Machine learning-based identification of symptomatic carotid atherosclerotic plaques with dual-energy computed tomography angiography

Ling-Jie Wang, Pei-Qing Zhai, Li-Li Xue, Cai-Yun Shi, Qian Zhang, and Hua Zhang,

Objective: This study aimed to develop and validate a machine learning model incorporating both dual-energy computed tomography (DECT) angiography quantitative parameters and clinically relevant risk factors for the identification of symptomatic carotid plaques to prevent acute cerebrovascular events. *Methods:* The data of 180 patients with carotid atherosclerosis plaques were analysed from January 2017 to December 2021; 110 patients (64.03±9.58 years old, 20 women, 90 men) were allocated to the symptomatic group, and 70 patients (64.70±9.89 years old, 50 women, 20 men) were allocated to the asymptomatic group. Overall, five machine learning models using the XGBoost algorithm, based on different CT and clinical features, were developed in the training cohort. The performances of all five models were assessed in the testing cohort using receiver operating characteristic curves, accuracy, recall rate, and F1 score. *Results:* The shapley additive explanation (SHAP) value ranking showed fat fraction (FF) as the highest among all CT and clinical features and normalised iodine density (NID) as the 10th. The model based on the top 10 features from the SHAP measurement showed optimal performance (area under the curve [AUC] .885, accuracy .833, recall rate .933, F1 score .861), compared with the other four models based on conventional CT features (AUC .588, accuracy .593, recall rate .767, F1 score .676), DECT features (AUC .685, accuracy .648, recall rate .667, F1 score .678), conventional CT and DECT features (AUC .819, accuracy .740, recall rate .867, F1 score .788), and all CT and clinical features (AUC .878, accuracy .833, recall rate .867, F1 score .852). *Conclusion:* FF and NID can serve as useful imaging markers of symptomatic carotid plaques. This tree-based machine learning model incorporating both DECT and clinical features could potentially comprise a non-invasive method for identification of symptomatic carotid plaques to guide clinical treatment strategies.

Keywords: Carotid plaque—Ischaemic stroke—Transient ischaemic attack—Machine learning—Dual-energy computed tomography angiography

Abbreviations: ROC, receiver operating characteristic curve; AUC, area under the receiver operating characteristic curve; DECT, dual-energy computed tomography; CTA, computed tomography angiography; NID, normalised iodine density; MPT, max soft-plaque thickness; FF, fat fraction; EAN, effective atomic number; SHAP, shapley additive explanations

From the Department of Radiology, First Hospital of Shanxi Medical University, Taiyuan, Shanxi Province 030001, PR China.

Received January 28, 2023; revision received May 30, 2023; accepted June 2, 2023.

What this paper adds: Many studies have revealed a strong association between the composition of carotid plaque and acute cerebrovascular events. And much of the recent literature has focused on how to accurately quantitatively assess the composition of carotid vulnerable plaque by noninvasive imaging methods, such as lipid necrotic core, intraplaque hemorrhage, and neovascularization. This study revealed that fat fraction and iodine density of carotid plaque obtained by dual-energy CT angiography could serve as powerful predictors of acute cerebrovascular events. Further, this study developed a machine learning model using the XGBoost algorithm based on dual-energy CT quantitative parameters and clinical risk factors for the identification of symptomatic carotid plaques, exhibiting excellent performance and might become a potential computer-aided tool for disease early therapeutic interventions to prevent cerebrovascular events.

Corresponding author at: First Hospital of Shanxi Medical University, No. 85, Jiefang South Road, Taiyuan, Shanxi, 03001, PR China. E-mails: 4617784@qq.com, 1060791457@qq.com, 1295184747@qq.com, 1696955652@qq.com, 499354932@qq.com, 13623665879@163.com.

1052-3057/$ - see front matter

https://doi.org/10.1016/j.jstrokecerebrovasdis.2023.107209

Introduction

Carotid atherosclerotic plaque is a predominant risk factor of ischaemic stroke, which is the leading cause of death and disability globally.[1,2] Some studies have revealed that ischaemic stroke in carotid atherosclerotic plaques results from hypoperfusion caused by lumen stenosis and embolisation or acute occlusion caused by plaque rupture among others.[3] The carotid-stenosis degree has been used as a determining parameter of stratifying the severity of carotid-artery atherosclerosis for prevention of stroke in patients with carotid plaques by European and US guidelines.[4] Although, symptomatic patients with ≥70% carotid stenosis can benefit from carotid endarterectomy or carotid angioplasty and stenting for ipsilateral stroke prevention,[5] the haemodynamic severity of carotid-artery stenosis is much less likely to be associated with stroke in asymptomatic patients.[3] Clearly, patients with symptomatic carotid plaques have a higher risk of ischaemic stroke than those with asymptomatic carotid plaques, and the risk of ischaemic stroke is highly likely to be caused by plaque vulnerability, independent of the degree of carotid stenosis.[4,6] Therefore, accurate identification of symptomatic carotid plaques is essential to the initial medical therapy for reducing disabling strokes and related deaths in these patients.

It has been demonstrated that both plaque composition and surrounding environment play a key role in plaque vulnerability.[5,7] Currently, conventional computed tomography angiography (CTA) of the neck is the most commonly used intravascular imaging modality for the assessment of carotid atherosclerosis, but it is difficult to discriminate and quantitate the plaque composition using conventional CTA, which can only consider low attenuation and irregular morphology as risk plaque features for stroke.[4,8] Developments in imaging techniques have enabled accurate characterisation of carotid-plaque features.[4] Compared with conventional CT, various quantitative post-processing techniques of dual-energy CT (DECT) can offer several quantitative plaque parameters, such as iodine density (ID), fat fraction (FF), and effective atomic number (EAN), which can objectively reflect plaque composition to facilitate optimised plaque risk assessment.[9,10]

Furthermore, histopathologic studies have revealed that carotid plaque composition is associated with clinical risk factors of cardiovascular disease, such as sex, age, body mass index, and relevant comorbidities.[5,11,12] Combined analysis of quantitative imaging descriptors and clinical characteristics is the most promising approach of identification of symptomatic carotid plaques. Machine learning is considered an important branch of artificial intelligence and has been widely used in the extraction of meaningful patterns from substantial medical data to make disease predictions.[13]

This study aimed to develop and validate a machine learning model that incorporates both quantitative parameters of DECT angiography and clinically relevant risk factors for identification of symptomatic and asymptomatic carotid plaque, and we also explored the quantitative analysis of plaque composition using DECT parameters.

Patients and methods

Patient population

Institutional review board approval (no. 2018K008) was obtained for this study, and the requirement for informed consent was waived because of the retrospective design. Data of patients with carotid atherosclerosis diagnosed using head-and-neck CTA were collected between January 2017 and December 2021 and retrospectively reviewed.

Presence of symptomatic carotid plaques was defined when there were ipsilateral acute ischaemic symptoms in the distribution of the anterior and middle cerebral arteries, including transient ischaemic attack (TIA) and ischaemic stroke. TIA was diagnosed as a brief (<24 h) episode of neurologic dysfunction, including monocular amaurosis fugax, dysarthria, and dysphasia, lacking evidence of acute cerebral infarction on medical imaging.[14]

The inclusion criteria were: (1) plaque thickness >1.5mm[15] and calcification <50%; (2) plaque in the common carotid artery (CCA) or extracranial segment of the internal carotid artery; (3) no multiple stenosis or occlusion of the contralateral anterior and middle cerebral arteries on head-and-neck CTA; (4) sufficient available medical records to determine whether ipsilateral stroke or TIA had occurred in the 2 weeks before the head-and-neck CTA.

The exclusion criteria were: (1) complete occlusion of the extracranial carotid artery; (2) ipsilateral stroke or TIA having occurred more than 2 weeks before the head-and-neck CTA; (3) atrial fibrillation, vascular disease, peripheral vascular disease, carotid web, surface ulcers in the ascending aorta, and/or aortic arch plaque significantly affect aortic arch branches; (4) history of head injury, intracranial artery bypass graft surgery, ipsilateral carotid endarterectomy, or carotid stenting; (5) diseases that can cause lumen stenosis or occlusions, such as vasculitis, moyamoya, fibromuscular dysplasia, and carotid artery dissection; and (6) impaired CTA image quality. The patients were divided into a symptomatic group and an asymptomatic group (Fig. 1).

DECT angiography protocol

All head and neck CTA was performed using a third-generation dual-source scanner (SOMATOM Force; Siemens Healthcare, Erlangen, Germany). The scanning

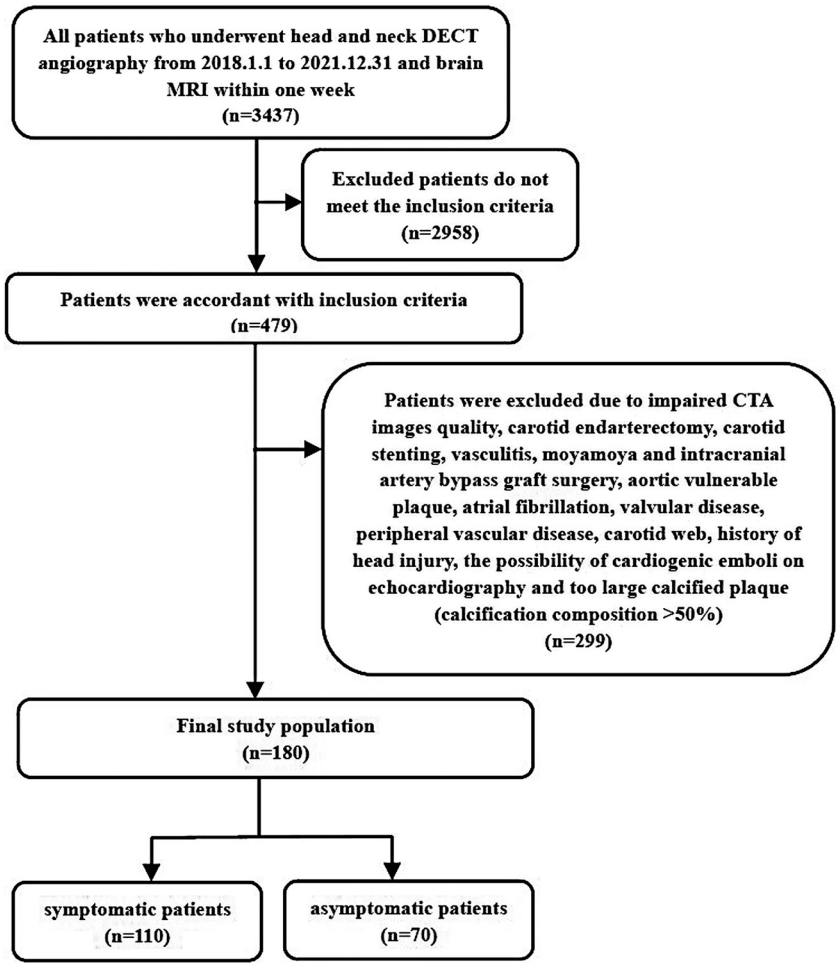

Fig. 1. *Flow diagram of patient recruitment for the final analyses*

conditions were as follows: tube voltage, tube A 100 kV and tube B Sn150 kV; tube current, automatically regulated; slice thickness, 0.75 mm; increment, 0.5 mm; rotation speed, 0.28 s; helical pitch, 0.7; and table speed, 109.7 mm/s. A non-ionic contrast agent (Ultravist 370) at a dose of 1.2 mL/kg (the maximum dose is 120 mL) was injected at a flow rate of 3.5−5.0 mL/s through a peripheral vein in the forearm, followed by injection of a 40−50 mL saline flush at the same flow rate. An automated trigger scan was performed when the attenuation reached 100 Hounsfield units (HU) in a region of interest in the ascending aorta with a 3-s delay. All obtained scan data were transferred to a dedicated workstation (syngo.via, version VB10B; Siemens Healthcare).

Carotid plaque image analysis

All carotid plaques in the CTA images were analysed by two radiologists with >5 years of experience in carotid plaque imaging. The two radiologists were blinded to magnetic resonance imaging (MRI) data and clinical information. We used the North American Symptomatic Carotid Endarterectomy Trial criteria to quantify the degree of carotid stenosis. The presence of a low attenuation plaque rim sign and plaque ulceration have been validated as markers of high-risk plaque features on CTA. Soft plaques were defined as having low density or lipid-rich cores, studies using a specific threshold of HU of <50 or 60.[16] If there is adventitial calcification (<2 mm thick) with internal soft plaque (≥2 mm thickness), the rim sign is defined as positive and otherwise as negative.[17] CT values of carotid plaques were measured by manually drawing a region of interest along the edge of the internal soft plaque while avoiding margins and partial volume effects from adjacent fat. Max soft-plaque thickness (MPT) and max soft-plaque length (MPL) were respectively measured on CTA axial and sagittal sections (Fig. 2A−C).

ID and FF values were measured on a dedicated workstation (syngo.via, version VB10B, Siemens Healthcare) with an iodine subtraction algorithm (Liver VNC, Siemens Healthcare). This software performs image-based analysis of the low- and high-energy keV-peak images and enables spectral analysis of materials. The spectral algorithm generates a map that encodes the iodine distribution in each voxel based on a three-material decomposition algorithm by assuming that every voxel is composed

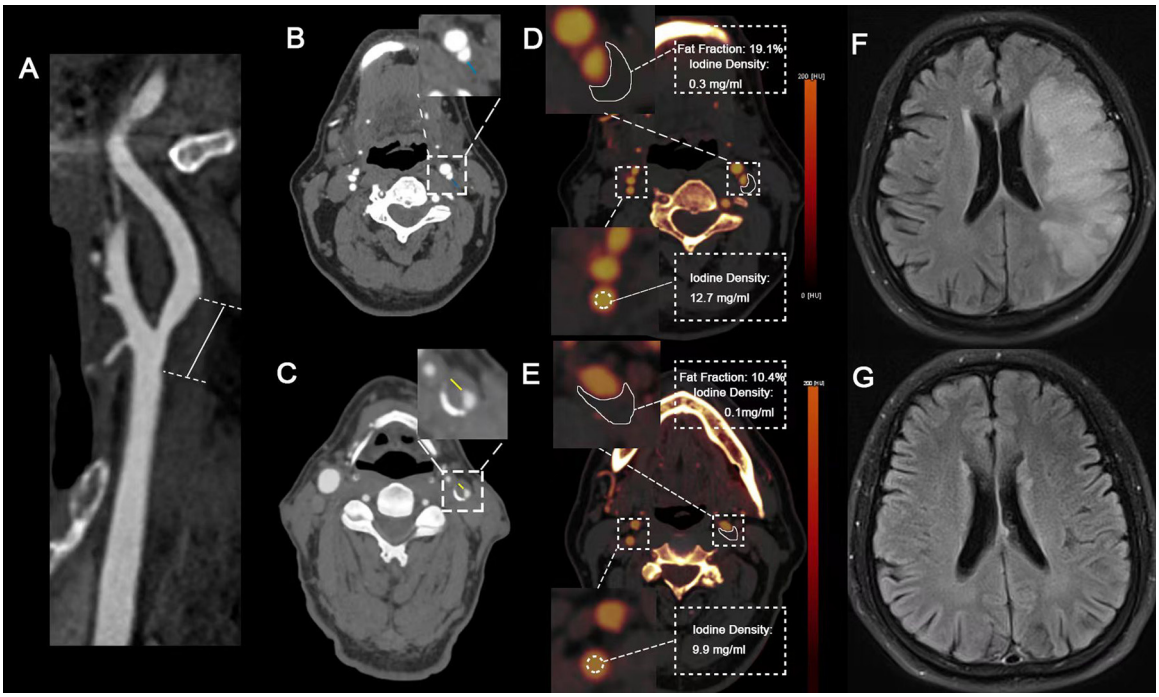

Fig. 2. *Methods of various computed tomography (CT) quantitative parameter assessment. Representative axial CT angiography image of the carotid artery max soft-plaque length (MPL) and max soft-plaque thickness (MPT) in a patient with non-calcified (A, B) or mixed (C) plaques. MPT in patients with non-calcified (blue line) or mixed (yellow line) plaques was measured. Fat fraction (FF) and iodine density (ID) measurement of the carotid plaque, and ID of the contralateral carotid artery at the same level in the symptomatic group (D, F) and symptomatic group (E, G).*

of fat, soft tissue, and iodine. Regions of interest of the ID and FF value measurements were identified for the soft-plaque attenuation values. In the ID measurements of the contralateral carotid lumen at the same imaging level, the region of interest was placed at the centre while avoiding the edges where volume averaging can occur. After the regions of interest were placed, the ID and FF values were generated automatically in the workstation software. All measurements were repeated three times and averaged to limit potential measurement inaccuracies (Fig. 2D, E). The normalised ID (NID) of carotid plaque was calculated as NID = ID of carotid plaque / ID of contralateral carotid artery.

Statistical analysis

All statistical analyses and model development were performed using SPSS25.0 and Python 3.7. The Shapiro–Wilk test was used to examine for the assumption of normal distribution. In the univariate analysis, continuous variables were compared using independent t tests or the Mann–Whitney U test, and categorical variables were compared using the chi-square test. Three or more groups were compared using the Kruskal–Wallis H test. A two-tailed p-value $<.05$ was considered statistically significant.

The intraclass correlation coefficient (ICC) and Cohen k statistics were used to assess the inter-observer and intra-observer agreements of all CT parameters. A good agreement was shown when ICC values $>.75$ and $k>.6$. To assess inter-observer agreement, two radiologists independently analysed 20 randomly selected DECT angiography images of carotid plaques. Moreover, to assess intra-observer agreement, one radiologist re-analysed 20 randomly selected DECT angiography images of carotid plaques 6 months later.

Development and comparison of machine learning models

We used a boosted ensemble algorithm (XGBoost) as previously described for identification of symptomatic carotid atherosclerotic plaques, which is insensitive to missing values and particularly useful for obtaining a more reliable predictive model.[18,19]

Although an XGBoost model has better prediction accuracy than a linear model, it is missing the interpretability of the linear model. Therefore, the XGBoost model is usually considered a black-box model. To address this problem, the shapley additive explanations (SHAP) method was used for interpreting predictions.[20] Here, SHAP values were used to explain the effects of all variable attributions for symptomatic plaques in each patient. The top 10 most relative features from the SHAP measurement were used in the XGBoost model for identification of symptomatic plaques.

The data were randomly divided into a training cohort and a testing cohort at a ratio of 7: 3. We assessed the predictive ability of machine learning models based on the

following factors: conventional CT features, DECT features (including FF, NID, and EAN values), conventional CT and DECT features, the top 10 of all CT and clinical features according to the SHAP values, and all CT and clinical features.

Validation of the models was performed using the testing cohort. Receiver operating characteristic (ROC) analysis, accuracy, recall rate, and the F1 score were used to evaluate model performance. Differences in the area under the ROC curve (AUC) between the various models were compared using the Delong test.

Results

Clinical and CT data of the patients

In total, 180 patients were finally enrolled among 3437 patients who underwent head-and-neck CTA (Fig. 1); 110 patients were allocated to the symptomatic group (64.03± 9.58 years old, 20 women, 90 men), and 70 were allocated to the asymptomatic group (64.70±9.89 years old, 50 women, 20 men). All clinical and CT data of the patients are summarised in Table 1. The stroke classification of patients based on National Institute of Health Stroke Scale

Table 1. *Features of Patients in the Training and Testing Cohorts.*

Features	symptomatic Group (n = 110)	asymptomatic Group (n =70)	P value
Clinical Features			
Demographic			
Gender, Male	90 (81.8%)	20(80.0%)	0.761
Age, y	64.03 ± 9.58	64.70 ± 9.89	0.651
BMI#, kg/m^2	24.56 ± 2.94	24.47 ± 3.08	0.845
Somking #	50 (47.2%)	30 (46.2%)	0.897
Drinking#	23(21.7%)	14 (21.5%)	0.980
Hypertension#	79 (75.2%)	40 (61.5%)	0.058
Diabetes#	37 (35.2%)	32 (49.2%)	0.071
Initial laboratory data			
Glycosylated hemoglobin#, %	0.06 ± 0.02	0.06 ± 0.02	0.958
Homocysteine#, μmol/L	21.84 ± 15.62	17.13 ± 0.94	0.002*
TG#, mmol/L	1.48 ± 0.69	2.46 ± 6.74	0.322
TC#, mmol/L	4.15 ± 1.32	4.16 ± 1.14	0.685
HDL-C#, mmol/L	1.05±0.24	1.06 ± 0.24	0.861
LDL-C#, mmol/L	2.61 ± 1.00	2.58 ± 0.84	0.826
Systolic Pressure#, mmHg	148.88 ± 23.21	140.08 ± 19.99	0.016*
Diastolic Pressure#, mmHg	86.49 ± 15.28	83.10 ± 13.50	0.159
Conventional CT Features			
Plaque location			0.001*
CCA	24 (21.8%)	34 (48.6%)	
Bifurcation of CCA	24 (21.8%)	12 (17.1%)	
ICA	62 (56.4%)	14 (34.3%)	
Lumen stenosis			0.003*
<=50%	47 (42.7%)	46 (65.7%)	
<50% and <=70%	47 (42.7%)	22(31.4%)	
>70%	16 (14.5%)	2 (2.9%)	
Plaque ulceration	26 (23.6%)	8 (11.4%)	0.761
Rim sign	70 (63.6%)	40 (36.4%)	0.006*
CT Value, HU	24.15 ± 8.93	30.07 ± 10.60	0.000*
Max soft-plaque thickness, mm	3.59 ± 0.89	3.14 ± 0.90	0.001*
Max soft-plaque length, mm	21.74±8.10	22.70 ± 7.95	0.380
Minimum luminal area, mm^2	12.85 ± 9.50	17.16 ± 8.77	0.001*
DECT Features			
Fat Fraction, %	28.81 ± 7.86	23.34 ± 7.44	0.000*
Normalized iodine density, 10^{-2}	0.05 ± 0.03	0.03 ± 0.02	0.001*
Effective atomic number	7.33 ± 0.26	7.40 ± 0.27	0.085

NOTE. Continuous variables are presented as mean ± standard deviation and categorical variables are presented as numbers with percentages in parentheses. BMI, body mass index. TG, triglyceride. TC, total cholesterol. HDL-C, high-density lipoprotein cholesterol. LDL-C, low-density lipoprotein cholesterol. CCA, common carotid artery. ICA, internal carotid artery. # There are missing values. * P value <0.05.

(NIHSS) is presented in the data supplement (Supplementary Table 3).

Inter-observer and intra-observer agreement

Inter-observer and intra-observer agreements of all CT parameters are shown in the data supplement (Supplementary Table S1). The inter-observer and intra-observer ICCs for MPT, MPL, minimum lumen area (MLA), lumen stenosis, CT value, FF, NID, and EAN were all >.8, showing good agreement. The respective inter-observer and intra-observer agreements were good for plaque ulceration (k=.857 and .773) and for rim sign (k=.773 and .828).

Comparison of population distribution and DECT quantitative parameters in different plaque locations and lumen stenosis

In this study cohort, there were more symptomatic patients with <70% carotid stenosis than with ≥70% stenosis (Supplementary Figure 1A) and symptomatic plaques were more often located in internal carotid artery (ICA) (Supplementary Figure 1B).

DECT quantitative parameters (FF, NID, and EAN) were further analysed (Supplementary Figure 2A−E). The EAN values were not significantly different between the symptomatic and asymptomatic patients, while carotid plaques had significantly higher FF and NID values in the symptomatic patients than in the asymptomatic patients. There were no statistically significant differences in FF, NID, and EAN values in symptomatic patients at different locations (CCA, carotid bifurcation, and ICA). The FF of plaques in symptomatic patients with 50%−69% carotid stenosis was higher than that in patients with <50% and ≥70% carotid stenosis, while the NID and EAN values were not significantly different in symptomatic patients with different degrees of carotid stenosis.

Performance of the machine learning models for the identification of symptomatic carotid plaques

The importance ranking of global variables obtained with the SHAP method is exhibited in Fig. 3. In the top 10 variables based on the SHAP values, FF was the highest ranked, followed by homocysteine, lumen stenosis, systolic blood pressure, MLA, MPL, triglycerides, MPT, glycosylated haemoglobin, and NID.

In the testing cohort (Fig. 4A), the conventional CT feature model, DECT feature model, conventional CT and DECT feature model, the top 10 of all CT and clinical feature model, and all CT and clinical feature model yielded AUCs of .588 (95% CI 0.43−0.94), .685 (95% CI 0.54−0.83), 0.819 (95% CI 0.71−0.93), .885 (95% CI 0.78−0.99), and .878 (95% CI 0.77−0.98), respectively. The Delong test results are shown in Supplementary Table 2. The accuracy values were .593, .648, .740, .833, and .833, respectively. The recall rates were .767, .667, .867, .933, and .867, respectively. The F1 scores were .676, .678, .788, .861, and .852, respectively. The bar chart in Fig. 4B shows that the model based on the top 10 of all CT and clinical features had the highest performance.

The calibration curve (Fig. 5A) of the five models for symptomatic carotid plaque identification demonstrated that the model based on the top 10 of all CT and clinical features has the best agreement. Further, as shown in Fig. 5B, the decision curves indicated that the model based on the top 10 of all CT and clinical features could add more benefit to the patients.

Discussion

Here, we developed a comprehensive machine learning model based on clinical risk factors and carotid DECT angiography quantitative parameters to identify carotid symptomatic plaques for prevention of cerebrovascular events. Our model exhibited excellent performance for the identification of symptomatic plaques. Global variable analysis using the SHAP method revealed that quantitative DECT parameters, FF and NID, representing plaque composition were independent predictors of cerebrovascular symptoms and could serve as useful imaging markers of symptomatic carotid plaques. Our findings may provide a promising means of computer-aided analysis of carotid plaques based on DECT angiography for early therapeutic interventions.

Clearly, the cerebrovascular symptoms of patients with carotid atherosclerosis are significantly associated with the degree of carotid stenosis.[3] However, in actual clinical practice, cerebrovascular events are seen more often in patients with <70% stenosis, even <50% stenosis. This study cohort also showed that carotid stenosis is mostly <70% in symptomatic patients. Attributable to the anatomical characteristics of the ICA directly entering the skull, hypoperfusion and microemboli caused by ICA plaques are more likely to cause cerebrovascular symptoms than those by CCA and carotid bifurcation plaques.

Histopathologic studies have revealed that features of carotid plaques are closely correlated with cerebrovascular events, such as large lipid cores, intraplaque haemorrhage (IPH), intraplaque neovascularisation (IPN), adventitial inflammation, irregular/ulcerated surface, and rim sign.[5] Although it is difficult to accurately assess some of these features using conventional CTA compared to high resolution MRI (HR-MRI), it has been demonstrated that the HU scale used with conventional CT has the potential to quantify lipid necrotic cores of plaques, and the CTA rim sign is highly associated with IPH.[17,21,22] Moreover, owing to the high cost, long time required, and numerous contraindications of HR-MRI, CTA has become a commonly use clinical assessment method of carotid diseases. This study showed that the CT value and rim sign of carotid plaques were statistically different between symptomatic and asymptomatic patients, but their lower

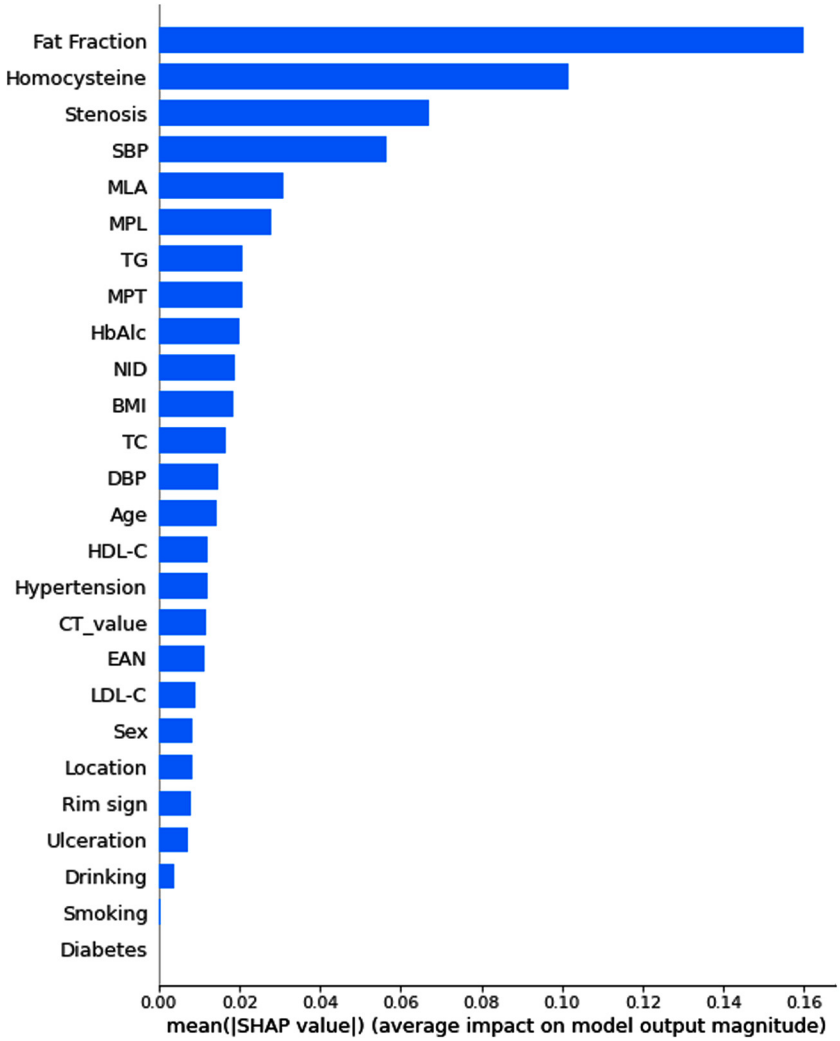

Fig. 3. *Shapley additive explanation (SHAP) value ranking of global variables.*

SHAP values reflect lower importance in the predictive model. These results revealed the limitations of conventional CTA for accurate assessment of intraplaque composition.

Currently, DECT has been extensively applied in clinical practice, and the measurement of ID and FF using the three-material decomposition technique of DECT is available to distinguish and evaluate different diseases.[23] Furthermore, there is evidence that DECT can accurately quantify lipid necrotic cores of plaques and IPN, which appear to be proportional to FF and ID, respectively.[24,25] It has been documented that large necrotic lipid cores and IPN play a key role in unstable plaque rupture,[1,26,27] and symptomatic plaques have a higher lipid content.[3] According to our results, the FF and NID of carotid plaques significantly differed between symptomatic and asymptomatic patients, and their higher SHAP values indicate they were very important variables in the predictive model; especially, the SHAP value of FF was the highest among all features. Hence, FF and NID can serve as useful imaging markers of necrotic lipid cores and IPN, respectively. EAN can reflect the atomic number of a material,[28] and the usefulness of EAN for quantification of carotid plaque composition was reported.[29] However, the EAN of carotid plaques may reflect the EAN mean of many materials because of the complex plaque composition, which could reduce the accuracy of quantitative analysis. This study also showed that EAN values were not significantly different between symptomatic and asymptomatic patients.

As shown by many previous studies, cardiovascular risk factors in clinical data are related to atherosclerotic plaque formation and progression. Therefore, after we incorporated these clinical risk factors into the model, the performance was improved.

The boosted ensemble algorithm, XGBoost, was first proposed by Chen and Guestrin.[30] As there were a few missing values in the clinical data of this study, XGBoost was used to develop the predictive model because it could handle sparse data better to develop a more reliable

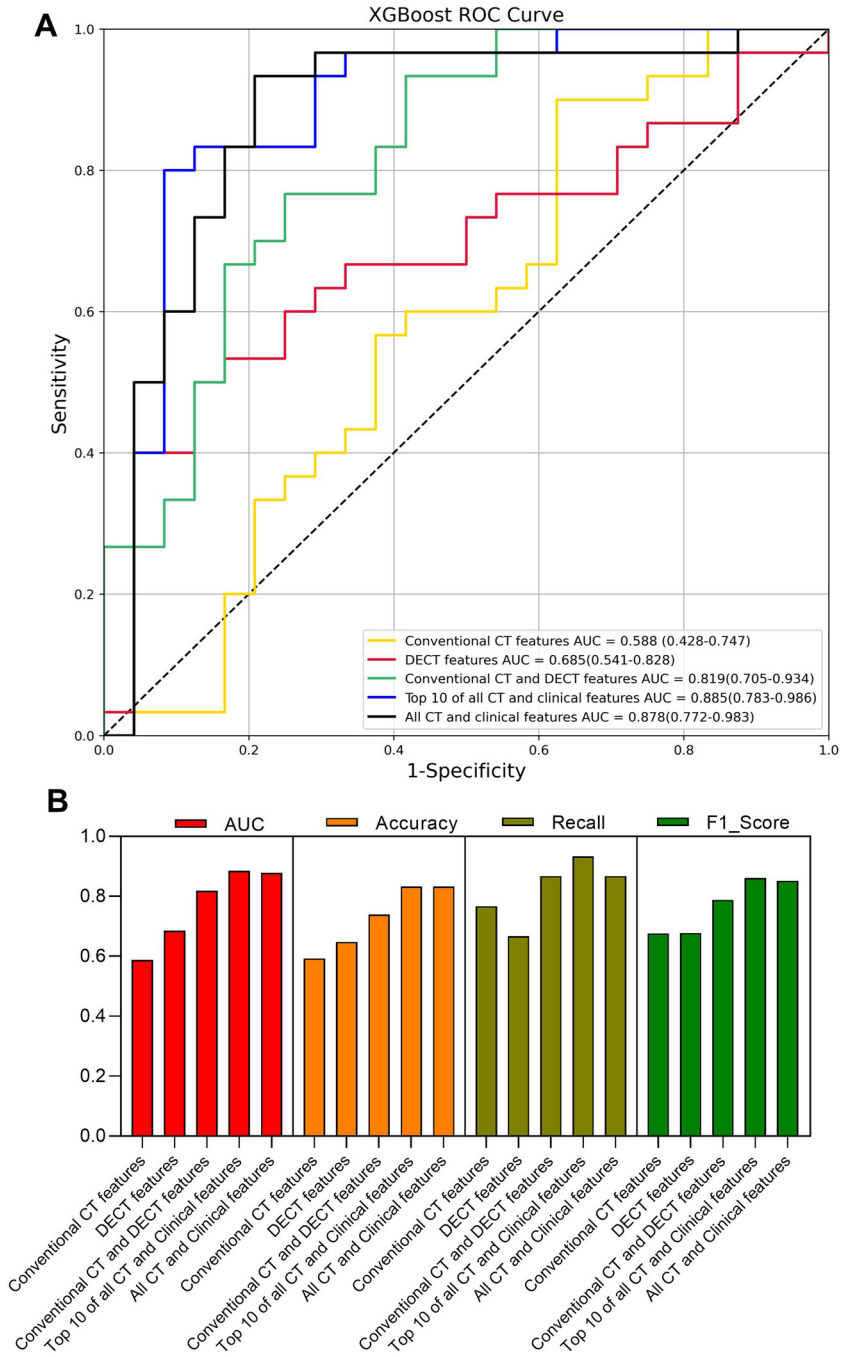

Fig. 4. *Receiver operating characteristic (ROC) curves of the five machine learning models based on different features in the testing cohort (A). Bar charts of the area under the ROC curve (AUC) values, accuracy, recall rate, and F1 score of the five models in the testing cohort (B).*

predictive model using a minimal amount of resources compared traditional logistic regression.[18,30] Although tree-based machine learning models, such as XGBoost, are widely used in the medical domain because of their excellent performance, their poor interpretability makes it impossible to know how the independent variables of the sample affect the final output. The SHAP method can accurately compute the contribution value (SHAP value) of each variable across the entire dataset to improve the interpretability of tree-based machine learning models.[31] Feature importance measurement and selection were performed using the SHAP method in this study, and the model based on the top 10 features according to the SHAP values had the best evaluation metrics (AUC, accuracy, recall rate, and F1 score), despite the lack of significant differences in the AUC values of the latter three models per the Delong test.

The major limitation of this study was its retrospective single-centre design. Therefore, a prospective multicentre study should be performed to estimate the validity of the

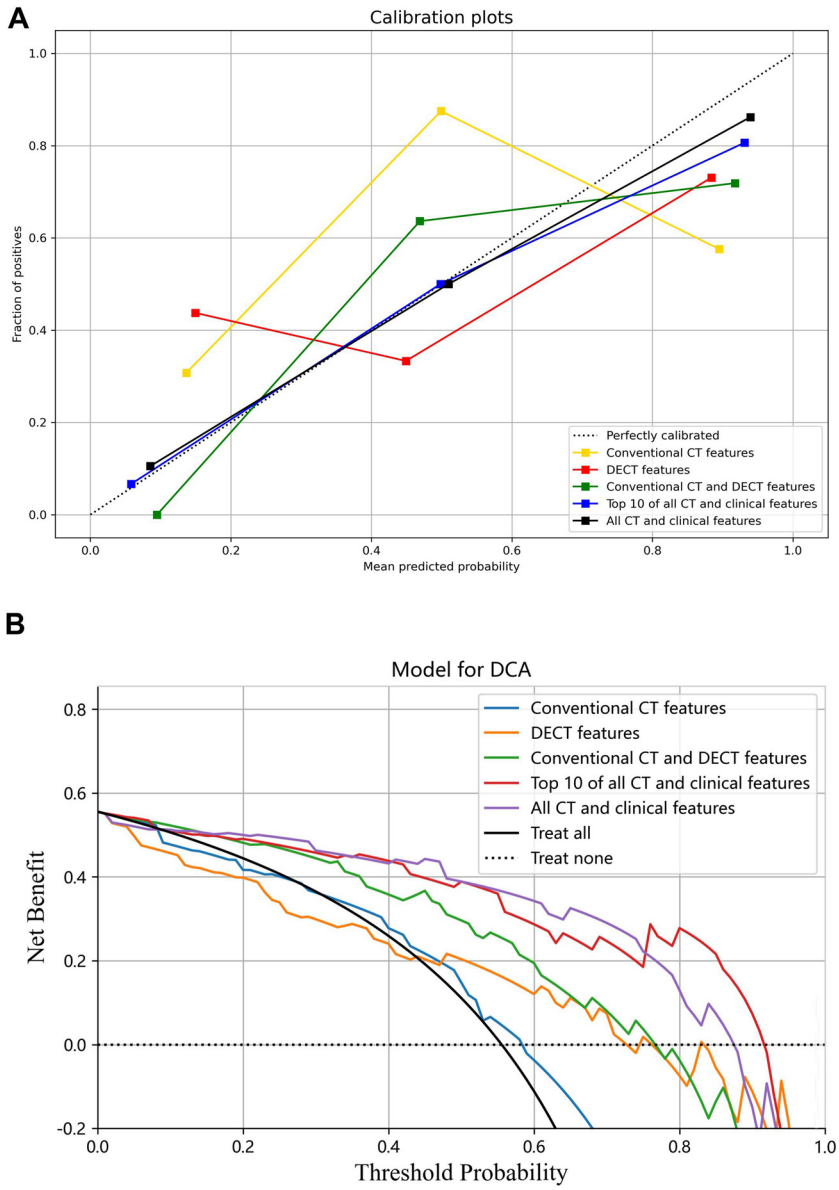

Fig. 5. *Calibration curves (A) and decision curve analysis (B) of the five machine learning models based on different features in the testing cohort. Calibration curves depict the calibration of each model in terms of the agreement between the predicted and observed outcomes of symptomatic carotid plaques. The y-axis represents the actual positive percentage and the x-axis represents the predicted probability. The diagonal dotted line represents a perfect prediction by an ideal model. Decision curves depict the relationship between the predicted probability threshold of symptomatic carotid plaques and maximum net benefit for patients. The y-axis represents the net benefit, and the x-axis represents the predicted probability threshold. The horizontal dotted line represents the assumption that no patients had symptomatic carotid plaques. The black line represents the assumption that all patients had symptomatic carotid plaques.*

predictive model, which is an area of planned future study.

Briefly, this study developed a tree-based machine learning model using XGBoost combining both DECT quantitative parameters and clinical risk factors for the identification of symptomatic carotid plaques, showing good predictive efficacy and might become a simple and useful computer-aided tool for early therapeutic interventions to prevent cerebrovascular events. The interpretability of the tree-based machine learning model was improved using the SHAP method. The SHAP value ranking of global variables revealed that the FF and NID of Energy spectrum parameters of DECT could serve as important imaging markers of symptomatic carotid plaques, which might reflect more accurately the lipid necrotic core and IPN of carotid plaque histopathological composition compared to the conventional CT parameters.

Supplementary Figure 1. Population distribution of symptomatic group and asymptomatic group in different carotid lumen stenosis degrees (A) and plaque locations (B).

Supplementary Figure 2. DECT quantitative parameters (FF, NID, and EAN) comparison of symptomatic

group and asymptomatic group in different carotid lumen stenosis degrees (A,C,E) and plaque locations (B, D, F).

Declaration of Competing Interest

The authors of this manuscript declare no relationships with any companies, whose products or services may be related to the subject matter of the article.

Acknowledgements: This work was financially supported by the Fundamental Research Program of Shanxi Province and the Key Research and Development (R&D) Project of Shanxi Province (grant numbers 20210302123253, 20210302123256 and 201803D31004).

Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.jstrokecerebrovasdis.2023.107209.

References

1. Zamani M, Skagen K, Scott H, Lindberg B, Russell D, Skjelland; M. Carotid Plaque Neovascularization Detected With Superb Microvascular Imaging Ultrasound Without Using Contrast Media. Stroke 2019;50(11):3121-3127.
2. Huang Z, Cheng XQ, Liu HY, Bi XJ, Liu YN, Lv WZ, Xiong L, Deng; YB. Relation of Carotid Plaque Features Detected with Ultrasonography-Based Radiomics to Clinical Symptoms. Transl Stroke Res 2021.
3. Golledge J, Greenhalgh RM, Davies; AH. The symptomatic carotid plaque. Stroke 2000;31(3):774-781.
4. Saba L, Saam T, Jager HR, Yuan C, Hatsukami TS, Saloner D, Wasserman BA, Bonati LH, Wintermark; M. Imaging biomarkers of vulnerable carotid plaques for stroke risk prediction and their potential clinical implications. Lancet Neurol 2019;18(6):559-572.
5. Howard DP, van Lammeren GW, Rothwell PM, Redgrave JN, Moll FL, de Vries JP, de Kleijn DP, den Ruijter HM, de Borst GJ, Pasterkamp; G. Symptomatic carotid atherosclerotic disease: correlations between plaque composition and ipsilateral stroke risk. Stroke 2015;46(1):182-189.
6. Yang Q, Guo H, Shi X, Xu X, Zha M, Cai H, Yang D, Huang F, Zhang X, Lv Q, Liu R, Liu; X. Identification of Symptomatic Carotid Artery Plaque: a Three-Item Scale Combined Angiography With Optical Coherence Tomography. Front Neurosci 2021;15:792437.
7. Saba L, Agarwal N, Cau R, Gerosa C, Sanfilippo R, Porcu M, Montisci R, Cerrone G, Qi Y, Balestrieri A, Lucatelli P, Politi C, Faa G, Suri; JS. Review of imaging biomarkers for the vulnerable carotid plaque. JVS Vasc Sci 2021;2:149-158.
8. Naim C, Douziech M, Therasse E, Robillard P, Giroux MF, Arsenault F, Cloutier G, Soulez; G. Vulnerable atherosclerotic carotid plaque evaluation by ultrasound, computed tomography angiography, and magnetic resonance imaging: an overview. Can Assoc Radiol J 2014;65(3):275-286.
9. Grnberg F, Lundberg J, Sjlin M, Persson M, Danielsson; M. Feasibility of unconstrained three-material decomposition: imaging an excised human heart using a prototype silicon photon-counting CT detector. Eur Radiol 2020(9).
10. Zainon R, Ronaldson JP, Janmale T, Scott NJ, Buckenham TM, Butler A, Butler PH, Doesburg RM, Gieseg SP, Roake; JA. Spectral CT of carotid atherosclerotic plaque: comparison with histology. Eur Radiol 2012;22(12):2581-2588.
11. Rashid, H.M.; H. Hassan; M. Khan; J. Khan; H. Khan; Small Dense Low-Density Lipoprotein as Risk Factor for Atherosclerosis in Type 2 Diabetes Mellitus. 2020.
12. van Dam-Nolen DHK, van Dijk AC, Crombag GAJC, Lucci C, Kooi ME, Hendrikse J, Nederkoorn PJ, Daemen MJAP, van der Steen AFW, Koudstaal PJ, Kronenberg F, Roeters van Lennep JE, Mulder MT, van der Lugt; A. Lipoprotein(a) levels and atherosclerotic plaque characteristics in the carotid artery: The Plaque at RISK (PARISK) study. Atherosclerosis 2021;329:22-29.
13. Erickson BJ, Korfiatis P, Akkus Z, Kline; TL. Machine Learning for Medical Imaging. Radiographics 2017;37(2):160130.
14. Saba L, Zucca S, Gupta A, Micheletti G, Suri JS, Balestrieri A, Porcu M, Crivelli P, Lanzino G, Qi Y. Perivascular Fat Density and Contrast Plaque Enhancement: Does a Correlation Exist? Am J Neuroradiol 2020;41(8).
15. Saba L, Lanzino G, Lucatelli P, Lavra F, Sanfilippo R, Montisci R, Suri JS, Yuan; C. Carotid Plaque CTA Analysis in Symptomatic Subjects with Bilateral Intraparenchymal Hemorrhage: A Preliminary Analysis. AJNR Am J Neuroradiol 2019;40(9):1538-1545.
16. Baradaran H, Al-Dasuqi K, Knight-Greenfield A, Giambrone A, Delgado D, Ebani EJ, Kamel H, Gupta; A. Association between Carotid Plaque Features on CTA and Cerebrovascular Ischemia: a Systematic Review and Meta-Analysis. Ajnr Am J Neuroradiol 2017. ajnr;ajnr.A5436v1.
17. Eisenmenger LB, Aldred BW, Kim SE, Stoddard GJ, de Havenon A, Treiman GS, Parker DL, McNally; JS. Prediction of Carotid Intraplaque Hemorrhage Using Adventitial Calcification and Plaque Thickness on CTA. AJNR Am J Neuroradiol 2016;37(8):1496-1503.
18. Wang Y, Chen H, Sun T, Li A, Wang S, Zhang J, Li S, Zhang Z, Zhu D, Wang X, Cao; F. Risk predicting for acute coronary syndrome based on machine learning model with kinetic plaque features from serial coronary computed tomography angiography. Eur Heart J Cardiovasc Imaging 2022;23(6):800-810.
19. Alim M, Ye GH, Guan P, Huang DS, Zhou BS, Wu; W. Comparison of ARIMA model and XGBoost model for prediction of human brucellosis in mainland China: a time-series study. BMJ Open 2020;10(12):e039676.
20. Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. Advances in neural information processing systems; 2017. p. 30.
21. Ito T, Terashima M, Kaneda H, Nasu K, Matsuo H, Ehara M, Kinoshita Y, Kimura M, Tanaka N, Habara; M. Comparison of in vivo assessment of vulnerable plaque by 64-slice multislice computed tomography versus optical coherence tomography. Am J Cardiol 2011;107(9):1270-1277.
22. 3rd Sheahan M, Ma X, Paik D, Obuchowski NA, St Pierre S, Newman WP, Rae G, Perlman ES, Rosol M, Keith Jr. JC, Buckler AJ. Atherosclerotic Plaque Tissue: noninvasive Quantitative Assessment of Characteristics with Software-aided Measurements from Conventional CT Angiography. Radiology 2018;286(2):622-631.
23. Martin SS, Weidinger S, Czwikla R, Kaltenbach B, Albrecht MH, Lenga L, Vogl TJ, Wichmann; JL. Iodine and fat quantification for differentiation of adrenal gland

adenomas from metastases using third-generation dual-source dual-energy computed tomography. Investig Radiol 2017:1.
24. Saba L, Lai ML, Montisci R, Tamponi E, Sanfilippo R, Faa G, Piga; M. Association between carotid plaque enhancement shown by multidetector CT angiography and histologically validated microvessel density. Eur Radiol 2012;22(10):2237-2245.
25. Demeure F, Bouzin C, Roelants V, Bol A, Verhelst R, Astarci P, Gerber BL, Pouleur; AC, Pasquet A, De MC. Head-to-Head Comparison of Inflammation and Neovascularization in Human Carotid Plaques: implications for the Imaging of Vulnerable Plaques. Circul Cardiovasc Imaging 2017;10(5):e005846.
26. Toutouzas K, Benetos G, Karanasos A, Chatzizisis YS, Giannopoulos AA, Tousoulis; D. Vulnerable plaque imaging: updates on new pathobiological mechanisms. Eur Heart J 2015;36(45):3147-3154.
27. Falk E, Nakano M, Bentzon JF, Finn AV, Virmani; R. Update on acute coronary syndromes: the pathologists' view. Eur Heart J 2013;34(10):719-728.
28. Nakajima S, Ito H, Mitsuhashi T, Kubo Y, Matsui K, Tanaka I, Fukui R, Omori H, Nakaoka T, Sakura H, Ueno E, Machida; H. Clinical application of effective atomic number for classifying non-calcified coronary plaques by dual-energy computed tomography. Atherosclerosis 2017;261:138-143.
29. Shinohara Y, Sakamoto M, Kuya K, Kishimoto J, Iwata N, Ohta Y, Fujii S, Watanabe T, Ogawa; T. Assessment of carotid plaque composition using fast-kV switching dual-energy CT with gemstone detector: comparison with extracorporeal and virtual histology-intravascular ultrasound. Neuroradiology 2015;57(9):889-895.
30. Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. ACM; 2016.
31. Lundberg SM, Erion G, Chen H, DeGrave A, Prutkin JM, Nair B, Katz R, Himmelfarb J, Bansal N, Lee SI. From local explanations to global understanding with explainable AI for Trees. Nat Mach Intell 2020;2(1):56-67.