelines outlined by the Society of Cardiovascular Computed Tomography.^{17–19} While there were no restrictions in scanner type (single-source, dual-source) or brand (Lightspeed VCT, GE Healthcare, Milwaukee, WI, USA; Somatom Definition CT, Siemens, Erlangen, Germany), all machines were required to possess a minimum of 64-detector rows. Level III-equivalent readers evaluated all patient scans and determined the extent of CAD in addition to providing a CACS using the Agatston method. Such a method is semi-automated to calculate a weighted sum of the area of coronary calcification, wherein each calcified area is multiplied by a local density factor determined by the Hounsfield unit (HU) of the calcium (0: 0–129 HU; 1: 130–199 HU; 2: 200–299 HU; 3: 300–399 HU; 4: >400 HU). The outcome of the present study was the presence of obstructive CAD on CCTA, defined as the detection of ≥50% diameter stenosis in any of the four major epicardial coronary arteries. A sensitivity analysis was further performed with the definition of obstructive CAD set at ≥70% diameter stenosis.

Machine learning

Patients included in our analyses were characterized by a total of 25 readily available demographic and clinical variables, including age, gender, risk factors (including diabetes mellitus, hypertension, dyslipidaemia), and baseline cholesterol levels (including total cholesterol, LDL and HDL values). Correlation coefficients between variables were obtained and are shown in [Supplementary material online, Figure S1](#). An ensemble ML algorithm was constructed to classify patients on the basis of the presence of obstructive CAD. Machine learning techniques were implemented in Python using open-source libraries.

A gradient boosting machine learning algorithm (XGBoost) was employed for a binary classification task based on the presence or absence of obstructive CAD. XGBoost is a novel boosting tree-based ensemble algorithm which has gained wide popularity in the ML community. XGBoost outlines the creation of classification and regression trees, in which classification accuracy is iteratively improved one level at a time through optimization of a customized objective function—an instance of a process otherwise known as ‘boosting’.²⁰ This algorithm was employed due to its state-of-the-art accuracy; ability to employ both continuous and categorical inputs, without need for scaling or other pre-processing modifications; capacity for handling of sparsity; interpretability; and lastly, high degree of internal optimization and relatively modest computational cost. Overall, the original dataset was randomly split into training (75%) and a held-out validation (25%) set, such that the ratio of obstructive to non-obstructive CAD was maintained across both the training and validation subsets. Further, the training set was divided into 10 equally sized folds roughly maintaining the ratio of event to non-events seen in the training set to select optimal model hyper-parameters by grid search through 10-fold cross-validation. Model hyper-parameters (e.g. number of trees, depth of each tree) were fine-tuned using 10-fold cross-validation on the training set. Cross-validation is an iterative process whereby the training data is partitioned into roughly equally sized subsets (e.g. 10 such subsets in 10-fold cross-validation), with training occurring using all but one of these subsets and validation being performed on that which is remaining. Employing such a tactic during the training phase can

Table 1 Participant-level characteristics of the study cohort

| Variables | Values |
|----------------------------------|--------------|
| Age (mean ± SD) | 58.0 ± 11.4 |
| Male participants (%) | 54% |
| Hypertension (%) | 52.9% |
| Hyperlipidaemia (%) | 56.9% |
| Diabetes mellitus (%) | 14.2% |
| Current smoker (%) | 17.4% |
| Mean serum creatinine (mg/dL) | 0.97 ± 0.7 |
| Mean total cholesterol (mg/dL) | 189.1 ± 43.6 |
| Low density lipoprotein (mg/dL) | 115.7 ± 36.8 |
| High density lipoprotein (mg/dL) | 52.2 ± 16.0 |
| Chest pain (%) | 69.5% |
| Shortness of breath (%) | 21.0% |

be beneficial for numerous reasons which have been well-characterized elsewhere, but its primary use is in empirically determining optimal model hyper-parameters without recourse to the validation set.²¹ Finally, classification performance of the ML model was measured using the area under the curve (AUC) and the associated 95% confidence interval (CI) and reported for the held-out validation set. Feature ranking was obtained by computing Shapley Additive Explanation values (SHAP), as previously described.²²

Statistical analysis

Performance of the ML model to classify participants was compared with commonly employed prediction scores such as the CAD consortium clinical score and the updated Diamond-Forrester (UDF) score. Further, CACS was added to the ML model and the CAD consortium clinical score given its widespread use as a screening modality. Calibration of the ML model (with and without CACS) was evaluated using the calibration slope and the Brier score (on a scale ranging from 0 to 1). The calibration slope was obtained by fitting a linear regression equation on the mean predicted probabilities vs. fraction of positives, and calculating its slope. The Brier score, on the other hand, calculates the difference between the estimated and observed risk for occurrence of obstructive CAD, with values closer to 0 indicating better calibration. Continuous net reclassification index (NRI) was performed in order to quantify how well the ML models reclassified subjects, either appropriately or inappropriately, compared with the traditional UDF score and the CAD consortium clinical score. Finally, continuous variables were expressed as the mean ± 1 SD, while categorical variables were expressed as counts (percentages) of the total population. Comparisons were considered statistically significant based on a two-sided *P*-value of <0.05.

Results

A total of 13 054 patients met the inclusion criteria and were included in the analysis. The occurrence of obstructive CAD was 18.2% (2380/13 054) within the studied cohort. Mean age 58.0 ± 11.4 and 54% were male patients. Hypertension (52.6%) and hyperlipidaemia (56.9%) were the two prevalent risk factors, while diabetes mellitus was present in 14.2% of the population and 17.4% were

active smokers. At the time of enrolment, mean total cholesterol (in mg/dL) was 189.1 ± 43.6 , mean LDL was 115.7 ± 36.8 and mean HDL

was 52.2 ± 16.0 . Of the total cohort, 68.0% had chest pain while 21.0% had shortness of breath as the presenting symptoms (Table 1).

Prediction of obstructive coronary artery disease on coronary computed tomography angiography

Machine learning produced the best performance in terms of predicting individuals with obstructive CAD, with an AUC of 0.773 (95% CI 0.757–0.791) compared with CAD consortium clinical score (AUC 0.734, 95% CI 0.717–0.751), and UDF score (AUC 0.682, 95% CI 0.662–0.702), $P < 0.05$ for all comparisons. With the addition of CACS, both ML model and CAD consortium clinical scores significantly improved in terms of prediction of the occurrence of obstructive vs. non-obstructive CAD (AUC 0.881, 95% CI 0.869–0.895 and an AUC 0.866, 95% CI 0.852–0.879, respectively) (Figure 1). Sensitivity analysis was performed with obstructive CAD defined as $\geq 70\%$ diameter stenosis with no significant change in discriminative performance across the five models (Supplementary material online, Figure S2). The sensitivity, specificity, positive predictive value, negative predictive value, and accuracy for the prediction of obstructive CAD (at a probability threshold of 15%) were 78.0%, 62.8%, 31.9%, 92.8%, and 65.6% for Model 1 (ML) and 80.0%, 81.5%, 49.1%, 94.8%, and 81.3% for Model 2 (ML + CACS), respectively.

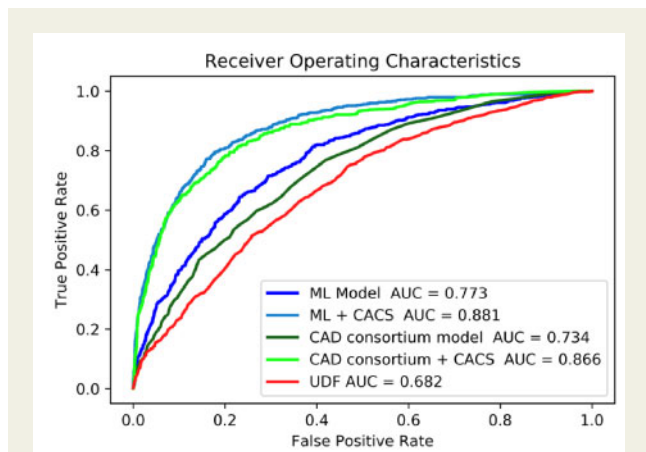


Figure 1 Area under the curve as a measure of individual model performance for the prediction of obstructive coronary artery disease on coronary computed tomography angiography. AUC, area under the curve; CACS, coronary artery calcium score; CAD, coronary artery disease; ML, machine learning; UDF, updated Diamond-Forrester score.

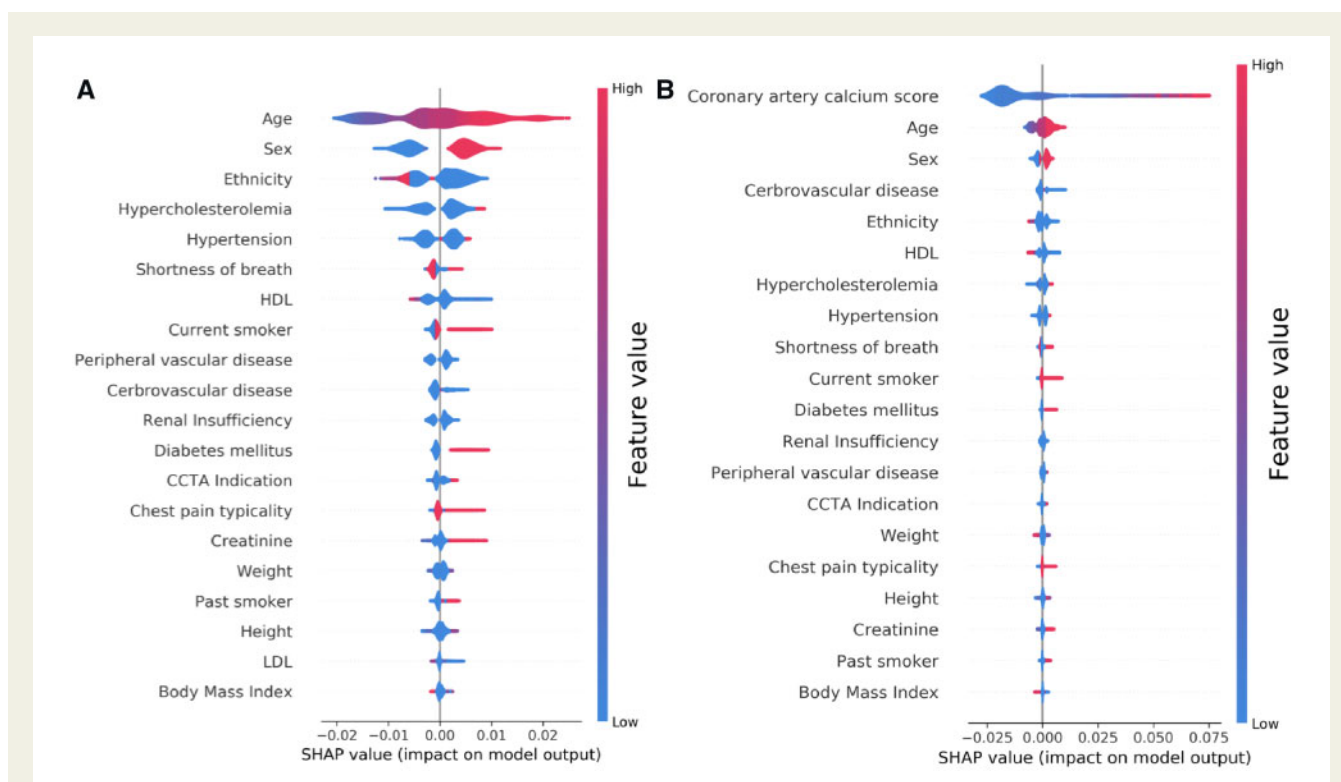


Figure 2 Feature importance plot for the (A) machine learning model and (B) machine learning model with coronary artery calcium score. The top 20 clinical variables are shown in this figure. The blue and red points in each row represent participants having low to high values of the specific variable, while the x-axis gives the SHAP value which gives the impact on the model [i.e. does it tend to drive the predictions towards event (positive value of SHAP) or non-event (negative value of SHAP)]. SHAP, Shapley Additive Explanation values.

The performance of the ML model was subsequently evaluated in select subgroups stratified by age, gender, presence of diabetes mellitus and/or chest pain typicality. The ML model showed improved discrimination for the detection of obstructive CAD in younger individuals (less than 65 years of age) with atypical chest pain (Supplementary material online, Figure S3). Interestingly, there appear to be gender-specific differences in the performance of the ML model since the ML model with CACS is better in males with atypical chest pain than in females with atypical chest pain.

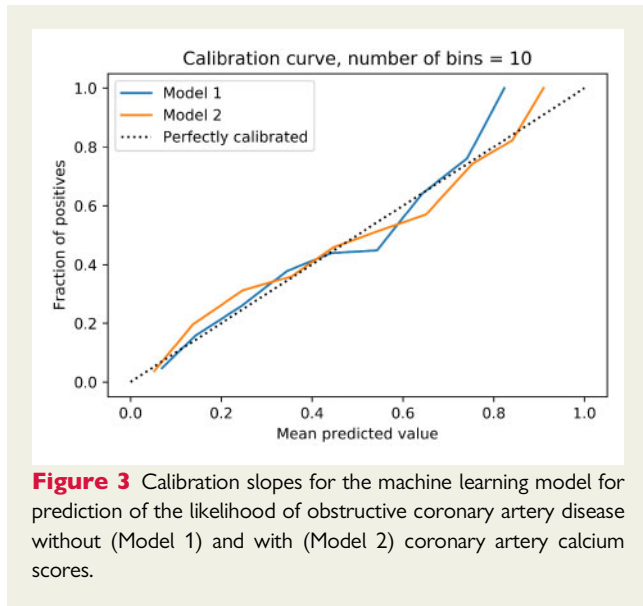


Figure 3 Calibration slopes for the machine learning model for prediction of the likelihood of obstructive coronary artery disease without (Model 1) and with (Model 2) coronary artery calcium scores.

Predictive features

As shown in Figure 2, age, ethnicity, and sex were the most predictive features in the ML model, followed by prior history of hypertension and hypercholesterolaemia. Interestingly, with the addition of CACS into the model, the most predictive features (after the CACS itself) were age and sex followed by history of cerebrovascular disease and the presence of shortness of breath as the presenting symptom. In both models, the presence of shortness of breath was more predictive of the presence of obstructive CAD than the presence of chest pain and typicality of symptoms. Coronary artery calcium score had the highest predictive value as low CACS values were likely to be associated with absence of obstructive CAD (blue colour), while very high values (red colour) were significantly associated with obstructive CAD (Figure 2B).

Calibration and net reclassification

Model calibration was performed in order to assess the certainty of a given new observation belonging to each of the already established classes (prediction of the presence or absence of CAD on CCTA). The calibration slope was 0.856 for the ML model and 0.992 after the addition of CACS, indicating minimal difference between the predicted and observed probability of obstructive CAD and hence good model fit (Figure 3). On the other hand, the Brier score for Model 1 (ML) was 0.205 before and 0.127 after calibration. Similarly to Model 1, Model 2 (ML with CACS) had a Brier score of 0.224 before and a score of 0.099 after calibration. Additionally, continuous NRI was performed comparing both ML models (Models 1 and 2) to the conventional comparator risk scores (Models 3, 4, and 5). Net reclassification index was 0.585 (95% CI 0.495–0.671) when Model 1 (ML)

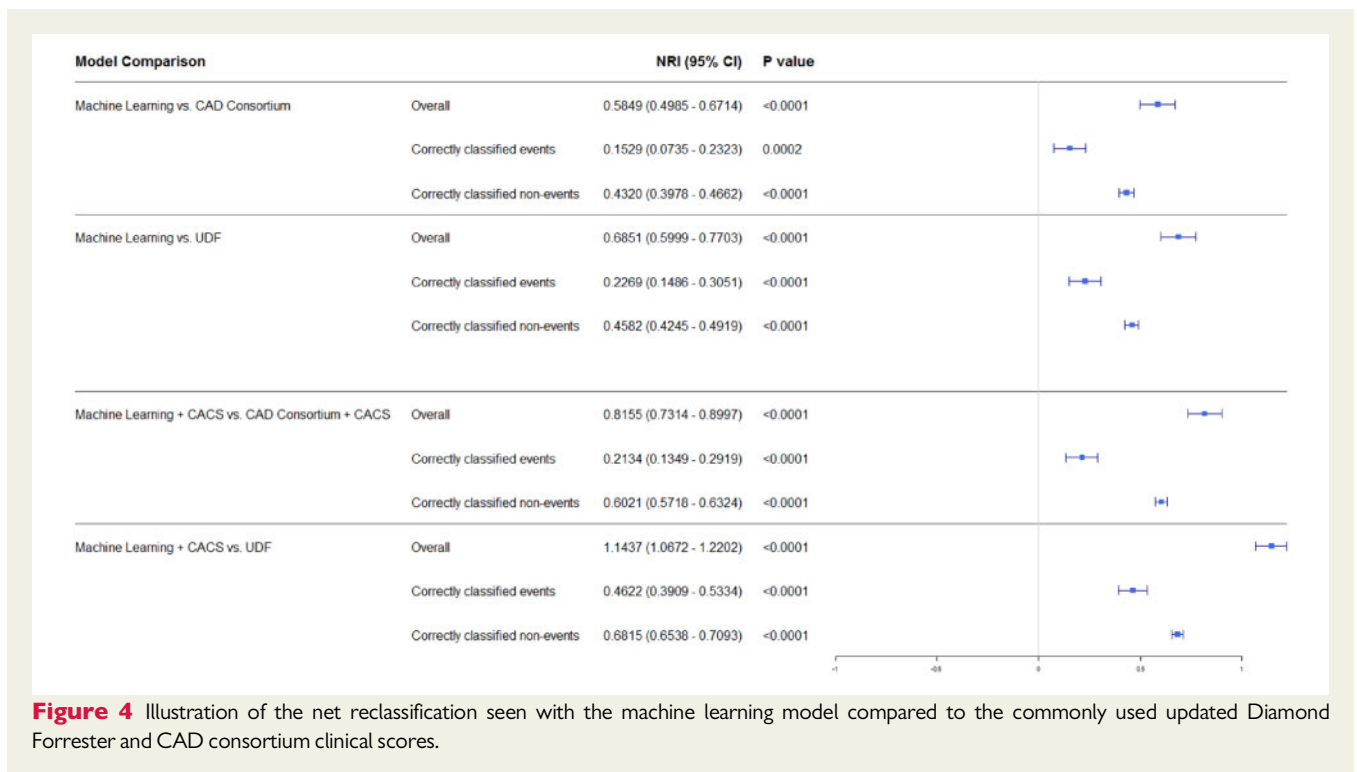
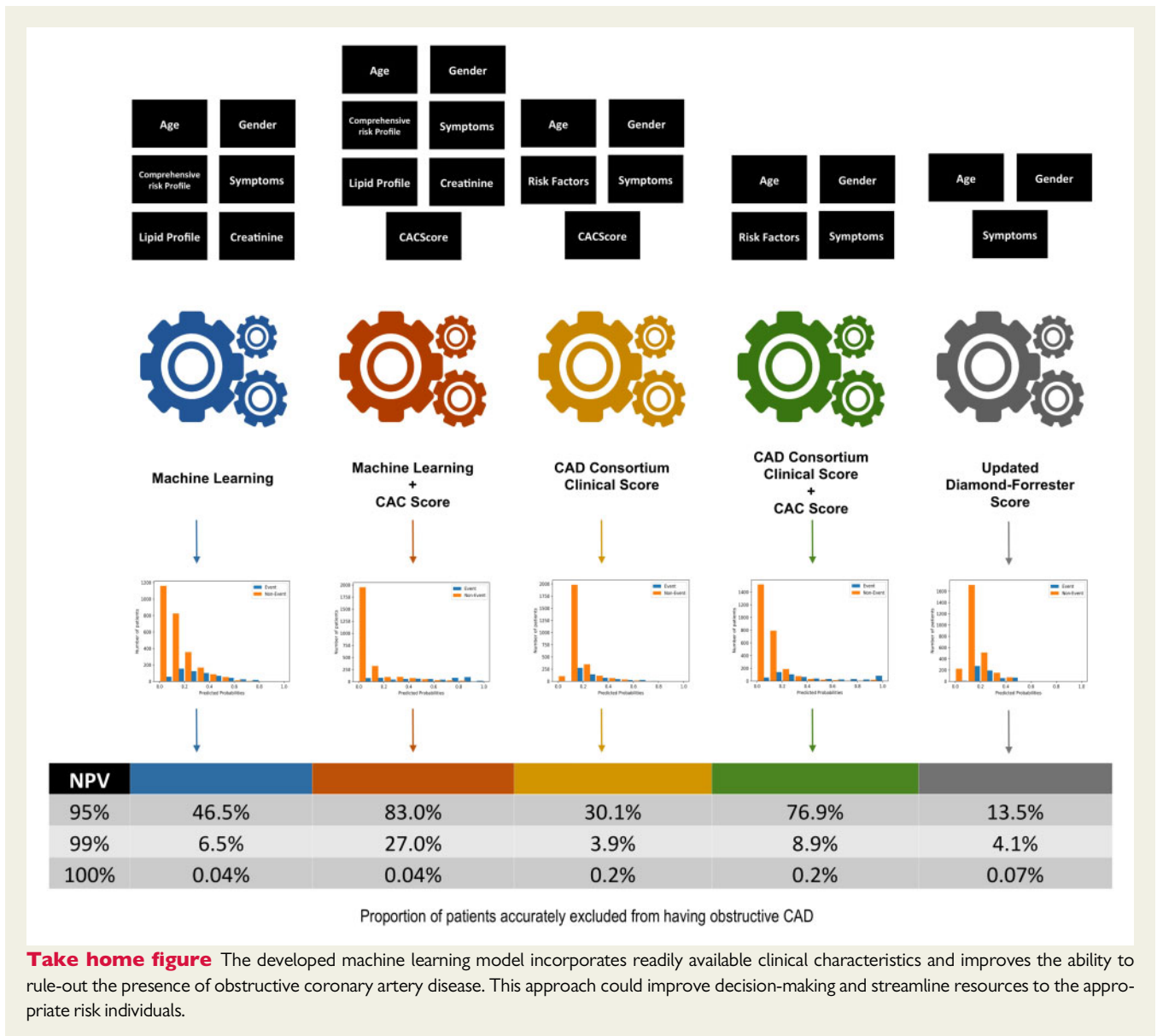


Figure 4 Illustration of the net reclassification seen with the machine learning model compared to the commonly used updated Diamond Forrester and CAD consortium clinical scores.



was compared with Model 3 (CAD consortium) and an NRI of 0.685 (95% CI 0.600–0.770) when compared with Model 5 (UDF). Similarly, Model 2 performed better than Model 4 (CAD consortium + CAC score: NRI 0.816; 95% CI 0.713–0.900) and Model 5 (UDF: NRI of 1.144; 95% CI 1.067–1.220) (Figure 4). All NRIs were driven both by an improvement in correct classification of both events (i.e. correct prediction of obstructive CAD) and non-events (i.e. correct prediction of non-obstructive CAD).

Discussion

Coronary artery disease is a commonly encountered disease entity associated with significant morbidity, mortality, and healthcare expenditure. A conventional routine in clinical practice over the years has been to employ validated diagnostic models of the PTP of stable,

albeit obstructive, CAD in order to direct downstream testing. A majority of existent models have modest performance (with remarkable overestimation of risk in certain subgroups such as women) while very few studies have data regarding the effect of PTP-based models on clinical decision-making regarding further testing or patient outcomes. Hence, there is a need for clinically based models that can predict the PTP of stable CAD and as a result function as gatekeepers to identify low-risk individuals who are unlikely to have obstructive CAD and unlikely to need further diagnostic testing. In the present investigation, we utilized readily available clinically characteristic in a large multicentre, multiethnic cohort undergoing clinically indicated CCTA for the diagnosis of CAD. We utilized ML as a novel analytic approach that is optimized towards the creation of accurate predictive models and found that the developed ML model predicts the occurrence of obstructive CAD on CCTA, specifically in younger individuals with atypical symptoms. Added to that was the finding of

appropriate calibration, improved reclassification and enhanced discrimination of non-events (*Take home figure*). The implementation of the ML model within a clinical setting could help automate the process of selecting for appropriate candidates for further diagnostic evaluation while circumventing more cumbersome routine clinical steps.

Overuse of diagnostic imaging modalities is frequently encountered. As a result, there has been emphasis on the utilization of risk stratification and PTP assessment prior to initiation of downstream testing. The ESC recommends the use of the CAD consortium clinical score while the ACC/AHA recommends the use of the Duke Clinical Score (DCS) or UDF as part of the clinical assessment of the PTP of suspected stable CAD in order to avoid unwarranted examinations.^{5,6} However, multiple investigations have shown that such risk assessment models have suboptimal performance in certain cohorts.^{23–25} Additionally, several models have been validated in more than one external population, with a trend towards lower discriminative ability over the past few years.²⁶ Indeed, differences in derivation (utilization of various imaging modalities as well as different cut-off values for the definition of obstructive CAD, utilization of imputation methodologies for missing values), model complexity as well as inconsistent external validation often exist, which limit their utilization in routine practice. In an ever-changing environment where populations are longitudinally evolving as a result of changing dietary habits, environmental exposures, primordial and preventative practices, there is a need for comprehensive models that evolve over time. To this end, ML has been increasingly applied across the cardiovascular domain. Machine learning involves algorithms that are specifically geared towards finding associations between data beyond the one-dimensional traditional statistical approaches currently utilized. In addition, ML takes advantage of the increasing availability of computational power as well as storage space to provide instantaneous outputs. Machine learning provides the perfect opportunity to use the increasingly complex data that is available while improving predictions in an era of precision medicine. Further, ML has proven itself to be a vastly more powerful tool for prediction across several cardiovascular applications.^{27–31}

The addition of CACS to prediction models has been previously shown to improve performance; similarly, our findings show that the best prediction was achieved with the addition of CACS to the ML model.^{11,15,32} For instance, the addition of the CACS to extended CAD consortium clinical score was found to significantly increase the C-statistic from 0.79 to 0.88 for the prediction of obstructive CAD on invasive coronary angiography.¹¹ Along the same lines, the addition of the Agatston score to the DCS improved the accuracy of prediction of obstructive CAD compared with the DCS by itself (AUC 0.806 vs. AUC 0.714, respectively, $P < 0.05$) in a cohort of 3939 individuals suspected of having CAD and undergoing CCTA evaluation.³² The fact that the CACS significantly improves the estimate of the probability of obstructive CAD is expected given the graded and positive correlation between increasing CACS and the presence of obstructive CAD.^{15,33} Most contemporary calcium score scans now result in less than a millisievert of radiation exposure. The ability of the ML model, with CACS, to correctly reclassify individuals without obstructive CAD could result in reduced radiation exposure and associated costs. However, that statement needs to be supported by

prospective randomized trials that could focus on the evaluation of the effectiveness and safety of such an approach.

The utility of likelihood analysis in the diagnosis of CAD, either on invasive coronary angiography or CCTA, has been an area of growing interest amongst the cardiovascular imaging community given the low prevalence of obstructive CAD in patients undergoing CCTA.^{13,14,26} The overarching goals of cardiovascular imaging, specifically within the context of suspected CAD, are to identify individuals with high-risk anatomy and/or myocardial ischaemia that would improve prognosis with revascularization therapy as well as to improve the utilization of preventative therapies. Data from the National Cardiovascular Data Registry (NCDR) reveal that the diagnostic yield of invasive coronary angiography is low with 149 739/398 978 (37.6%) having obstructive CAD despite an 83.9% pre-catheterization rate of noninvasive testing.³⁴ In a subsequent investigation using the NCDR cohort, CCTA was superior to other noninvasive imaging modalities for the detection of obstructive CAD and thus functioned as an effective gatekeeper to the performance of cardiac catheterization (69.6% rate of detection of obstructive CAD on CCTA compared with 44.5% on single-photon emission computed tomography and 43.8% on stress echocardiography).³⁵ The ML + CACS approach as a routine in symptomatic patients with an intermediate likelihood of obstructive CAD could lead to improved patient outcomes as well as to guide downstream testing. The dramatic reduction in myocardial infarction and cardiac death noted in the SCOT-Heart trial has been attributed to both institution of preventive therapies as well as the use of early revascularization. Current guidelines suggest that CACS can be used to guide the use of preventive therapies in asymptomatic patients at intermediate risk. In clinical practice, it is also used to guide the use of these therapies in symptomatic patients. For patients in whom the ML + CACS were to fall below the level that indicated the need for further testing, preventive therapy resulting from the combined score might lead to improved outcomes. Similarly, if the result of the ML + CACS prompted the use of stress testing, the CACS could lead to appropriate preventive therapies in high proportion of patients in whom the stress testing is negative.

An important consideration for the use of 'big data' for predictive modelling within clinical practice is the need for existence of standardization as well as quality control measures for acquisition and processing of diagnostic imaging results. In the example of CCTA, there exist numerous image acquisition and processing protocols between various sites and institutions. Such variability can be attributed to both hardware and software, as well as variations in interpretation and reporting of imaging findings. There have been several studies that evaluated the influence of intra- and inter-scanner variability on cardiovascular findings on CCTA. As such, the creation of standardized systems for both image acquisition and assessment of findings is of paramount importance especially for the creation of useful and reproducible imaging-based prediction scores. Furthermore, the integration of such risk assessment tools into electronic health record systems would permit continuous optimization of scores at a system level and thereafter help guide with clinical decision-making. Incorporation of a multitude of clinical systems, such as qualitative and quantitative imaging findings, 'omic' data such as genomics and proteomics, will advance abilities further, as will data capture from novel technology, such as wearable gadgets.

There are several limitations of the present investigation that are noteworthy to mention. Firstly, the CONFIRM cohort includes participants referred for CCTA for the evaluation of suspected CAD. As a result of the limitations associated with such a design, there could be a significant referral bias that results in a predictive model that does not apply to a community or general cohort of individuals. However, the goal of utilizing ML was to introduce the concept that such an approach is specifically tailored towards big data, wherein model parameters can be automatically updated and recalibrated as more data becomes available. Secondly, external validation in an independent cohort was not done in the present investigation but is planned for a subsequent analysis on well-validated cohorts of stable chest pain such as PROMISE and SCOT-HEART. Thirdly, UDF and CAD clinical consortium scores were available as comparator scores, while other commonly utilized stratification scores, such as the DCS, were not available for comparison. Nevertheless, the present study included the largest cohort, to date, used for the development of a predictive score for the presence of obstructive CAD on CCTA from a multicentre and multiethnic cohort. Additionally, all included participants had the outcome of interest (i.e. determination of CAD severity on CCTA) without the need to apply imputation techniques. We acknowledge that in the presence of severe calcification (i.e. a high CACS), CCTA overestimates % stenosis, hence our study endpoint (>50% stenosis by CCTA) does not reflect the effective >50% stenosis by coronary angiography. Given the increasing overestimation of degree of % stenosis by CCTA along with increasing CACS, it is not surprising that Model 2 (ML + CACS) outperformed Model 1 (ML). We recognize that using >50% stenosis by coronary angiography as endpoint, different results may be found.

In conclusion, we developed a ML model based on baseline demographic and clinical characteristics for the prediction of obstructive CAD on CCTA that is highly accurate and results in improved net reclassification as a result of correct reclassification of both obstructive and non-obstructive CAD on CCTA. Additionally, the incorporation of CACS further improves risk stratification, such that an ML score that incorporates the CACS and clinical variables may be optimal for initial assessment of younger individuals with atypical symptoms. The utilization of such models may improve decisions in low to intermediate risk patients regarding the need for further testing such as CCTA, as well as for the need for preventive therapies.

Supplementary material

Supplementary material is available at *European Heart Journal* online.

Funding

The research reported in this publication was funded, in part, by the National Institute of Health (Bethesda, MD, USA) under award number R01 HL115150. This research was also supported, in part, by the Dalio Institute of Cardiovascular Imaging (New York, NY, USA) and the Michael Wolk Foundation (New York, NY, USA).

Conflict of interest: All authors have completed the ICMJE uniform disclosure form and declare the following: J.K.M. receives funding from the Dalio Foundation, National Institutes of Health, and GE Healthcare, serves on the scientific advisory board of Arineta and GE Healthcare and has an equity interest in Cleerly. Matthew Budoff receives grant support

from the National Institutes of Health and General Electric. D.S.B. receives a research grant from Heartflow and software royalties from Cedars-Sinai Medical Center. Gianluca Pontone has a research grant and/or honorarium as speaker from GE, Bracco, Bayer, Medtronic, and Heartflow. All other authors have reported that they have no relationships relevant to the contents of this paper to disclose.

References

- Budoff MJ, Dowe D, Jollis JG, Gitter M, Sutherland J, Halamert E, Scherer M, Bellinger R, Martin A, Benton R, Delago A, Min JK. Diagnostic performance of 64-multidetector row coronary computed tomographic angiography for evaluation of coronary artery stenosis in individuals without known coronary artery disease: results from the prospective multicenter ACCURACY (Assessment by Coronary Computed Tomographic Angiography of Individuals Undergoing Invasive Coronary Angiography) trial. *J Am Coll Cardiol* 2008;**52**:1724–1732.
- Miller JM, Rochitte CE, Dewey M, Arbab-Zadeh A, Niinuma H, Gottlieb I, Paul N, Clouse ME, Shapiro EP, Hoe J, Lardo AC, Bush DE, de Roos A, Cox C, Brinker J, Lima JA. Diagnostic performance of coronary angiography by 64-row CT. *N Engl J Med* 2008;**359**:2324–2336.
- Cho I, Chang HJ, Ó Hartaigh B, Shin S, Sung JM, Lin FY, Achenbach S, Heo R, Berman DS, Budoff MJ, Callister TQ, Al-Mallah MH, Cademartiri F, Chinnaiyan K, Chow BJ, Dunning AM, DeLago A, Villines TC, Hadamitzky M, Hausleiter J, Leipsic J, Shaw LJ, Kaufmann PA, Cury RC, Feuchtnr G, Kim YJ, Maffei E, Raff G, Pontone G, Andreini D, Min JK. Incremental prognostic utility of coronary CT angiography for asymptomatic patients based upon extent and severity of coronary artery calcium: results from the COronary CT Angiography EvaluatiON For Clinical Outcomes InteRnational Multicenter (CONFIRM) study. *Eur Heart J* 2015;**36**:501–508.
- Lin FY, Shaw LJ, Dunning AM, Labounty TM, Choi JH, Weinsaft JW, Koduru S, Gomez MJ, Delago AJ, Callister TQ, Berman DS, Min JK. Mortality risk in symptomatic patients with nonobstructive coronary artery disease: a prospective 2-center study of 2,583 patients undergoing 64-detector row coronary computed tomographic angiography. *J Am Coll Cardiol* 2011;**58**:510–519.
- Montalescot G, Sechtem U, Achenbach S, Andreotti F, Arden C, Budaj A, Bugiardini R, Crea F, Cuisset T, Di Mario C, Ferreira JR, Gersh BJ, Gitt AK, Hulot JS, Marx N, Opie LH, Pfisterer M, Prescott E, Ruschitzka F, Sabatè M, Senior R, Taggart DP, van der Wall EE, Vrints CJ, ESC Committee for Practice Guidelines, Zamorano JL, Achenbach S, Baumgartner H, Bax JJ, Bueno H, Dean V, Deaton C, Erol C, Fagard R, Ferrari R, Hasdai D, Hoes AV, Kirchhof P, Knuuti J, Kolh P, Lancellotti P, Linhart A, Nihoyannopoulos P, Piepoli MF, Ponikowski P, Sirnes PA, Tamargo JL, Tendera M, Torbicki A, Wijns W, Windecker S, Document Reviewers, Knuuti J, Valgimigli M, Bueno H, Claeys MJ, Donner-Banzhoff N, Erol C, Frank H, Funck-Brentano C, Gaemperli O, Gonzalez-Juanatey JR, Hamilos M, Hasdai D, Husted S, James SK, Kervinen K, Kolh P, Kristensen SD, Lancellotti P, Maggioni AP, Piepoli MF, Pries AR, Romeo F, Rydén L, Simoons ML, Sirnes PA, Steg PG, Timmis A, Wijns W, Windecker S, Yildirir A, Zamorano JL. 2013 ESC guidelines on the management of stable coronary artery disease: the Task Force on the management of stable coronary artery disease of the European Society of Cardiology. *Eur Heart J* 2013;**34**:2949–3003.
- Fihn SD, Gardin JM, Abrams J, Berra K, Blankenship JC, Dallas AP, Douglas PS, Foody JM, Gerber TC, Hinderliter AL, King SB 3rd, Kligfield PD, Krumholz HM, Kwong RY, Lim MJ, Linderbaum JA, Mack MJ, Munger MA, Prager RL, Sabik JF, Shaw LJ, Sikkema JD, Smith CR Jr, Smith SC Jr, Spertus JA, Williams SV. 2012 ACCF/AHA/ACP/AATS/PCNA/SCAI/STS Guideline for the diagnosis and management of patients with stable ischemic heart disease: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines, and the American College of Physicians, American Association for Thoracic Surgery, Preventive Cardiovascular Nurses Association, Society for Cardiovascular Angiography and Interventions, and Society of Thoracic Surgeons. *J Am Coll Cardiol* 2012;**60**:e44–e164.
- Ferencik M, Liu T, Mayrhofer T, Puchner SB, Lu MT, Maurovich-Horvat P, Pope JH, Truong QA, Udelson JE, Peacock WF, White CS, Woodard PK, Fleg JL, Nagurny JT, Januzzi JL, Hoffmann U. hs-Troponin I followed by CT angiography improves acute coronary syndrome risk stratification accuracy and work-up in acute chest pain patients: results from ROMICAT II trial. *JACC Cardiovasc Imaging* 2015;**8**:1272–1281.
- Meinel FG, Schoepf UJ, Townsend JC, Flowers BA, Geyer LL, Ebersberger U, Krzinski AW, Kunz WG, Thierfelder KM, Baker DW, Khan AM, Fernandes VL, O'Brien TX. Diagnostic yield and accuracy of coronary CT angiography after abnormal nuclear myocardial perfusion imaging. *Sci Rep* 2018;**8**:9228.
- Chinnaiyan KM, Peyser P, Goraya T, Ananthasubramaniam K, Gallagher M, Depetris A, Boura JA, Kazerooni E, Poopat C, Al-Mallah M, Saba S, Patel S, Girard S, Song T, Share D, Raff G. Impact of a continuous quality improvement initiative on appropriate use of coronary computed tomography angiography.

- Results from a multicenter, statewide registry, the Advanced Cardiovascular Imaging Consortium. *J Am Coll Cardiol* 2012;**60**:1185–1191.
10. Genders TS, Steyerberg EW, Alkadhi H, Leschka S, Desbiolles L, Nieman K, Galema TW, Meijboom WB, Mollet NR, de Feyter PJ, Cademartiri F, Maffei E, Dewey M, Zimmermann E, Laule M, Pugliese F, Barbagallo R, Sinitsyn V, Bogaert J, Goetschalckx K, Schoepf UJ, Rowe GW, Schuijf JD, Bax JJ, de Graaf FR, Knuuti J, Kajander S, van Mieghem CA, Meijs MF, Cramer MJ, Gopalan D, Feuchtner G, Friedrich G, Krestin GP, Hunink MG. A clinical prediction rule for the diagnosis of coronary artery disease: validation, updating, and extension. *Eur Heart J* 2011;**32**:1316–1330.
 11. 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.
 12. Bittencourt MS, Hulten E, Polonsky TS, Hoffman U, Nasir K, Abbara S, Di Carli M, Blankstein R. European Society of Cardiology-recommended coronary artery disease consortium pretest probability scores more accurately predict obstructive coronary disease and cardiovascular events than the Diamond and Forrester Score: the Partners Registry. *Circulation* 2016;**134**:201–211.
 13. Douglas PS, Hoffmann U, Patel MR, Mark DB, Al-Khalidi HR, Cavanaugh B, Cole J, Dolor RJ, Fordyce CB, Huang M, Khan MA, Kosinski AS, Krucoff MW, Malhotra V, Picard MH, Udelson JE, Velazquez EJ, Yow E, Cooper LS, Lee KL. Outcomes of anatomical versus functional testing for coronary artery disease. *N Engl J Med* 2015;**372**:1291–1300.
 14. SCOT-HEART investigators. CT coronary angiography in patients with suspected angina due to coronary heart disease (SCOT-HEART): an open-label, parallel-group, multicentre trial. *Lancet* 2015;**385**:2383–2391.
 15. Budoff MJ, Diamond GA, Raggi P, Arad Y, Guerci AD, Callister TQ, Berman D. Continuous probabilistic prediction of angiographically significant coronary artery disease using electron beam tomography. *Circulation* 2002;**105**:1791–1796.
 16. Min JK, Dunning A, Lin FY, Achenbach S, Al-Mallah MH, Berman DS, Budoff MJ, Cademartiri F, Callister TQ, Chang HJ, Cheng V, Chinnaiyan KM, Chow B, Delago A, Hadamitzky M, Hausleiter J, Karlsberg RP, Kaufmann P, Maffei E, Nasir K, Pencina MJ, Raff GL, Shaw LJ, Villines TC. Rationale and design of the CONFIRM (COronary CT Angiography Evaluation For Clinical Outcomes: an International Multicenter) registry. *J Cardiovasc Comput Tomogr* 2011;**5**:84–92.
 17. Abbara S, Blanke P, Maroules CD, Cheezum M, Choi AD, Han BK, Marwan M, Naoum C, Norgaard BL, Rubinshtein R, Schoenhagen P, Villines T, Leipsic J. SCCT guidelines for the performance and acquisition of coronary computed tomographic angiography: a report of the Society of Cardiovascular Computed Tomography Guidelines Committee: endorsed by the North American Society for Cardiovascular Imaging (NASCI). *J Cardiovasc Comput Tomogr* 2016;**10**:435–449.
 18. Wu FZ, Wu MT. 2014 SCCT guidelines for the interpretation and reporting of coronary CT angiography: a report of the Society of Cardiovascular Computed Tomography Guidelines Committee. *J Cardiovasc Comput Tomogr* 2015;**9**:e3.
 19. Hecht HS, Cronin P, Blaha MJ, Budoff MJ, Kazerooni EA, Narula J, Yankelevitz D, Abbara S. 2016 SCCT/STR guidelines for coronary artery calcium scoring of noncontrast noncardiac chest CT scans: a report of the Society of Cardiovascular Computed Tomography and Society of Thoracic Radiology. *J Cardiovasc Comput Tomogr* 2017;**11**:74–84.
 20. Chen T, Guestrin, C. Xgboost: A scalable tree boosting system. In: *Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2016. ACM.
 21. Al'Aref SJ, Anchouche K, Singh G, Slomka PJ, Kolli KK, Kumar A, Pandey M, Maliakal G, van Rosendaal AR, Beecy AN, Berman DS, Leipsic J, Nieman K, Andreini D, Pontone G, Schoepf UJ, Shaw LJ, Chang HJ, Narula J, Bax JJ, Guan Y, Min JK. Clinical applications of machine learning in cardiovascular disease and its relevance to cardiac imaging. *Eur Heart J* 2019;**40**:1975–1986.
 22. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. In: *Advances in Neural Information Processing Systems 30 NIPS* 2017. <https://papers.nips.cc/paper/7062-a-unified-approach-to-interpreting-model-predictions.pdf>.
 23. Isma'eel HA, Serhan M, Sakr GE, Lamaa N, Garabedian T, Elhaji I, Skouri H, Abchee A. Diamond-Forrester and Morise risk models perform poorly in predicting obstructive coronary disease in Middle Eastern Cohort. *Int J Cardiol* 2016;**203**:803–805.
 24. Baskaran L, Danad I, Gransar H, Hartaigh B Ó, Schulman-Marcus J, Lin FY, Peña JM, Hunter A, Newby DE, Adamson PD, Min JK. A comparison of the updated Diamond-Forrester, CAD Consortium, and CONFIRM history-based risk scores for predicting obstructive coronary artery disease in patients with stable chest pain: the SCOT-HEART Coronary CTA Cohort. *JACC Cardiovasc Imaging* 2019;**12**:1392–1400.
 25. Zhou J, Liu Y, Huang L, Tan Y, Li X, Zhang H, Ma Y, Zhang Y. Validation and comparison of four models to calculate pretest probability of obstructive coronary artery disease in a Chinese population: a coronary computed tomographic angiography study. *J Cardiovasc Comput Tomogr* 2017;**11**:317–323.
 26. He T, Liu X, Xu N, Li Y, Wu Q, Liu M, Yuan H. Diagnostic models of the pretest probability of stable coronary artery disease: a systematic review. *Clinics (Sao Paulo)* 2017;**72**:188–196.
 27. Motwani M, Dey D, Berman DS, Germano G, Achenbach S, Al-Mallah MH, Andreini D, Budoff MJ, Cademartiri F, Callister TQ, Chang HJ, Chinnaiyan K, Chow BJ, Cury RC, Delago A, Gomez M, Gransar H, Hadamitzky M, Hausleiter J, Hindoyan N, Feuchtner G, Kaufmann PA, Kim YJ, Leipsic J, Lin FY, Maffei E, Marques H, Pontone G, Raff G, Rubinshtein R, Shaw LJ, Stehli J, Villines TC, Dunning A, Min JK, Slomka PJ. Machine learning for prediction of all-cause mortality in patients with suspected coronary artery disease: a 5-year multicentre prospective registry analysis. *Eur Heart J* 2017;**38**:500–507.
 28. van Rosendaal AR, Maliakal G, Kolli KK, Beecy A, Al'Aref SJ, Dwivedi A, Singh G, Pandey M, Kumar A, Ma X, Achenbach S, Al-Mallah MH, Andreini D, Bax JJ, Berman DS, Budoff MJ, Cademartiri F, Callister TQ, Chang H-J, Chinnaiyan K, Chow BJW, Cury RC, DeLago A, Feuchtner G, Hadamitzky M, Hausleiter J, Kaufmann PA, Kim Y-J, Leipsic JA, Maffei E, Marques H, Pontone G, Raff GL, Rubinshtein R, Shaw LJ, Villines TC, Gransar H, Lu Y, Jones EC, Peña JM, Lin FY, Min JK. Maximization of the usage of coronary CTA derived plaque information using a machine learning based algorithm to improve risk stratification: insights from the CONFIRM registry. *J Cardiovasc Comput Tomogr* 2018;**12**:204–209.
 29. Coenen E, Kim YH, Kruk M, Tesche C, De Geer J, Kurata A, Lubbers ML, Daemen J, Itu L, Rapaka S, Sharma P, Schwemmer C, Persson A, Schoepf UJ, Kepka C, Hyun Yang D, Nieman K. Diagnostic accuracy of a machine-learning approach to coronary computed tomographic angiography-based fractional flow reserve: result from the MACHINE consortium. *Circ Cardiovasc Imaging* 2018;**11**:e007217.
 30. Kalscheur MM, Kipp RT, Tattersall MC, Mei C, Buhr KA, DeMets DL, Field ME, Eckhardt LL, Page CD. Machine learning algorithm predicts cardiac resynchronization therapy outcomes: lessons from the COMPANION trial. *Circ Arrhythm Electrophysiol* 2018;**11**:e005499.
 31. Mortazavi BJ, Downing NS, Bucholz EM, Dharmarajan K, Manhapra A, Li SX, Negabhan SN, Krumholz HM. Analysis of machine learning techniques for heart failure readmissions. *Circ Cardiovasc Qual Outcomes* 2016;**9**:629–640.
 32. Takamura K, Kondo T, Fujimoto S, Hiki M, Matsumori R, Kawaguchi Y, Amanuma M, Takase S, Daida H. Incremental predictive value for obstructive coronary artery disease by combination of Duke Clinical Score and Agatston score. *Eur Heart J Cardiovasc Imaging* 2016;**17**:550–556.
 33. Haberl R, Becker A, Leber A, Knez A, Becker C, Lang C, Brünig 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.
 34. 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.
 35. Patel MR, Dai D, Hernandez AF, Douglas PS, Messenger J, Garratt KN, Maddox TM, Peterson ED, Roe MT. Prevalence and predictors of nonobstructive coronary artery disease identified with coronary angiography in contemporary clinical practice. *Am Heart J* 2014;**167**:846–852.e2.