Improving cardiovascular risk prediction with machine learning: a focus on perivascular adipose tissue characteristics

Cong He[1], Fangye Wu[1], Linfeng Fu[1], Lingting Kong[1], Zefeng Lu[1], Yingpeng Qi[1] and Hongwei Xu[1*]

*Correspondence:
chinaxhw@163.com

[1] Department of Radiology, Shaoxing Second Hospital, 123 Yanan Rd, Shaoxing 312000, Zhejiang, China

Abstract

Background: Timely prevention of major adverse cardiovascular events (MACEs) is imperative for reducing cardiovascular diseases-related mortality. Perivascular adipose tissue (PVAT), the adipose tissue surrounding coronary arteries, has attracted increased amounts of attention. Developing a model for predicting the incidence of MACE utilizing machine learning (ML) integrating clinical and PVAT features may facilitate targeted preventive interventions and improve patient outcomes.

Methods: From January 2017 to December 2019, we analyzed a cohort of 1077 individuals who underwent coronary CT scanning at our facility. Clinical features were collected alongside imaging features, such as coronary artery calcium (CAC) scores and perivascular adipose tissue (PVAT) characteristics. Logistic regression (LR), Framingham Risk Score, and ML algorithms were employed for MACE prediction.

Results: We screened seven critical features to improve the practicability of the model. MACE patients tended to be older, smokers, and hypertensive. Imaging biomarkers such as CAC scores and PVAT characteristics differed significantly between patients with and without a 3-year MACE risk in a population that did not exhibit disparities in laboratory results. The ensemble model, which leverages multiple ML algorithms, demonstrated superior predictive performance compared with the other models. Finally, the ensemble model was used for risk stratification prediction to explore its clinical application value.

Conclusions: The developed ensemble model effectively predicted MACE incidence based on clinical and imaging features, highlighting the potential of ML algorithms in cardiovascular risk prediction and personalized medicine. Early identification of high-risk patients may facilitate targeted preventive interventions and improve patient outcomes.

Keywords: Perivascular adipose tissue, Coronary artery calcium, Major adverse cardiovascular event, Ensemble model, Computed tomography

Introduction

Cardiovascular diseases (CVDs), including myocardial infarction, acute coronary syndrome and stroke, are the primary contributors to mortality on a global scale. Early prevention and treatment of major adverse cardiovascular events (MACEs) are crucial for reducing the morbidity and mortality associated with CVDs [1]. Computed tomography angiography (CTA) has emerged as a noninvasive tool for screening major cardiovascular events, providing valuable insights into atherosclerosis and risk prediction [2].

Machine learning (ML), a branch of computer science, incorporates numerous variables using algorithms to identify their nonlinear relationships and complex interactions [3]. ML algorithms have been explored for their ability to predict heart failure [4], arrhythmia [5], postsurgical mortality [6], and composite CVD risk [7]. A prior investigation demonstrated that the ML exhibited superior performance compared to the logistic regression (LR), a traditional statistical technique, in forecasting patient outcomes [8]. The reason may stem from the inherent limitations of LR, which is known to be susceptible to multicollinearity. While some ML models, such as Linear Discriminant Analysis (LDA) and linear Support Vector Machine (SVM), are limited in capturing nonlinear relationships, others like Adaptive Boosting (AdaBoost) and ensemble methods are effective in modeling complex interactions among features. Additionally, multicollinearity, which poses challenges for traditional models like LR and LDA, is better managed by advanced ML models.

In the field of MACE prediction, the use of ML could improve CVD risk prediction by agnostically discovering new predictors of risk and learning about their complex interactions [9]. Previous studies have integrated myocardial perfusion imaging into the prediction of MACEs using ML techniques, demonstrating a moderate level of accuracy with area under the curve (AUC) values ranging from 0.73 to 0.79 [10, 11]. However, the clinical utilization of myocardial perfusion imaging remains limited. Recently, perivascular adipose tissue (PVAT), the adipose tissue surrounding the coronary artery, has attracted increased amounts of attention due to its hypothetical role in the pathogenesis of CVDs [12]. Using CTA images, PVAT characteristics can potentially be evaluated to determine if individuals are at increased risk of adverse cardiovascular outcomes [13, 14]. In addition, ensemble learning has been proposed because it is more effective and has a better generalization ability than single algorithms [15]. Coupling PVAT imaging features with the ML algorithm may hold promise for improving cardiovascular prevention by enhancing early detection capabilities [16]. However, the effectiveness of ensemble machine learning algorithms utilizing PVAT features in predicting patients at risk of MACEs remains uncertain.

Hence, we developed a novel ensemble model to predict MACE incidence utilizing ML algorithms. Through the utilization of stacking algorithms, our ensemble model incorporating both clinical and imaging features was established. The ability of the prediction model for MACE risk stratification may facilitate optimal preventive treatment intervention in future.

Methods

Participants and data collection

From January 2017 to December 2019, we analyzed a cohort of 1077 individuals who underwent coronary CT scanning at our facility. Exclusion criteria included the following: (1) coronary artery stenosis greater than 20% in at least one coronary major vessel;

(2) history of stroke, serious heart disease or tumor; (3) incomplete clinical and imaging data; and (4) loss to follow-up. After exclusion, 228 subjects were included. Figure 1 illustrates the flowchart of the subject selection.

The baseline features of the subjects, including age and sex, were collected. Risk factors, including hypertension, smoking history, alcohol consumption, diabetes, body mass index (BMI) and laboratory findings. The BMI was obtained using the equation weight (kg) divided by squared height (m^2). A TyG index was calculated by integrating fasting triglyceride (TG) by fasting blood glucose (FBG). The definition of hypertension was to have a systolic blood pressure (SBP) higher than 140 mmHg or a diastolic blood pressure (DPB) higher than 90 mmHg on two separate occasions or to be receiving antihypertensive medication. The definition of diabetes was to have an FBG $\geq$ 126 mg/dL or to be receiving any antidiabetic medication.

Follow-up and end points

Follow-ups through clinical visits were conducted for up to three years. The endpoint was the occurrence of MACEs [17], such as myocardial infarction, unstable angina, or stroke, which were confirmed through clinical presentation, imaging data, or other
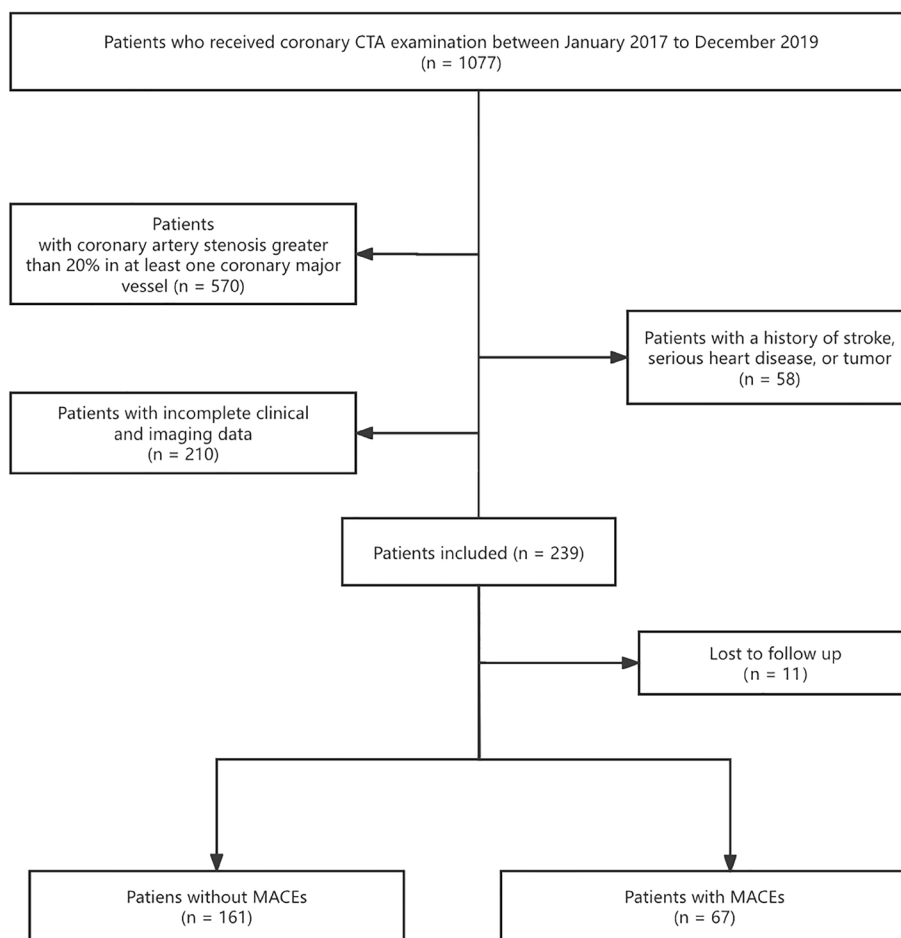

Fig. 1 Patients selection flowchart

pertinent data collected during each follow-up session and diagnosed by specialized cardiology physicians or neurologists.

CTA multiparameter analysis

The CTA scan was performed on the Siemens SOMATOM Definition AS (Siemens Medical Systems, Erlangen, Germany). The acquired images were postprocessed by Deepwise Coronary Artery Analysis Software (Deepwise Inc., Beijing, China). The coronary artery calcium (CAC) scores were calculated based on a sequence within the CTA examination. Prior to the injection of the contrast agent, a sequence was scanned during diastole using prospective gated acquisition. This non-contrast sequence was used to extract the CAC scores, which were automatically calculated by the software and confirmed manually.

The segmentation of PVAT was performed by manual adjustment on the basis of automatic segmentation. Specifically, if non-adipose tissues surrounding the coronary artery were encompassed as the region of interest during the automatic segmentation procedure, the radiologist will remove these tissues manually to ensure that only PVAT was included.

For the analysis of PVAT, the region of interests (ROIs) were automatically tracked by the software based on the fat threshold (from -190 to -30 HU). The ROI extends from 10 to 40 mm proximal to the coronary artery, with the radial distance of the ROI equivalent to the diameter of the coronary artery. Two experienced radiologists confirmed the results of the segmentation. Then, the fat attenuation index (FAI) and fat volume (FV) of PCAT could be automatically generated by the software. CAC scores, PCAT-FAI, and PCAT-FV were recorded as coronary artery imaging features. PVAT is visualized with an adipose tissue Hounsfield unit color table.

Model development

Feature selection

The features screened for model development included both clinical features and coronary artery imaging features. First, samples with 80% or above of missing values were excluded, among which were 186 subjects with a high percentage of missing values and 24 subjects with missing critical images. Afterwards, the remaining missing values were imputed by multiple imputation. After normalizing the continuous variables using z score normalization, 23 features were considered for selection. The least absolute shrinkage and selection operator (LASSO) regression was employed for feature selection in order to automatically remove redundant and irrelevant collinear features. The correlation between the selected features was determined by Spearman's correlation coefficient. Finally, the most predictive features were confirmed for further model development.

ML algorithm

We employed five ML algorithms to predict MACEs in our study based on the selected features. In addition to four commonly used models, namely, AdaBoost, GNB, LDA, and SVM models [18], we introduced a stacking ensemble model. The ensemble model was built by stacking the classification probabilities of the outputs of the other four machine learning models. Compared with other ML models, this approach can enhance the

prediction accuracy by combining diverse models, reducing the possibility of overfitting, and improving the model stability and generalizability [19]. Besides, the traditional LR model was also developed. Significant variables identified in the univariate analysis were incorporated into the multivariable LR analysis.

Model fitting

A fivefold cross-validation method was used to assign the training and validation cohorts using the same random seed across all splits to ensure consistency in groupings. Herein, based on the occurrence of MACEs within the first three years of initial examination, the MACE prediction task was conceptualized as a binary classification problem, with all machine learning models generating a normalized risk probability within the range of 0 to 1. The combination of hyperparameters that yielded the best average AUC was selected for model fitting.

Model evaluation

Model performance was evaluated using the receiver operating characteristic (ROC) curves. Calibration curves were employed for qualitative evaluation of model calibration, while decision curve analysis (DCA) was utilized to assess the utility of the models across various threshold probabilities. Confusion matrices were employed for comparative analysis of model performances. Furthermore, model efficiency metrics were computed to quantitatively assess the predictive capabilities of each model.

Additionally, in order to further validate the model's clinical applicability, model with the best performance was used to predict 3-year MACE incidence using Kaplan–Meier survival curves. By applying the cut-off values determinate by the X-tile software [20], the subjects were divided into high, middle, and low ML risk groups. The overall log-rank test was applied to compare the differences between Kaplan–Meier survival curves. Figure 2 shows the study flowchart.

Statistical analysis

SPSS (version 26.0), R software (version 4.0.2), X-tile (version 3.6.1), and MedCalc (version 19.5.6) used to analyze the data. A comparison of variables between patients with and without MACE was carried out based on the results of the follow-up. A normal distribution was determined by the Kolmogorov–Smirnov test. A continuous variable is expressed as the mean+standard deviation in normal distributions and median in non-normal distributions. A categorical variable is expressed as frequency counts and percentages. T test and a Chi-square test were used to analyze differences between groups. The AUCs of the different models were compared using DeLong tests. The significance of a study is determined by *P* values that have two tails no greater than 0.05. We set the *P*-value threshold at 0.05 because it is a widely accepted standard in many scientific fields and balances statistical rigor with the practical significance of our findings.

Results

Clinical features

Out of the initial dataset of 1077 patients, 210 patients were excluded due to a high missing values or missing critical images for the analysis. A bias analysis was
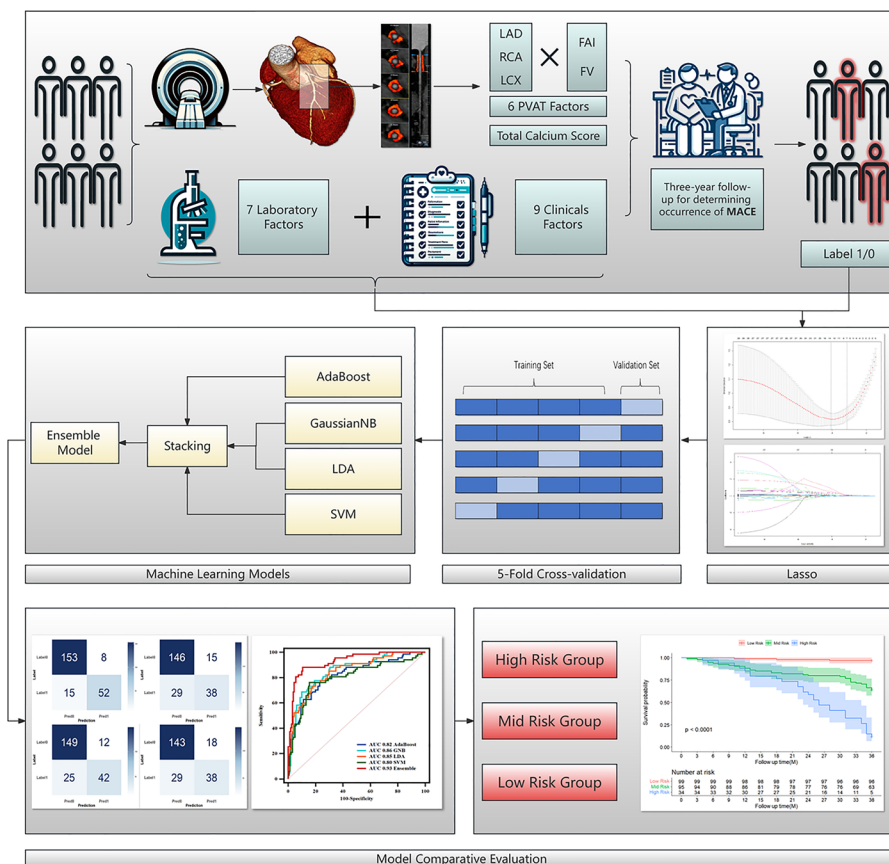
Fig. 2 Schematic of the workflow of the study

conducted to examine the demographic characteristics of patients with complete data in comparison to those with missing data. The results indicated no significant differences between the two, suggesting that the missing data were likely random and did not introduce systematic biases (Supplementary materials, Table S1).

In our analysis, 228 participants were involved with an average age of (62.58 ± 11.63) years and 146 men (64.0%). Regarding factors associated with arteriosclerosis, 56.6% of patients had hypertension, 25.9% were current smokers, 27.2% currently consumed alcohol, and 28.9% had diabetes.

Patients who experienced MACEs were older $(68.13 \pm 10.23$ vs. 60.26 ± 11.42 years, $P < 0.001$) compared to those who did not experience MACEs. Besides, there was a higher proportion of males (74.6% vs. 59.6%, $P = 0.032$) among patients who experienced MACEs compared to those who did not. Furthermore, individuals in the MACE group were more inclined to engage in unhealthy lifestyle behaviors (smoking and alcohol use, 35.8% vs. 21.7%, $P = 0.027$ and 38.8% vs. 22.4%, $P = 0.011$, respectively), and the prevalence of hypertension was notably higher in the MACE group compared to the non-MACE group (76.1% vs. 48.4%, $P < 0.001$). The clinical characteristics of the patients are outlined in Table 1, and are in line with known literature [17].

Coronary artery imaging features

As presented in Table 2, compared with patients without MACE, those with MACE outcomes had a much greater CAC score (78.76 [5.43–371.10] vs. 0 [0–47.38], $P < 0.001$). With regard to PVAT features, the median PVAT-FAI was greater in the MACE group compared to the non-MACE group (-71.56 [-76.87–66.51] vs. -80.58 [-87.34–75.05] HU for the RCA, -65.25 [-71.70–61.38] vs. -71.94 [-76.87–66.67] HU for the LCX, and -67.68 [-72.59–63.09] vs. 76.17 [-81.37–71.31] HU for the LAD, respectively, all $P < 0.001$). In addition, the PVAT volume was smaller in the MACE group (2.13 [1.62–2.94] mm^3 for the RCA, 1.27 [0.95–1.65] mm^3 for the LCX, and 1.65 [1.32–2.34] mm^3 for the LAD) than in the non-MACE group (2.84 [2.38–3.37] mm^3 for the RCA, 1.59 [1.19–1.94] mm^3 for the LCX, and 2.24 [1.86–2.82] mm^3 for the LAD, all $P < 0.001$). Figure 3 shows the typical CTA images obtained during the analysis.

Development and performance of models

By using LASSO, irrelevant or redundant features were eliminated. Finally, seven features were chosen for ML model development, including 3 clinical features (age, hypertension status, and smoking status) and 4 coronary artery imaging features (CAC scores, PCAT-FAI$_{RCA}$, PCAT-FAI$_{LAD}$, and PCAT-FV$_{RCA}$). Feature importance was determined using LASSO regression (Supplementary materials, Figure S1). The reason to use LASSO for all the ML models was driven by the need for a model-agnostic feature selection process that could be consistently applied across different ML algorithms, ensuring a fair comparison. Correlations among the selected features were evaluated using

Table 1 Clinical features of the study participants

Variables	Overall (n = 228)	Non-MACE group (n = 161)	MACE group (n = 67)	P value
Age, years	62.58 ± 11.63	60.26 ± 11.42	68.13 ± 10.23	< 0.001**
Men, n (%)	146 (64.0)	96 (59.6)	50 (74.6)	0.032*
Hypertension, n (%)	129 (56.6)	78 (48.4)	51 (76.1)	< 0.001**
Smoker, n (%)	59 (25.9)	35 (21.7)	24 (35.8)	0.027*
Alcohol use, n (%)	62 (27.2)	36 (22.4)	26 (38.8)	0.011*
Diabetics, n (%)	66 (28.9)	42 (26.1)	24 (35.8)	0.140
TyG index, value	4.74 (4.58–4.92)	4.74 (4.57–4.92)	4.75 (4.59–4.95)	0.820
BMI, kg/m^2	24.30 (22.50–26.53)	24.20 (22.40–26.50)	24.90 (22.90–26.75)	0.102
SBP, mmHg	136.48 ± 19.92	134.83 ± 19.87	140.43 ± 19.60	0.053
DBP, mmHg	80.37 ± 11.14	80.77 ± 11.18	79.40 ± 11.05	0.400
Total cholesterol, mmol/L	4.49 (3.76–5.44)	4.46 (3.82–5.40)	4.55 (3.65–5.49)	0.649
Triglyceride, mmol/L	1.48 (1.11–1.96)	1.47 (1.07–1.98)	1.49 (1.19–1.95)	0.909
HDL, mmol/L	1.09 (0.89–1.28)	1.08 (0.89–1.27)	1.09 (0.90–1.30)	0.912
LDL, mmol/L	2.90 (2.38–3.61)	2.89 (2.43–3.54)	2.91 (2.34–3.85)	0.882
Hs-CRP, mg/L	0.90 (0.50–1.80)	1.00 (0.50–2.40)	0.90 (0.50–1.70)	0.472
FBG, mmol/L	5.25 (4.78–6.18)	5.23 (4.82–6.05)	5.37 (4.70–6.35)	0.892

Data are presented as number (%), median (P25–P75) or means ± SD

MACE major adverse cardiovascular event, *TyG* triglyceride–glucose index, *BMI* body mass index, *SBP* systolic blood pressure, *DBP* diastolic blood pressure, *HDL* high density lipoprotein, *LDL* low density lipoprotein, *hs-CRP* high sensitivity C-reactive protein, *FBG* fasting blood glucose

* Significant difference (P < 0.05)

** Extremely significant difference (P < 0.001)

Table 2 Coronary artery imaging features of the study participants

Variables	Overall (n = 228)	Non-MACE group (n = 161)	MACE group (n = 67)	P value
CAC Score	4.79 (0–125.88)	0 (0–47.38)	78.76 (5.43–371.10)	< 0.001[**]
PVAT-FAI, HU				
RCA	− 77.89 (− 84.83–71.47)	− 80.58 (− 87.34–75.05)	− 71.56 (− 76.87–66.51)	< 0.001[**]
LCX	− 70.45 (− 76.11–64.56)	− 71.94 (− 76.87–66.67)	− 65.25 (− 71.70–61.38)	< 0.001[**]
LAD	− 73.978 (− 79.13–68.28)	− 76.17 (− 81.37–71.31)	− 67.68 (− 72.59–63.09)	< 0.001[**]
PVAT-FV, cm^3				
RCA	2.76 (2.17–3.25)	2.84 (2.38–3.37)	2.13 (1.62–2.94)	< 0.001[**]
LCX	1.48 (1.11–1.88)	1.59 (1.19–1.94)	1.27 (0.95–1.65)	< 0.001[**]
LAD	2.17 (1.63–2.74)	2.24 (1.86–2.82)	1.65 (1.32–2.34)	< 0.001[**]

Data are presented as median (P25–P75)

MACE major adverse cardiovascular event, *CAC* coronary artery calcium, *FAI* fat attenuation index, *FV* fat volume, *HU* Hounsfield unit, *PVAT* perivascular adipose tissue, *RCA* right coronary artery, *LAD* left anterior descending artery, *LCX* left circumflex artery

[**] extremely significant difference (*P* < 0.001)

Spearman's correlation coefficients (Supplementary materials, Figure S2). Additionally, in LR analysis, variables that exhibited statistical significance in the univariate analysis were subsequently incorporated into the multivariate logistic regression analysis, which is a common approach to develop a classical LR model. Age, hypertension, CAC scores, PCAT-FAI$_{LAD}$, and PCAT-FAI$_{RCA}$ were independent risk factors for MACE prediction (Supplementary materials, Table S2).

As Adaboost inherently calculates feature importance during its training process, a supplementary analysis was conducted using AdaBoost for the feature selection. There was no significant difference of the AUCs between LASSO and AdaBoost selections (Supplementary materials, Table S3). Besides, a supplementary analysis using LASSO for the LR feature selection was also conducted. No significant difference of the AUCs was found between LASSO and univariate-multivariate analysis (Supplementary materials, Table S4).

We compared the performance of the ML models to classical LR and a well-established model, the Framingham Risk Score [21]. Table 3 summarizes commonly used metrics for the model evaluation. Upon evaluation of LR, Framingham Risk Score, AdaBoost, GNB, LDA, SVM, and an ensemble algorithm, it was determined that the ensemble algorithm exhibited superior predictive performance metrics. Specifically, the ensemble algorithm demonstrated an AUC of 0.94, accuracy of 0.87, precision of 0.74, recall of 0.87, F1-score of 0.80, sensitivity of 0.87, specificity of 0.88, PPV of 0.74, and NPV of 0.94 in the training cohort. In the validation cohort, the ensemble algorithm maintained strong performance with an AUC of 0.93, accuracy of 0.88, precision of 0.77, recall of 0.85, F1-score of 0.81, sensitivity of 0.85, specificity of 0.89, PPV of 0.77, and NPV of 0.94. (Fig. 4 and Figure S3). The model parameters and settings were provided in detail in Supplementary Appendix S1. The formulas of the ensemble model are available in Supplementary Appendix S2.

As an additional demonstration of the effectiveness of the models, the ROC curves are plotted in Fig. 5A. Although the AUC of the ensemble model did not exhibit the most substantial advantage in the training cohort when compared with all the other models,
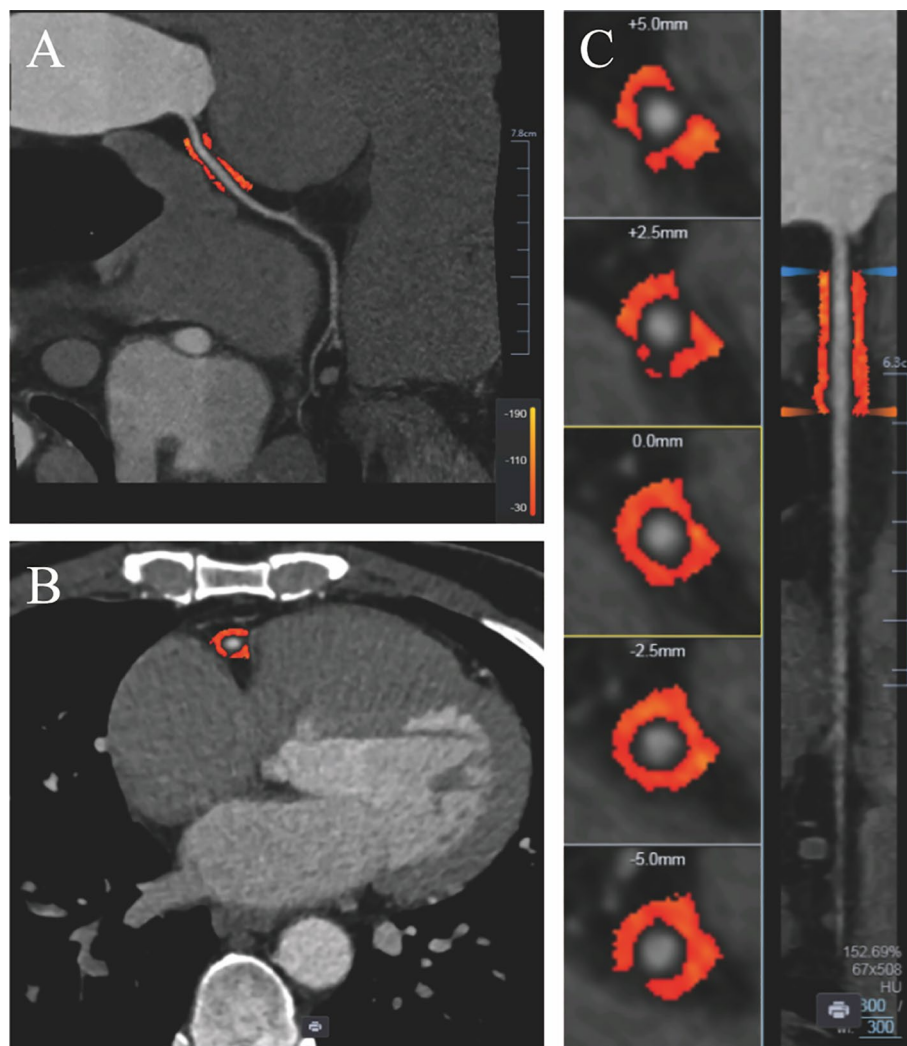
Fig. 3 PVAT analysis of the RCA. **A** Longitudinal view of RCA PVAT measurements. **B** Cross-sectional view of RCA PVAT measurements. **C** Curved multiplane review of RCA PVAT measurements. An adipose tissue HU color table is shown with a color bar to visualize PVAT. Based on the fat threshold (from − 190 to − 30 HU), the ROI was automatically tracked using the software. The length of the ROI is 10–40 mm proximal to the coronary artery. The radial distance of the ROI is equal to the diameter of the coronary artery. The PVAT-FAI was − 109 HU. The PVAT volume was the total volume of adipose voxels within the ROI. RCA, right coronary artery; PVAT, perivascular adipose tissue; HU, Hounsfield unit; ROI, region of interest; FAI, fat attenuation index

it demonstrated superior performance in the validation cohort (Table 4). The ensemble model's predicted probability was close to the actual probability according to calibration curves, which describe how well the prediction agrees with reality (Supplementary materials online, Figure S4). The clinical net benefit can be calculated using DCA to determine whether using a model is beneficial. In our study, the ensemble model demonstrated superior decision-making capabilities when compared to the other four algorithms, as illustrated in Fig. 5. Consequently, the ensemble algorithm was chosen for further model development.

Subsequently, the ensemble model was utilized for risk stratification prediction in order to investigate its potential clinical utility. Patients were categorized into low-risk, intermediate-risk, and high-risk cohorts for the prediction of MACE. Cut-off values of

Table 3 Performance comparisons of logistic regression model and machine learning models

	AUC	95% CI	ACC	PRE	REC	F1-score	SENS	SPEC	PPV	NPV
Training cohort										
LR	0.89	0.84–0.93	0.81	0.63	0.82	0.71	0.82	0.80	0.63	0.91
FS	0.62	0.54–0.70	0.55	0.35	0.66	0.46	0.67	0.50	0.35	0.78
AdaBoost	0.95	0.92–0.98	0.87	0.87	0.67	0.76	0.67	0.96	0.87	0.88
GNB	0.88	0.83–0.93	0.83	0.75	0.63	0.68	0.63	0.91	0.75	0.85
LDA	0.90	0.85–0.94	0.84	0.78	0.63	0.69	0.63	0.93	0.78	0.86
SVM	0.93	0.89–0.97	0.90	0.87	0.78	0.82	0.78	0.95	0.87	0.91
Ensemble	0.94	0.91–0.97	0.87	0.74	0.87	0.80	0.87	0.88	0.74	0.94
Validation cohort										
LR	0.84	0.79–0.90	0.75	0.56	0.73	0.63	0.73	0.76	0.57	0.87
FS	0.61	0.53–0.69	0.55	0.35	0.66	0.46	0.67	0.50	0.35	0.78
AdaBoost	0.82	0.76–0.88	0.78	0.68	0.48	0.56	0.48	0.91	0.68	0.81
GNB	0.86	0.80–0.91	0.82	0.76	0.58	0.66	0.58	0.93	0.76	0.84
LDA	0.85	0.79–0.90	0.81	0.72	0.57	0.63	0.57	0.91	0.72	0.83
SVM	0.80	0.93–0.87	0.79	0.68	0.57	0.62	0.57	0.89	0.68	0.83
Ensemble	0.93	0.90–0.97	0.88	0.77	0.85	0.81	0.85	0.89	0.77	0.94

AUC, area under the curve, *ACC*, accuracy, *CI* confidence interval, *PRE* precision, *SENS* sensitivity, *SPEC* specificity, *REC* recall, *PPV* positive prediction value, *NPV* negative prediction value, *AdaBoost* adaptive boosting, *GNB* Gaussian naive Bayes, *LDA* linear discriminant analysis, *SVM* linear support vector machine, *LR* logistic regression, *FS* Framingham Score

the three risk groups were 0.14 and 0.92. As presented in Fig. 6, the risk probability differed significantly among the three groups ($P < 0.0001$). The 3-year MACE risk for the middle-risk group was found to be 10.05 times (95% CI 8.46–23.33) higher compared to the low-risk group. Similarly, the 3 year MACE risk for the high-risk group was 51.21 times higher (95% CI 22.15–118.42) than that of the low-risk group. These findings suggest that the ensemble model could offer more valuable information and thus be more applicable in clinical settings.

Discussion

Our study developed an ensemble model based on clinical and imaging features to predict MACEs accurately. Older age, current smoker, and a higher incidence of hypertension were identified as significant risk factors for MACEs, consistent with previous research. Additionally, patients with MACEs exhibited significantly higher CAC scores and median PVAT values, suggesting these imaging biomarkers may play a role in risk stratification and prognosis assessment. As compared to the other algorithms, the ensemble model performed the best in terms of prediction accuracy.

Seven common features were utilized in constructing the predictive ML models in our study. The features can be classified into three categories: general information of patients (age); lifestyle habits (hypertension and smoking); and CTA imaging data (CAC scores, PCAT-FAI$_{RCA}$, PCAT-FAI$_{LAD}$, and PCAT-FV$_{RCA}$). For comparison, a LR model was limited to only five features using univariate and multivariate regression analysis. As presented in our study, the inclusion of additional features has the potential to enhance prediction accuracy by capturing more complex relationships and interactions within the data. Furthermore, the incorporation of more features may enhance the model's resilience to data fluctuations and enhance its applicability to novel datasets, thereby
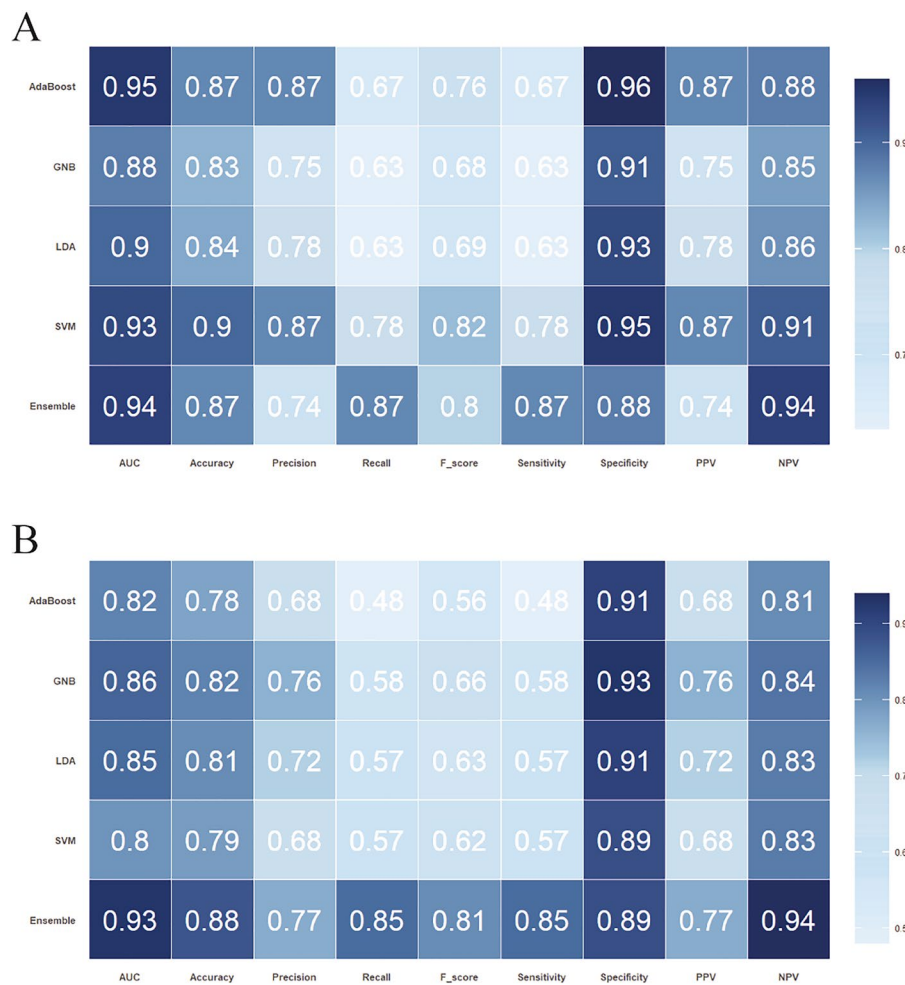
Fig. 4 Heatmap comparing the performance of different ML algorithm models in the training cohort (**A**) and validation cohort (**B**). AUC, area under the curve; AdaBoost, adaptive boosting; GNB, Gaussian naive Bayes; LDA, linear discriminant analysis; SVM, linear support vector machine; PPV, positive prediction value; NPV, negative prediction value

mitigating the risk of overfitting. Additionally, it was observed that the laboratory results were not incorporated into the models, potentially attributable to the early stage of the disease that leads to lack of significant variations in the laboratory findings. It highlights a strength of our ML-based approach to dimensionality reduction. By preserving features that remain unaffected by laboratory findings, our model demonstrates robustness and relevance, even at early stages of the disease when laboratory variations are minimal. The clinical features observed in our study align with established risk factors for cardiovascular events. Older age and smoker have long been recognized as significant contributors to cardiovascular morbidity and mortality [22, 23]. Moreover, the increase in the incidence of hypertension among patients with MACE highlights the continued importance of blood pressure management in mitigating cardiovascular risk [24]. Our findings reinforce the need for targeted interventions aimed at addressing modifiable risk factors like smoking and drinking alcohol, to reduce the burden of MACE in high-risk populations [25, 26].
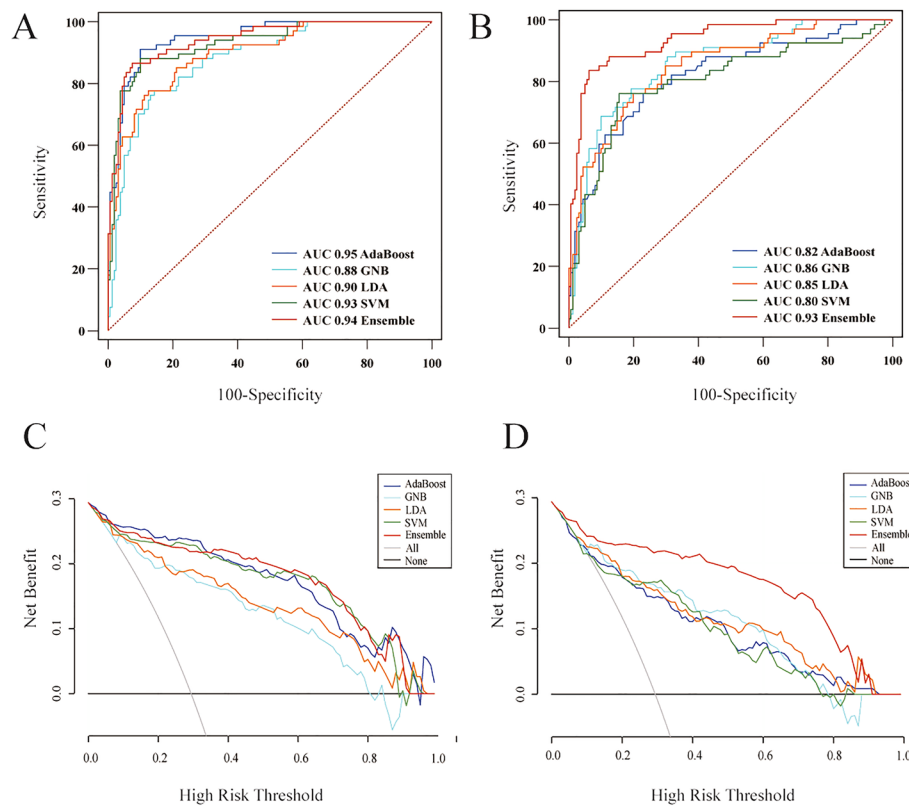
Fig. 5 Performances of different ML algorithm models. Plots showing the ROC curves of different models in the training cohort (**A**) and validation cohort (**B**). Decision curve analysis of different models in the training cohort (**C**) and validation cohort (**D**). The X-axis represents the threshold probability; the Y-axis represents the net benefit. AUC, area under the curve; AdaBoost, adaptive boosting; GNB, Gaussian naive Bayes; LDA, linear discriminant analysis; SVM, linear support vector machine

Table 4 DeLong tests to compare the AUCs of logistic regression model and machine learning models

Comparison	Training cohort		Validation cohort	
	AUC	*P* value	AUC	*P* value
Ensemble vs LR	0.94	< 0.0001**	0.93	< 0.0001**
	0.89		0.84	
Ensemble vs FS	0.94	< 0.0001**	0.93	< 0.0001**
	0.62		0.61	
Ensemble vs AdaBoost	0.94	0.03944*	0.93	< 0.0001**
	0.95		0.82	
Ensemble vs GNB	0.94	0.0001**	0.93	< 0.0001**
	0.88		0.86	
Ensemble vs LDA	0.94	0.0004**	0.93	< 0.0001**
	0.90		0.85	
Ensemble vs SVM	0.94	0.2118	0.93	< 0.0001**
	0.93		0.80	

AUC area under the curve, *AdaBoost* adaptive boosting, *GNB* Gaussian naive Bayes, *LDA* linear discriminant analysis, *SVM* linear support vector machine, *LR* logistic regression, *FS* Framingham Score

* significant difference (*P* < 0.05)

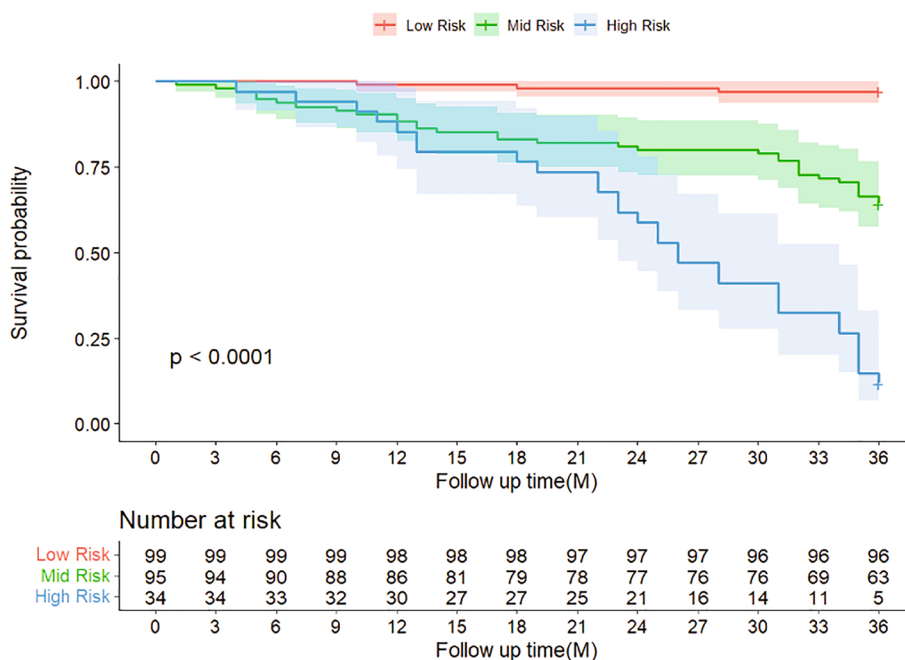** extremely significant difference (*P* < 0.001)

Fig. 6 Survival curves by risk group for the validation cohort stratified into low-risk, middle-risk and high-risk groups according to the optimal cut-off points

Coronary artery imagines, particularrly CTA, can contribute significantly to risk stratification and prognosis. Our analysis revealed distinct patterns in CAC scores and PVAT characteristics between patients with and without a 3-year MACE risk in a population that did not exhibit disparities in laboratory results. The heightened CAC burden observed in the MACE group signifies advanced atherosclerotic processes and plaque vulnerability, consistent with previous studies linking CAC scores to adverse cardiovascular outcomes [27, 28]. Additionally, alterations in PVAT composition, as reflected in our findings, underscores PVAT's emerging role as a marker of metabolic disorders and inflammation [29, 30]. It has been reported that an increased LAD and RCA PVAT attenuation is associated with an increased CVD risk [31]. In our research, we substantiated the additional predictive significance of PCAT-FAI$_{RCA}$, PCAT-FAI$_{LAD}$, and PCAT-FV$_{RCA}$ in patients experiencing MACEs. The comparative analysis of our models with the Framingham Risk Score highlights the added value of incorporating CTA characteristics into cardiovascular risk prediction models. The attenuation of PVAT and alterations in intracellular lipid deposition are attributed to early and persistent inflammation [32]. Thus, PVAT features may provide important complementary information about coronary artery disease [33].

The development of a predictive model integrating clinical and imaging features represents a significant advancement in personalized risk assessment. Recently, ML algorithms have been used to develop regression and classification models for clinical prediction [34, 35]. In this study, we compare LR, a classical statistical method, with ML algorithms including ensemble methods (AdaBoost), probabilistic models (Gaussian Naive Bayes), and linear classifiers (LDA and SVM). ML algorithms, particularly those using ensemble techniques, are generally less affected by multicollinearity due to their

ability to combine multiple models and their inherent design to handle high-dimensional data. While the traditional LR model proved to be effective in predicting MACEs, the utilization of ML model, specifically the ensemble model, showcases the potential of data-driven approaches in health care decision support systems. Clinical data are inherently diverse and often suffer from imbalances among classes or outcomes. Although the cross-validation in our study attempts to mitigate this issue by partitioning data into multiple folds, the results in each fold may still not adequately representing the full spectrum of clinical scenarios. That's the reason why the four commonly used ML models seem to perform worse in the validation cohort in our study [36]. The ensemble model, on the other hand, could mitigate this issue by amalgamating multiple base models to enhance overall performance. In our study, multiple base classifiers were employed as level-0 models. The predictions of these base classifiers were then used as input features for a level-1 metamodel, which generated the final prediction. By aggregating predictions from multiple models, the ensemble model effectively smooths out noise and leads to consistent performance across both training and validation cohorts [37]. The model's high accuracy, precision, and sensitivity underscore its potential clinical utility in guiding risk management strategies and optimizing patient outcomes [19, 38].

In our study, we utilized fivefold cross-validation to evaluate the efficacy and reliability of our predictive model. Cross-validation involves splitting the dataset into multiple subsets and using each fold as both training and validation data. In each iteration, the model was trained on four folds and validated on onefold. The scores from the five validation folds were averaged to obtain a single representative score, which was used for final model prediction. Afterwards, the performance metrics are derived from the final model. This approach enables more efficient utilization of the available data.

In light of MACEs' high mortality risk, early detection of MACEs may facilitate risk stratification, clinical decisions, and improved patient outcomes. For instance, implementing stringent lifestyle management practices, such as controlling blood pressure and cessation of smoking, adhering to a nutritious and well-rounded diet, engaging in regular physical activity to decrease overall body fat, and utilizing appropriate medication therapies (such as GLP-1 agonists) that focus on lipid metabolism and diminish fat build-up, may be efficaciously employed in the prevention of MACEs in patients at high risk. GLP-1 agonists, traditionally used for glycemic control in type 2 diabetes, have shown significant cardiovascular benefits. Combining the pharmacological advancements with stringent lifestyle interventions may offer a synergistic approach to significantly mitigate the risk of MACEs in high-risk patients.

Our study exhibits certain limitations. First, since the individuals under study were exclusively sourced from a single medical facility, there exists a potential for bias in the evaluation of predictive efficacy. To reduce the likelihood of overfitting and providing robust estimates of predictive accuracy, we utilized cross-validation to evaluate the efficacy and reliability of our predictive model. Second, directly comparing the ensemble ML model to individual variables can be challenging due to its inherent complexity and nonlinear relationships. The ensemble model combines the predictions of multiple base ML models, which obfuscates the contribution of each individual variable to the overall model performance. Third, the performance/validation metrics for the automatic segmentation tool of PVAT were not calculated. However, in our study the automatic

segmentation is based on well-established rules of PVAT segmentation [39]. Given that the segmentation is rule-based, there was no need to independently verify its accuracy. Besides, in our research process, there were no instances requiring manual adjustments, which further attests to its robustness and reliability. Finally, although the ensemble model demonstrated promising performance in risk stratification, the translation of the model's predictions into meaningful clinical decisions and interventions needs to be further explored.

Conclusion

Our study established an ensemble model for the prediction of the occurrence of MACEs using clinical and PVAT features. By integrating clinical data, advanced imaging modalities such as CTA, and machine learning techniques, we advance towards a more nuanced understanding of cardiovascular risk prediction. These findings may improve patient outcomes through early disease prevention and therapeutic interventions.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12938-024-01273-5.

Supplementary Material 1.

Declarations

Ethics approval and consent to participate
This study was approved by the Ethics Committee of Shaoxing Second Hospital. The data were anonymous, so informed consent was waived.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

References

1. Chen W, Li R, Yin K, Liang J, Li H, Chen X, Sheng Z, Yu H, Mu D. Clinical feasibility of using effective atomic number maps derived from non-contrast spectral computed tomography to identify non-calcified atherosclerotic plaques: a preliminary study. Quant Imag Med Surg. 2022;12:2280–7.
2. Chen Q, Pan T, Wang YN, Schoepf UJ, Bidwell SL, Qiao H, Feng Y, Xu C, Xu H, Xie G, et al. A coronary CT angiography radiomics model to identify vulnerable plaque and predict cardiovascular events. Radiology. 2023;307:e221693.
3. Liu W, Laranjo L, Klimis H, Chiang J, Yue J, Marschner S, Quiroz JC, Jorm L, Chow CK. Machine-learning versus traditional approaches for atherosclerotic cardiovascular risk prognostication in primary prevention cohorts: a systematic review and meta-analysis. Eur Heart J Qual Care Clin Outcomes. 2023;9:310–22.

4.   Hamatani Y, Nishi H, Iguchi M, Esato M, Tsuji H, Wada H, Hasegawa K, Ogawa H, Abe M, Fukuda S, Akao M. Machine learning risk prediction for incident heart failure in patients with atrial fibrillation. JACC Asia. 2022;2:706–16.

5.   Li L, Zhang Z, Zhou L, Zhang Z, Xiong Y, Hu Z, Yao Y. Use of machine learning algorithms to predict life-threatening ventricular arrhythmia in sepsis. Eur Heart J Digit Health. 2023;4:245–53.

6.   Cho SM, Austin PC, Ross HJ, Abdel-Qadir H, Chicco D, Tomlinson G, Taheri C, Foroutan F, Lawler PR, Billia F, et al. Machine learning compared with conventional statistical models for predicting myocardial infarction readmission and mortality: a systematic review. Can J Cardiol. 2021;37:1207–14.

7.   Gautam N, Mueller J, Alqaisi O, Gandhi T, Malkawi A, Tarun T, Alturkmani HJ, Zulqarnain MA, Pontone G, Al'Aref SJ. Machine learning in cardiovascular risk prediction and precision preventive approaches. Curr Atheroscler Rep. 2023;25:1069–81.

8.   Leonard G, South C, Balentine C, Porembka M, Mansour J, Wang S, Yopp A, Polanco P, Zeh H, Augustine M. Machine learning improves prediction over logistic regression on resected colon cancer patients. J Surg Res. 2022;275:181–93.

9.   Alaa AM, Bolton T, Di Angelantonio E, Rudd JHF, van der Schaar M. Cardiovascular disease risk prediction using automated machine learning: a prospective study of 423,604 UK biobank participants. PLoS ONE. 2019;14:e0213653.

10. Singh A, Miller RJH, Otaki Y, Kavanagh P, Hauser MT, Tzolos E, Kwiecinski J, Van Kriekinge S, Wei CC, Sharir T, et al. Direct risk assessment from myocardial perfusion imaging using explainable deep learning. JACC Cardiovasc Imag. 2023;16:209–20.

11. Rios R, Miller RJH, Hu LH, Otaki Y, Singh A, Diniz M, Sharir T, Einstein AJ, Fish MB, Ruddy TD, et al. Determining a minimum set of variables for machine learning cardiovascular event prediction: results from REFINE SPECT registry. Cardiovasc Res. 2022;118:2152–64.

12. Rachwalik M, Matusiewicz M, Jasiński M, Hurkacz M. Evaluation of the usefulness of determining the level of selected inflammatory biomarkers and resistin concentration in perivascular adipose tissue and plasma for predicting postoperative atrial fibrillation in patients who underwent myocardial revascularisation. Lipids Health Dis. 2023;22:2.

13. Kurata A. Deep learning-based CT noise reduction for perivascular adipose tissue evaluation. Acad Radiol. 2024;31:446–7.

14. Mancio J, Oikonomou EK, Antoniades C. Perivascular adipose tissue and coronary atherosclerosis. Heart. 2018;104:1654–62.

15. Zhang Z, Liu J, Xi J, Gong Y, Zeng L, Ma P. Derivation and validation of an ensemble model for the prediction of agitation in mechanically ventilated patients maintained under light sedation. Crit Care Med. 2021;49:e279–90.

16. Oikonomou EK, Williams MC, Kotanidis CP, Desai MY, Marwan M, Antonopoulos AS, Thomas KE, Thomas S, Akoumi-anakis I, Fan LM, et al. A novel machine learning-derived radiotranscriptomic signature of perivascular fat improves cardiac risk prediction using coronary CT angiography. Eur Heart J. 2019;40:3529–43.

17. Heo J, Yoo J, Lee H, Lee IH, Kim JS, Park E, Kim YD, Nam HS. Prediction of hidden coronary artery disease using machine learning in patients with acute ischemic stroke. Neurology. 2022;99:e55–65.

18. Wang K, Tian J, Zheng C, Yang H, Ren J, Liu Y, Han Q, Zhang Y. Interpretable prediction of 3-year all-cause mortality in patients with heart failure caused by coronary heart disease based on machine learning and SHAP. Comput Biol Med. 2021;137:104813.

19. Zhou J, Hu B, Feng W, Zhang Z, Fu X, Shao H, Wang H, Jin L, Ai S, Ji Y. An ensemble deep learning model for risk stratification of invasive lung adenocarcinoma using thin-slice CT. NPJ Digit Med. 2023;6:119.

20. Camp RL, Dolled-Filhart M, Rimm DL. X-tile: a new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. Clin Cancer Res. 2004;10:7252–9.

21. Atehortúa A, Gkontra P, Camacho M, Diaz O, Bulgheroni M, Simonetti V, Chadeau-Hyam M, Felix JF, Sebert S, Lekadir K. Cardiometabolic risk estimation using exposome data and machine learning. Int J Med Inform. 2023;179:105209.

22. Jiang ZZ, Zhu JB, Shen HL, Zhao SS, Tang YY, Tang SQ, Liu XT, Jiang TA. A high triglyceride-glucose index value is associated with an increased risk of carotid plaque burden in subjects with prediabetes and new-onset type 2 diabetes: a real-world study. Front Cardiovasc Med. 2022;9:832491.

23. Gomes D, Wang S, Goodspeed L, Turk KE, Wietecha T, Liu Y, Bornfeldt KE, O'Brien KD, Chait A, den Hartigh LJ. Comparison between genetic and pharmaceutical disruption of Ldlr expression for the development of atherosclerosis. J Lipid Res. 2022;63:100174.

24. Kuwabara M, Kodama T, Ae R, Kanbay M, Andres-Hernando A, Borghi C, Hisatome I, Lanaspa MA. Update in uric acid, hypertension, and cardiovascular diseases. Hypertens Res. 2023;46:1714–26.

25. Møller AL, Andersson C. Importance of smoking cessation for cardiovascular risk reduction. Eur Heart J. 2021;42:4154–6.

26. Day E, Rudd JHF. Alcohol use disorders and the heart. Addiction. 2019;114:1670–8.

27. Bell KJL, White S, Hassan O, Zhu L, Scott AM, Clark J, Glasziou P. Evaluation of the incremental value of a coronary artery calcium score beyond traditional cardiovascular risk assessment: a systematic review and meta-analysis. JAMA Intern Med. 2022;182:634–42.

28. Hussain B, Mahmood A, Flynn MG, Alexander T. Coronary artery calcium scoring in asymptomatic patients. HCA Healthc J Med. 2023;4:341–52.

29. Adachi Y, Ueda K, Nomura S, Ito K, Katoh M, Katagiri M, Yamada S, Hashimoto M, Zhai B, Numata G, et al. Beiging of perivascular adipose tissue regulates its inflammation and vascular remodeling. Nat Commun. 2022;13:5117.

30. Koenen M, Hill MA, Cohen P, Sowers JR. Obesity, adipose tissue and vascular dysfunction. Circ Res. 2021;128:951–68.

31. Oikonomou EK, Marwan M, Desai MY, Mancio J, Alashi A, Hutt Centeno E, Thomas S, Herdman L, Kotanidis CP, Thomas KE, et al. Non-invasive detection of coronary inflammation using computed tomography and prediction of residual cardiovascular risk (the CRISP CT study): a post-hoc analysis of prospective outcome data. Lancet. 2018;392:929–39.

32. Antoniades C, Antonopoulos AS, Deanfield J. Imaging residual inflammatory cardiovascular risk. Eur Heart J. 2020;41:748–58.

33. Ichikawa K, Miyoshi T, Osawa K, Nakashima M, Miki T, Nishihara T, Toda H, Yoshida M, Ito H. High pericoronary adipose tissue attenuation on computed tomography angiography predicts cardiovascular events in patients with type 2 diabetes mellitus: post-hoc analysis from a prospective cohort study. Cardiovasc Diabetol. 2022;21:44.
34. Silva GFS, Fagundes TP, Teixeira BC, Chiavegatto Filho ADP. Machine learning for hypertension prediction: a systematic review. Curr Hypertens Rep. 2022;24:523–33.
35. Jiang Z, Yuan F, Zhang Q, Zhu J, Xu M, Hu Y, Hou C, Liu X. Classification of superficial suspected lymph nodes: noninvasive radiomic model based on multiphase contrast-enhanced ultrasound for therapeutic options of lymphadenopathy. Quant Imag Med Surg. 2024;14:1507–25.
36. Xu Y, Goodacre R. On splitting training and validation set: a comparative study of cross-validation, bootstrap and systematic sampling for estimating the generalization performance of supervised learning. J Anal Test. 2018;2:249–62.
37. Ju C, Bibaut A, van der Laan M. The relative performance of ensemble methods with deep convolutional neural networks for image classification. J Appl Stat. 2018;45:2800–18.
38. Zhang L, Wang Z, Zhou Z, Li S, Huang T, Yin H, Lyu J. Developing an ensemble machine learning model for early prediction of sepsis-associated acute kidney injury. iScience. 2022;25:104932.
39. Goeller M, Achenbach S, Cadet S, Kwan AC, Commandeur F, Slomka PJ, Gransar H, Albrecht MH, Tamarappoo BK, Berman DS, et al. Pericoronary adipose tissue computed tomography attenuation and high-risk plaque characteristics in acute coronary syndrome compared with stable coronary artery disease. JAMA Cardiol. 2018;3:858–63.

Publisher's Note