

RESEARCH

Open Access



# Predictive value of CAC score combined with clinical features for obstructive coronary heart disease on coronary computed tomography angiography: a machine learning method

Yongkui Ren<sup>1,5</sup>, Yulin Li<sup>1,2,3</sup>, Weili Pan<sup>5</sup>, Da Yin<sup>4\*\*†</sup> and Jie Du<sup>1,2,3\*\*†</sup>

## Abstract

**Objective:** We investigated the predictive value of clinical factors combined with coronary artery calcium (CAC) score based on a machine learning method for obstructive coronary heart disease (CAD) on coronary computed tomography angiography (CCTA) in individuals with atypical chest pain.

**Methods:** The study included data from 1,906 individuals undergoing CCTA and CAC scanning because of atypical chest pain and without evidence for the previous CAD. A total of 63 variables including traditional cardiovascular risk factors, CAC score, laboratory results, and imaging parameters were used to build the Random forests (RF) model. Among all the participants, 70% were randomly selected to train the models on which fivefold cross-validation was done and the remaining 30% were regarded as a validation set. The prediction performance of the RF model was compared with two traditional logistic regression (LR) models.

**Results:** The incidence of obstructive CAD was 16.4%. The area under the receiver operator characteristic (ROC) for obstructive CAD of the RF model was 0.841 (95% CI 0.820–0.860), the CACS model was 0.746 (95% CI 0.722–0.769), and the clinical model was 0.810 (95% CI 0.788–0.831). The RF model was significantly superior to the other two models ( $p < 0.05$ ). Furthermore, the calibration curve and Hosmer–Lemeshow test showed that the RF model had good classification performance ( $p = 0.556$ ). CAC score, age, glucose, homocysteine, and neutrophil were the top five important variables in the RF model.

**Conclusion:** RF model was superior to the traditional models in the prediction of obstructive CAD. In clinical practice, the RF model may improve risk stratification and optimize individual management.

**Keywords:** Machine learning, Random forest, Coronary artery calcification score, Obstructive coronary artery disease

<sup>†</sup>Da Yin and Jie Du contributed equally to this study

\*Correspondence: dlyinda@hotmail.com; jiedu@ccmu.edu.cn

<sup>1</sup> Beijing Anzhen Hospital, Capital Medical University, No. 2 Anzhen Road, Chaoyang District, Beijing, China

<sup>4</sup> Department of Cardiology, Shenzhen People's Hospital, 2nd Clinical Medical College of JINAN University, 1st Affiliated Hospital of Southern University of Science and Technology, ShenZhen, China

Full list of author information is available at the end of the article

## Introduction

Individuals with atypical chest pain are common in clinical practice. It is a big challenge for cardiologists to more accurately distinguish patients with obstructive coronary artery disease (CAD) from atypical patients. Coronary computed tomography angiography (CCTA) can non-invasively evaluate the severity of CAD [1, 2]. However,



intensive X-ray and contrast usage and high-cost limit wide applications of CCTA in routine screenings. Similarly, several guidelines have recommended using the Diamond and Forrester model (DF) or the Duke clinical score (DCS) to estimate the pretest probability of CAD in patients with chest pain [3, 4]. Nevertheless, recent studies have proved that DF and DCS score tend to overestimate the probability of obstructive CAD [5, 6]. Given the above, it is necessary to seek a new prediction model for obstructive CAD.

Coronary artery calcium (CAC), known as a biomarker of subclinical atherosclerosis, is tightly related to the occurrence of future cardiovascular events and all-cause mortality [7–10]. The addition of CAC scores to prediction models has been reported to enhance performance for obstructive CAD [11, 12].

Machine learning (ML) has emerged as a neoteric category of artificial intelligence and is widely applied to healthcare data analysis [13–15]. Random forest (RF), as a classic ML algorithm, is good at using out-of-bag estimates and an internal bootstrap to reduce and select predictive features and avoid over-fitting[16]. In this study, we sought to develop the RF model to predict patients with obstructive CAD and compare the prediction performance between the RF model and two traditional logistic regression (LR) models.

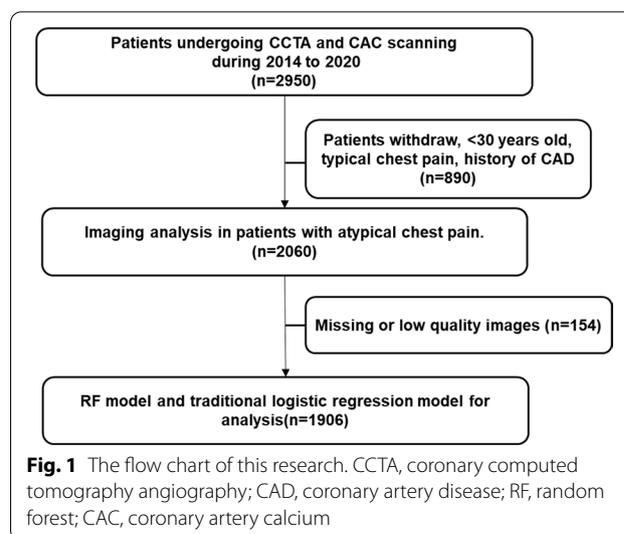
## Methods

### Study population

We retrospectively screened patients admitted to the First Affiliated Hospital of Dalian Medical University from January 2014 to December 2020 and the detailed flow chart was shown in Fig. 1. The data includes demographics, CAC scores, and clinical and imaging parameters. Inclusion criteria were patients with  $\geq 30$  years old, atypical chest pain, no history of CAD, and underwent CAC and CCTA scanning. Individuals with missing data of scan identifiers, no-dedicated CAC score, no-CAC scanning, uncertain date of birth, and uncertain data of scan were excluded. A total of 1,906 patients were enrolled in the study. This study was approved by the Ethics Committee Board of the First Affiliated Hospital of Dalian Medical University.

### Coronary computed tomography angiography and coronary artery calcium scanning

The scanner (dual-source, Somatom Definition CT, Siemens, Erlangen, Germany) was applied to acquire and process the CCTA images and as well as CAC scores. All the processes strictly followed the guidelines recommended by the Society of Cardiovascular Computed

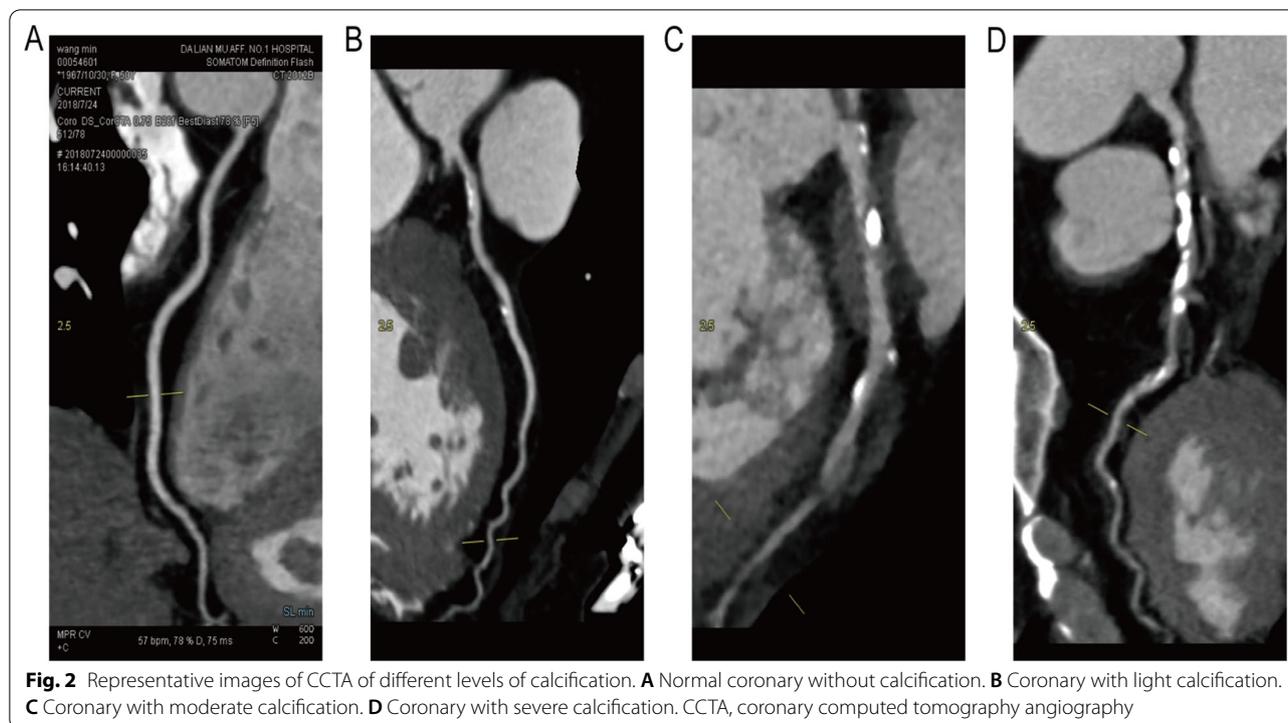


Tomography [17]. Meanwhile, two independent and professional imaging physicians assessed all images, identified the severity of CAD, and determined a CAC score based on the Agatston method [18]. Furthermore, representative CCTA images of different levels of calcification were shown in Fig. 2. The presence of diameter stenosis  $\geq 50\%$  in any of the four major epicardial coronary arteries detected on CCTA was defined as obstructive CAD and the outcome of the present study.

### Building a machine learning model

A total of 63 available demographic and clinical variables of screened individuals were analyzed in this study, including age, sex, cardiovascular risk factors (the history of hypertension and diabetes mellitus, current smoker), baseline plasma lipid levels (total cholesterol, high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C) and triglyceride), electrocardiogram parameters, imaging parameters (ultrasonic cardiogram, carotid ultrasonography) and so on (shown in Additional file 3: Table S1).

RF, based on decision tree mechanisms, combines various decision tree classifiers to provide the final classification and improve classification accuracy. In this study, the RF model included all available variables, and all included individuals were randomly divided into the training set (70%) and the validation sets (30%). The random processes were continuously repeated until all the data equally distributed in both sets. Meanwhile, tuning was considered to avoid overfitting for ML-based models and the optimal hyper-parameter in the training process for ML models was fivefold cross-validation.



### Comparison of predictive performance between the RF model and LR models

To evaluate the predictive performance and clinical value of our proposed RF model, we compared the RF model with two traditional Logistic Regression (LR) models which were constructed as follows: (1) a model (regarded as CACS model) trained with CAC score alone; and (2) the other model (considered as a clinical model) trained with cardiovascular risk factors (age, sex, the history of hypertension and diabetes mellitus, current smoker, total cholesterol, LDL-C, HDL-C, and CAC score). To make paired comparisons, the same folds and cross-validation procedures were performed in the training and evaluation of the two LR models as the RF model.

### Statistical analysis

Continuous variables are reported as mean  $\pm$  SD for normally distributed variables or median (interquartile range) for non-normally distributed data and were compared with one-way ANOVA or non-parametric test. Categorical variables are expressed as a number (percentage) and were compared with chi-square tests or Fisher's exact test.

The RF model was compared with the traditional LR models using the calibration curve and Hosmer–Lemeshow test. The area under the receiver operator characteristics curve (AUC) was applied to assess

the performance of predictive models. To examine the added improvement of the RF model in predicting obstructive CAD, we utilized the continuous net reclassification index (NRI) and integrated discrimination improvement (IDI) analysis. The decision curve analysis (DCA) was used to compare the benefit of the three models at different threshold probabilities. Finally, The Youden index was also provided to summarize performance predictions. All statistical analyses were performed by R software (Version 4.0.3 R Foundation for Statistical Computing, Vienna, Austria). The R packages “caret,” “e1071,” “random-forest,” “dplyr,” “nricens,” “rmda,” “GLM,” “pROC” were used in this study. Statistical significance was defined as a  $p < 0.05$  on a 2-tailed test.

## Results

### Demographic features

A total of 1,906 patients were finally included in the study. The occurrence of obstructive CAD was 16.4% (313 out of 1,906) within the studied cohort (as shown in Table 1; Additional file 4: Table S2). The mean age of the cohort was 57.4 years, 54.6% were men, 48.0% had hypertension, 26.1% had diabetes mellitus, and 28.9% were current smokers. The proportion of CAC score = 0 in this study was 42.4%, and the proportion of severe calcification (CAC score  $\geq 400$ ) was 11.4%.

**Table 1** Baseline characteristics of the study population

	Total (n = 1906)
Age, years	57.4 ± 14.53
Male, n (%)	1041 (54.6)
Hypertention, n (%)	914 (48.0)
Diabetes mellitus, n (%)	498 (26.1)
Current smoker, n (%)	551 (28.9)
TC (mmol/L)	4.79 ± 1.05
TG (mmol/L)	1.40 (1.02–2.03)
HDL-C (mmol/L)	1.10 (0.93–1.31)
LDL-C (mmol/L)	2.63 ± 0.72
Serum creatinine, (μmol/L)	65.0 (54.0–78.0)
LVEF, %	57.81 ± 4.92
OCAD, n (%)	313 (16.4)
CACS, AU	3.35 (0.00–109.63)
CACS (n = 1906)	
0	850 (44.5)
1–99	563 (29.5)
100–399	275 (14.4)
≥ 400	218 (11.4)

Values are presented as mean ± SD, median (25th–75th percentiles) or n (%)

TC, total cholesterol; TG, triglyceride; HDL-C, high density lipoprotein cholesterol; LDL-C, low density lipoprotein cholesterol; LVEF, left ventricular ejection fraction; CACS, coronary artery calcium score; OCAD, obstructive coronary artery disease

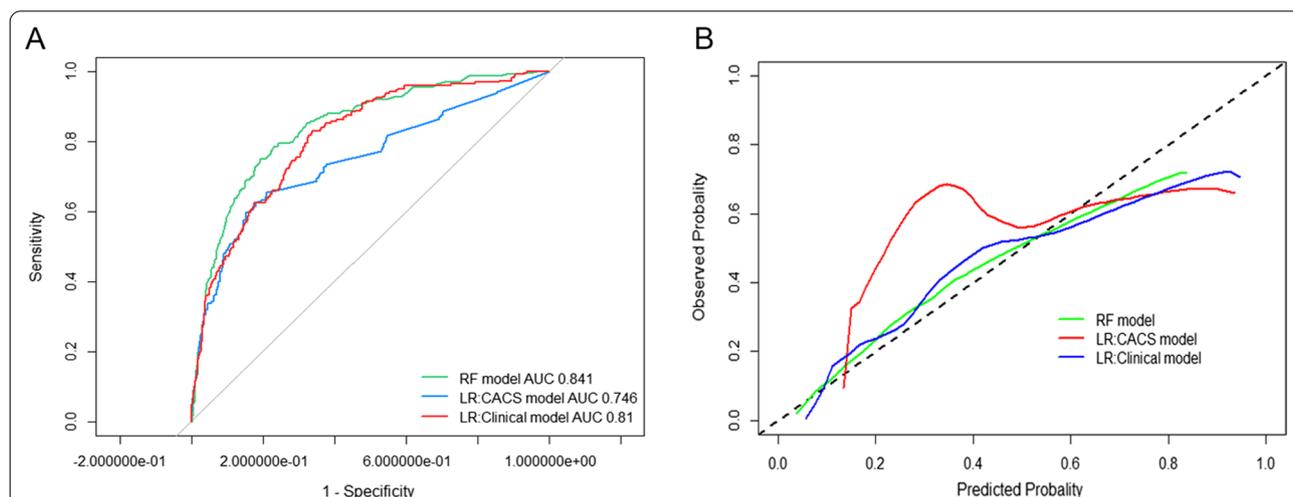
**Comparison of predictive performance for obstructive CAD between the RF model and traditional LR models**

To avoid overfitting of the model, we explored correlation coefficients between variables and showed them as a matrix in Additional file 1: Figure S1. Furthermore,

within the test set, the predictive performances of the three models were compared and detailed in Fig. 3A, which were evaluated based on the area under the receiver operating characteristics curve (area under AUC). Interestingly, the RF model produced the best performance in terms of predicting individuals with obstructive CAD, with an AUC of 0.841 (95% CI 0.820–0.860) compared with the CACS model (AUC 0.746, 95% CI 0.722–0.769), and clinical model (AUC 0.810, 95% CI 0.788–0.831), *p* < 0.05 for all comparisons.

Model calibration was deployed to evaluate the possibility of a given new observation belonging to each of the already established classes (the presence or absence of CAD on CCTA). RF model indicated a minimal difference between the predicted and observed probability of obstructive CAD. Therefore, the RF model achieved a good model fit (as shown in Fig. 3B). Furthermore, the Hosmer–Lemeshow test indicated that the RF model had a high calibration (*p* = 0.556), while the other two models were disappointing (*p* < 0.05). Additionally, continuous NRI was 0.18 (95% CI 0.040–0.327), and IDI was 0.03 (95% CI 0.005–0.058) when the RF model was compared with the clinical model (as shown in Table 2), with *p* < 0.05 for all comparisons.

Table 3 showed the Youden index for each model, with our RF model having the greatest Youden index for obstructive CAD. At the optimal cutoff, we observed 74.9% sensitivity, 80.6% specificity, 43.6% positive predictive value, 94.1% negative predictive value, and 77.8% accuracy. Moreover, the RF model still displayed strong



**Fig. 3** AUC (A) and calibration curve (B) of the different models for the prediction of obstructive coronary artery disease on CCTA. CACS model including CAC score only; the clinical model includes atherosclerotic cardiovascular disease risk factors and CAC score. AUC, the area under the receiver operating characteristics curve; CCTA, coronary computed tomography angiography; LR, logistic regression; RF, random forest; CAC, coronary artery calcium

predictive capabilities in individuals stratified by age and sex (as shown the Additional file 2: Figure S2).

**The relative importance of variables in RF algorithm**

We selected the variables that rank high in both mean decrease Gini and mean decrease accuracy. As shown in Fig. 4, the probability of the prevalence of obstructive CAD increased, with CACS increasing. That is,

CACS had the highest predictive value for the presence of obstructive CAD. Moreover, the most predictive features (after the CAC score itself) were age and fasting glucose levels followed by plasma homocysteine levels and the number of neutrophils. Intriguingly, left atrial dimension, carotid intima-media thickness (IMT) and QT interval might also have positive effects on the incidence of obstructive CAD among imaging parameters.

**Table 2** The comparisons of NRI and IDI between the different models

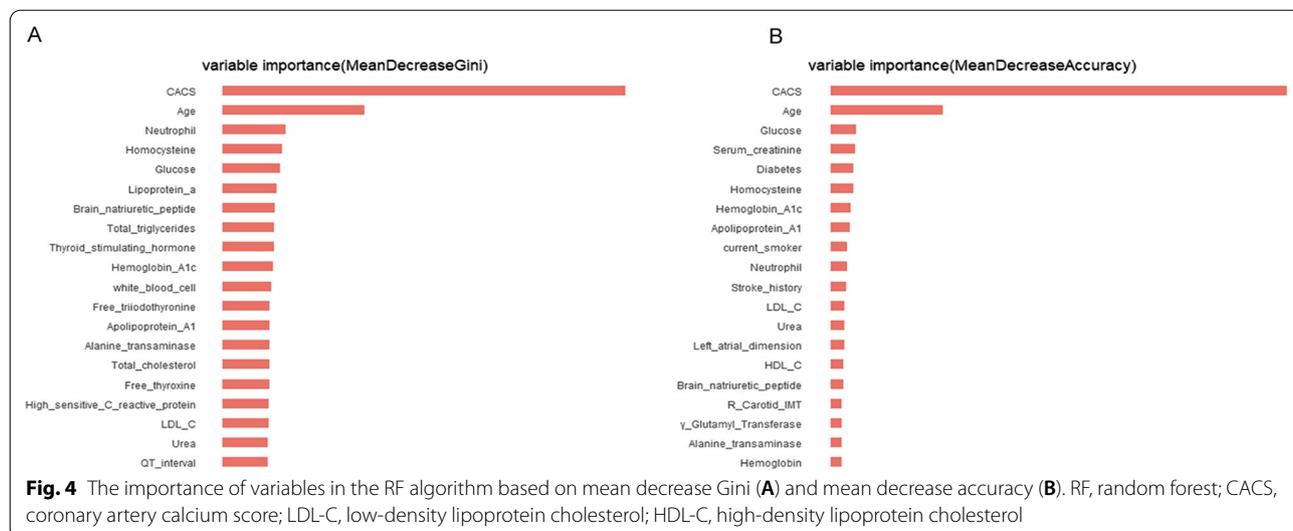
	NRI (95%CI)	P	IDI (95%CI)	P
RF model versus clinical model	0.18 (0.040–0.327)	0.012	0.03 (0.005–0.059)	0.021
Events	0.08 (– 0.053 to 0.221)	0.218		
Non-events	0.09 (0.035–0.158)	0.001		
RF model versus CACS model	0.82 (0.696–0.953)	< 0.0001	0.13 (0.10–0.155)	< 0.0001
Events	0.47 (0.352–0.578)	< 0.0001		
Non-events	0.35 (0.298–0.406)	< 0.0001		
Clinical model versus CACS model	0.64 (0.501–0.767)	< 0.0001	0.09 (0.075–0.117)	< 0.0001
Events	0.28 (0.153–0.393)	< 0.0001		
Non-events	0.36 (0.303–0.412)	< 0.0001		

CACS model including CAC score. Clinical model including ASCVD risk factors and CAC score  
 ASCVD, atherosclerotic cardiovascular disease; LR, logistic regression; RF, random forest; NRI

**Table3** Optimal cutpoints and youden index for each model

	Cutpoint	SE	SP	PPV	NPV	Accuracy	Youden index
RF model	0.229	0.749	0.806	0.436	0.941	0.778	0.555
CACS model	0.136	0.655	0.789	0.384	0.919	0.722	0.444
Clinical model	0.129	0.825	0.644	0.317	0.948	0.734	0.468

SE, sensitivity; SP, specificity; PPV, positive predictive value; NPV, negative predictive value; RF, random forest; CACS model, CAC score only; clinical model, atherosclerotic cardiovascular disease risk factors and CAC score



### The DCA of the three prediction models

The DCA was used to compare the benefit of the RF model, CACS model, and Clinical model. We found that if the threshold probability in the clinical decision was >10%, the patients would benefit more from the RF model than either the CACS model or the Clinical model (as shown in Fig. 5).

### Discussion

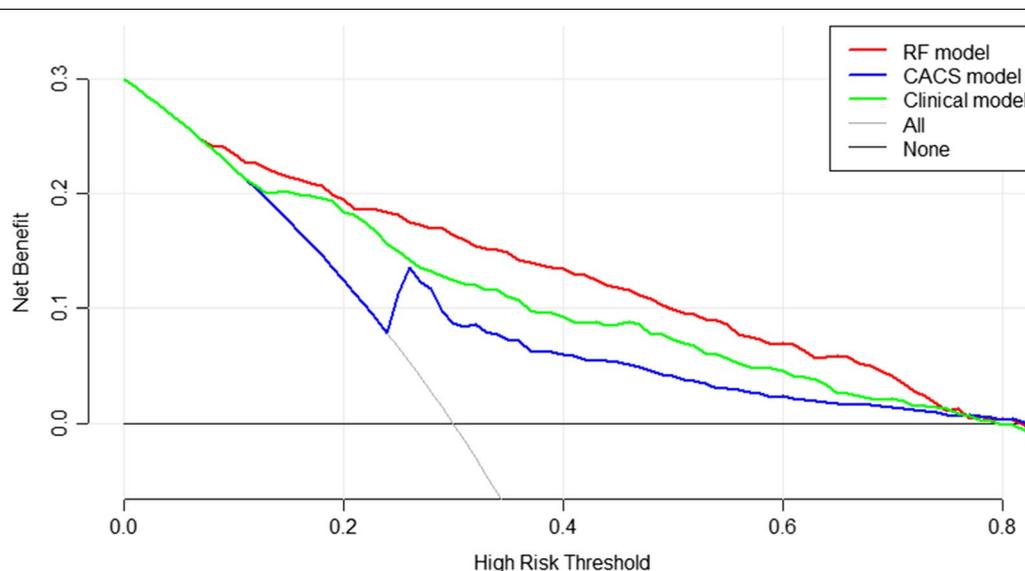
In this study, we have shown that the RF model integrating clinical variables and CAC score can obtain superior prognostic performance than the traditional LR models for obstructive CAD on CCTA. In addition, our comprehensive RF model obtained high concordance between the predicted risk and actual observed risk. CAC score was the most important variable in the RF model, followed by age, fasting glucose levels, plasma homocysteine levels, and the number of neutrophils.

Obstructive CAD is the most common etiology of atypical chest pain, which significantly increase mortality and healthcare expenditure. To noninvasively predict the occurrence of CAD, many models have been developed, such as CCTA, CACS, and cardiac magnetic resonance angiography [3, 4]. Nevertheless, the performance of many existent models is limited in the presence of obstructive CAD [19, 20]. Beyond that, the discriminative ability of some models has become lower in more than one external population in an ever-changing world [6]. Therefore, there is an urgent need for optimal predictive

models for obstructive CAD in individuals with atypical chest pain.

ML, as a scientific algorithm, can make data-driven predictions by learning from the training set and finishing subsequent prediction tasks in an independent set [21]. Compared with other ML algorithms, such as neural network (NNET) and support vector machine (SVM), RF does not need to select features in advance and prevent over-fitting [16, 22]; Compared with the traditional LR models, RF, a classic ML algorithm, can account for non-linear and higher dimensional relationships between a multitude of variables that could potentially lead to an improved explanatory model. Similarly, our research found that although the LR models containing the CAC score have moderate predictive power, however, the calibration curve fitting did not achieve well which has been proved by the Hosmer–Lemeshow. On the contrary, the RF model showed a better predictive performance for obstructive CAD. Additionally, RF models have shown equal or better performance than humans in medical practices such as diagnosis, decision-making, and risk prediction in cardiology [16]. Our findings uphold the RF model based on all available information and CAC scores can more accurately identify high-risk individuals and improve the clinical use of the CAC scanning in risk assessment and guiding management decisions [11, 23–25].

In the order of variable importance, consistent with the previous studies, the CAC score is superior to



**Fig. 5** The decision curve analysis (DCA) of the three prediction models. Theoretically, all patients without obstructive coronary artery disease are represented by a black line, and all patients with obstructive coronary artery disease are represented by a gray line. The x-axis presented the Threshold probability. CACS model including CAC score only; The clinical model including atherosclerotic cardiovascular disease risk factors and CAC score

traditional cardiovascular risk factors, such as age, sex, smoking, the presence of diabetes mellitus and hyperlipidemia, and so on. The CAC score measured by non-contrast cardiac-gated computed tomography (CT) provides an evaluation of the global burden of coronary atherosclerosis. Furthermore, the CAC score can provide a long-term and independent prognosis for the clinical risk of cardiovascular disease (CVD) and CAD events [7–9]. Therefore, accurate coronary calcification detection and assessment can aid in clinical decision-making. Recent research has shown that deep learning techniques, irrespective of picture quality and calcification, can precisely estimate coronary artery calcification from CT angiography images [15, 26]. In future clinical applications, it might have a significant impact.

A good model should take into account not only its diagnostic effectiveness but also its repeatability, non-invasiveness, and simplicity. In our study, other CT variables such as the total number of calcified coronary lesions, plaque density, the presence of thoracic aorta calcification, and so on, which have been revealed to increase the predictive potency of CAC for CVD events were not included in the present prediction model [27, 28]. However, the prediction results of the RF model in our investigation were similar to those of the previously reported Extreme Gradient Boosting (XGBoost) model [11]. Additionally, our preliminary experiments showed that the RF model had better calibration than XGBoost. Considering its effect on the insensitivity of missing values and the advantages of dealing with high-dimensional data make it easier to generalize in clinical practice. Last but not least, current guidelines have recommended that CACS can be used to guide preventive therapies in asymptomatic individuals at intermediate risk for CVD events [29, 30]. Given the above, patients at lower risk in the RF model may not require further testing, such as CCTA or Coronary angiography.

Several limitations of the present study should be paid more attention to. Firstly, the present investigation was lack of external validation in an independent cohort, which was planned for subsequent analysis. Secondly, the presence of severe calcification may lead to overestimates % stenosis on CCTA. Hence, more than 50% stenosis on CCTA may not represent the accuracy > 50% stenosis evaluated by coronary angiography. Thirdly, further study with long follow-up times is very necessary to assess the long-term predictive role of the CAC score. Fourthly, all the screened subjects were all from China, thus, the predictive model may be not suitable for other ethnic groups. Finally, we did not use multiple ML algorithms for this research, but the RF model has shown better predictive ability in previous studies.

## Conclusion

The predictive performance of the RF model integrating clinical variables and CAC score is superior to models combining traditional risks and CAC score for the presence of obstructive CAD in patients with atypical chest pain. It may be unreasonable for individuals at a low risk assessed by the prediction model to receive a further invasive examination.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12872-022-03022-9>.

**Additional file 1. Supplementary Figure 1. A matrix diagram of several variables.** Red means positive correlation, blue means negative correlation. LDL-C, low-density lipoprotein cholesterol; HDL-C, high-density lipoprotein cholesterol.

**Additional file 2. Supplementary Figure 2. Subgroup analysis of different prediction models.** CACS model including CAC score only. Clinical model including atherosclerotic cardiovascular disease risk factors and CAC score. AUC, area under the receiver operating characteristics curve; LR, Logistic Regression; RF, Random Forest; CAC, coronary artery calcium.

**Additional file 3. Supplementary Table 1.** Pre-implant clinical features included in the analysis.

**Additional file 4. Supplementary Table 2.** Comparisons between the test cohort and the training cohort.

## Acknowledgements

The authors thank Yidu Cloud Technology Co., Ltd for their assistance in data searching, extraction, and processing.

## Author contributions

JD and DY made substantial contributions to the conception and design of the study. YR participated in the study design, analyzed the data, and drafted the manuscript. YL participated in the collection and analysis of data. WP revised the manuscript. All authors have approved the final version of the manuscript. All authors read and approved the final manuscript.

## Funding

This study was funded by Liaoning Province Natural Science Foundation (2019-ZD-0635) to Dr. Ren.

## Availability of data and materials

The data are available from the corresponding authors upon reasonable request.

## Declarations

### Ethics approval and consent to participate

All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation and with the Helsinki Declaration of 1975, as revised in 2000. This study was approved by the Ethics Committee Board of the First Affiliated Hospital of Dalian Medical University. All participants gave their written informed consent.

### Consent for publication

Not applicable.

### Competing interests

The authors declare that they have no conflict of interest.

### Author details

<sup>1</sup>Beijing Anzhen Hospital, Capital Medical University, No. 2 Anzhen Road, Chaoyang District, Beijing, China. <sup>2</sup>Key Laboratory of Remodeling-Related

Cardiovascular Diseases, Ministry of Education, Beijing, China. <sup>3</sup>Beijing Institute of Heart, Lung, and Blood Vessel Disease, Beijing, China. <sup>4</sup>Department of Cardiology, Shenzhen People's Hospital, 2nd Clinical Medical College of JINAN University, 1st Affiliated Hospital of Southern University of Science and Technology, Shenzhen, China. <sup>5</sup>Department of Cardiology, 1st Affiliated Hospital of Dalian Medical University, Dalian, China.

Received: 14 April 2022 Accepted: 19 December 2022

Published online: 26 December 2022

## References

- Budoff MJ, et al. 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(21):1724–32.
- Miller JM, et al. Diagnostic performance of coronary angiography by 64-row CT. *N Engl J Med*. 2008;359(22):2324–36.
- Fox K, et al. Guidelines on the management of stable angina pectoris: executive summary: the task force on the management of stable angina pectoris of the european society of cardiology. *Eur Heart J*. 2006;27(11):1341–81.
- Fihn SD, et al. 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(24):e44–164.
- Genders TS, et al. A clinical prediction rule for the diagnosis of coronary artery disease: validation, updating, and extension. *Eur Heart J*. 2011;32(11):1316–30.
- He T, et al. Diagnostic models of the pre-test probability of stable coronary artery disease: a systematic review. *Clinics (Sao Paulo)*. 2017;72(3):188–96.
- Shemesh J, et al. Coronary calcium by spiral computed tomography predicts cardiovascular events in high-risk hypertensive patients. *J Hypertens*. 2004;22(3):605–10.
- Nakanishi R, et al. The relationship between coronary artery calcium score and the long-term mortality among patients with minimal or absent coronary artery risk factors. *Int J Cardiol*. 2015;185:275–81.
- Nakanishi R, et al. All-cause mortality by age and gender based on coronary artery calcium scores. *Eur Heart J Cardiovasc Imaging*. 2016;17(11):1305–14.
- McClelland RL, et al. 10-Year coronary heart disease risk prediction using coronary artery calcium and traditional risk factors: derivation in the MESA (multi-ethnic study of atherosclerosis) with validation in the HNR (Heinz Nixdorf recall) study and the DHS (Dallas Heart Study). *J Am Coll Cardiol*. 2015;66(15):1643–53.
- Al'Aref SJ, et al. 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. *Eur Heart J*. 2020;41(3):359–67.
- Nakanishi R, et al. Machine learning adds to clinical and CAC assessments in predicting 10-year CHD and CVD deaths. *JACC Cardiovasc Imaging*. 2021;14(3):615–25.
- Singal AG, et al. Machine learning algorithms outperform conventional regression models in predicting development of hepatocellular carcinoma. *Am J Gastroenterol*. 2013;108(11):1723–30.
- Wu Y, et al. Machine learning algorithms for the prediction of central lymph node metastasis in patients with papillary thyroid cancer. *Front Endocrinol (Lausanne)*. 2020;11:577537.
- Liu CY, et al. Deep learning powered coronary CT angiography for detecting obstructive coronary artery disease: the effect of reader experience, calcification and image quality. *Eur J Radiol*. 2021;142:109835.
- Breiman L. Random forests. *Mach Learn*. 2001;45(1):5–32.
- 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(2):e3.
- Willeminck MJ, et al. Coronary artery calcium: a technical argument for a new scoring method. *J Cardiovasc Comput Tomogr*. 2019;13(6):347–52.
- DeFilippis AP, et al. An analysis of calibration and discrimination among multiple cardiovascular risk scores in a modern multiethnic cohort. *Ann Intern Med*. 2015;162(4):266–75.
- Rana JS, et al. Accuracy of the atherosclerotic cardiovascular risk equation in a large contemporary. Multiethnic Popul J Am Coll Cardiol. 2016;67(18):2118–30.
- Al'Aref SJ, et al. Clinical applications of machine learning in cardiovascular disease and its relevance to cardiac imaging. *Eur Heart J*. 2019;40(24):1975–86.
- Mortazavi BJ, et al. Analysis of machine learning techniques for heart failure readmissions. *Circ Cardiovasc Qual Outcomes*. 2016;9(6):629–40.
- Arsanjani R, et al. Improved accuracy of myocardial perfusion SPECT for detection of coronary artery disease by machine learning in a large population. *J Nucl Cardiol*. 2013;20(4):553–62.
- Arsanjani R, et al. Prediction of revascularization after myocardial perfusion SPECT by machine learning in a large population. *J Nucl Cardiol*. 2015;22(5):877–84.
- Motwani M, et al. 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(7):500–7.
- Mu D, et al. Calcium scoring at coronary CT angiography using deep learning. *Radiology*. 2022;302(2):309–16.
- Craiem D, et al. Association between thoracic aorta calcium and thoracic aorta geometry in a cohort of asymptomatic participants at increased cardiovascular risk. *Rev Esp Cardiol (Engl Ed)*. 2016;69(9):827–35.
- Criqui MH, et al. Calcium density of coronary artery plaque and risk of incident cardiovascular events. *JAMA*. 2014;311(3):271–8.
- Hecht H, et al. 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(2):157–68.
- Hlatky MA, et al. Future directions for cardiovascular disease comparative effectiveness research: report of a workshop sponsored by the National Heart, Lung, and Blood Institute. *J Am Coll Cardiol*. 2012;60(7):569–80.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

