Ultrasound in Medicine & Biology 49 (2023) 2437–2445

Contents lists available at ScienceDirect

Ultrasound in Medicine & Biology

journal homepage: www.elsevier.com/locate/ultrasmedbio

Review

Radiomics in Carotid Plaque: A Systematic Review and Radiomics Quality Score Assessment

Chao Hou [a,b], Xin-Yao Liu [b], Yue Du [b], Ling-Gang Cheng [b], Lu-Ping Liu [a], Fang Nie [a], Wei Zhang [b,1], Wen He [a,b,*,1]

[a] Department of Ultrasound, Lanzhou University Second Hospital, Lanzhou, Gansu Province, China
[b] Department of Ultrasound, Beijing Tiantan Hospital, Capital Medical University, Beijing, China

ARTICLE INFO

Keywords:
Radiomics
Carotid artery
Plaque
Systematic review
Radiomics quality score

Imaging modalities provide information on plaque morphology and vulnerability; however, they are operator dependent and miss a great deal of microscopic information. Recently, many radiomics models for carotid plaque that identify unstable plaques and predict cardiovascular outcomes have been proposed. This systematic review was aimed at assessing whether radiomics is a reliable and reproducible method for the clinical prediction of carotid plaque. A systematic search was conducted to identify studies published in PubMed and Cochrane library from January 1, 2001, to September 30, 2022. Both retrospective and prospective studies that developed and/or validated machine learning models based on radiomics data to classify or predict carotid plaques were included. The general characteristics of each included study were selected, and the methodological quality of radiomics reports and risk of bias were evaluated using the radiomics quality score (RQS) tool and Quality Assessment of Diagnostic Accuracy Studies-2, respectively. Two investigators independently reviewed each study, and the consensus data were used for analysis. A total of 2429 patients from 16 studies were included. The mean area under the curve of radiomics models for diagnostic or predictive performance of the included studies was 0.88 ± 0.02, with a range of 0.741–0.989. The mean RQS was 9.25 (standard deviation: 6.04), representing 25.7% of the possible maximum value of 36, whereas the lowest point was −2, and the highest score was 22. Radiomics models have revealed additional information on patients with carotid plaque, but with respect to methodological quality, radiomics reports are still in their infancy, and many hurdles need to be overcome.

Introduction

Carotid atherosclerosis (CAS) is a chronic and progressive disease characterized by focal fibrosis, lipid accumulation and plaque formation, which account for 7%–18% of ischemic strokes [1] and affect the health of about one in four people during their lifetime [2]. Traditionally, risk stratification and therapeutic management of CAS are based mainly on the severity of luminal stenosis [3]. However, growing evidence indicates that the detection of carotid plaque composition by imaging protocols can provide more reliable risk prediction than calculation of stenosis degree alone [4,5]. Therefore, non-invasive identification of patients with high-risk carotid plaques is very important in determining the severity of CAS and risk stratification of subsequent vascular events.

Although imaging technologies such as magnetic resonance imaging (MRI), computed tomography angiography (CTA) and ultrasonography (US) provide a wealth of information on vascular lesions, lumen stenosis, plaque morphological changes and components with high sensitivity and specificity [4,6,7], they are operator dependent and miss a great deal of microscopic information (such as macrophage subsets, collagen type and protein phenotype expressed in plaques), which cannot be recognized by the naked eye. Radiomics, a hot topic first proposed by Lambin et al. in 2012 [8] and an emerging field that extracts high-throughput feature information from medical images and explores their correlation with clinical events non-invasively, has shown considerable potential for diagnostic and differential diagnosis and prognostic prediction in the field of oncology [9]. Indeed, imaging-based radiomics models have proven to be a valuable tool in the cardiovascular system [10]. However, to date, cardiovascular radiomics has lagged behind oncology.

Recently, several studies have attempted to assess the prediction performance and reporting quality of radiomics models based on the radiomics quality score (RQS) tool [11–13]. However, there are no available records of the use of the RQS in carotid plaque to date. Therefore, this review aimed to assess whether radiomics is a reliable and reproducible method for the clinical management of carotid plaque by identifying the role of radiomics models and evaluating the methodological quality of current radiomics studies in carotid plaque.

* Corresponding author. Department of Ultrasound, Lanzhou University Second Hospital, Lanzhou Gansu, 730030, China
E-mail address: hewen@bjtth.org (W. He).
[1] Wei Zhang and Wen He contributed equally to this work.

https://doi.org/10.1016/j.ultrasmedbio.2023.06.008
Received 27 March 2023; Revised 9 May 2023; Accepted 8 June 2023

Methods

Search strategy

This review was performed according to Preferred Reporting Items for Systematic Review and Meta-Analyses (PRISMA) guidelines [14]. We systematically searched PubMed and Cochrane library for articles published in English from January 1, 2001, to September 30, 2022. We used the following key terms: "radiomics/radiomic/texture/textural," "carotid artery/carotid" and "atherosclerosis/plaque/stenosis." The search details in each database are outlined in Table S1 (online only). The protocol for this systematic review was registered with the International Prospective Register of Systematic Reviews (PROSPERO, Registration No. CRD42023407441).

Study selection

Both prospective and retrospective studies satisfying the following criteria were considered eligible: (i) participants: patients with carotid artery plaque; (ii) intervention and control: development and/or validation diagnostic or prognostic model based on radiomics or texture analysis; (iii) outcomes: model performance, including area under the receiver operating characteristic curve (AUC) of both training and testing groups, sensitivity, specificity, positive or negative predictive values, cutoff values, validation and calibration of models; (iv) study design: peer-reviewed scientific reports published in English until the search date. Exclusion criteria were as follows: (i) non-human participants (animals or modeling data generated algorithmically); (ii) deep-learning research without any texture feature in the model, assessment of the predictive value of a single feature without any prediction model; (iii) study type such as letters, reviews, case reports, abstracts, editorial or other informal publication types. Two reviewers (X.Y.L. and Y.D.) independently reviewed the title and abstract for initial selection, then performed a full-text assessment of each article to decide the final studies to be included; any controversy was resolved by consensus with a senior reviewer (L.G.C.).

Data extraction

For each study, the following data were extracted from the full text: first author, publication year, country, study type, number of patients, gender, mean age, prevalence history, imaging modality, modeling method, number of features in the optimal model, main objective and conclusion and predictive power of the study. If there were several prediction models in a study, we selected the one with the best performance in the test/training cohort.

Quality assessment of each study

We applied the RQS tool to evaluate the methodological quality of the included studies. The RQS comprises six domains of 16 different items, with the score in each part ranging from −5 to 7; the total score is 36 (100%) [15]. Two reviewers (L.G.C. and L.P.L.) completed the RQS assessment, and any disagreement was resolved by consensus.

The Prediction Model Risk of Bias Assessment Tool (PROBAST) [16] and the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) [17] are tools used to assess the risk of bias and the applicability of studies. Both tools consist of four domains, the former including participants, predictors, outcomes and analysis, and the latter including patient selection, index test, reference standard and flow and timing. However, PROBAST deals primarily with regression-based clinical predictive models rather than radiomics models, as reported in some reviews; the results of RQS and PROBAST were not parallel [13]. As the radiomics model is a diagnostic tool, we assessed the risk of bias and applicability concerns with QUADAS-2. QUADAS-2 contains 14 standards with an answer as "yes/no/unclear" or "high concern/low concern/unclear concern." Two reviewers (X.Y.L. and Y.D.) completed the QUADAS-2 assessment, and any disagreement was resolved by consensus.

Statistical analysis

The percentage of articles according to the score category of each of the six domains of the RQS was extracted, the total RQS score was calculated (score range, −8 to 36) and the basic adherence rate (range: 2.7% −100%) was assigned when a score of at least one point was obtained. Measurement data were expressed as the mean ± standard deviation, and count data were expressed as a percentage or rate. All statistical analyses were performed using SPSS version 25 (IBM, Armonk, NY, USA). Risk of bias and applicability concerns were generated by Cochran RevMan, version 5.3 (Informatics and Knowledge Management Department, Cochrane, London, UK).

Results

Literature selection

A total of 302 studies were identified using the aforementioned search stratagem. After removal of 133 duplicates, 15 inappropriate types of publications, 5 non-English articles and 3 animal studies, 87 studies irrelevant to the review objective were excluded through more detailed evaluation, and finally, 16 studies that met the inclusion criteria were included in this review after a full-text screen [18–33]. Figure 1 is a flowchart of the study.

General characteristics

The 16 studies included were single-center research studies published between July 2014 and June 2022, with 68.7% (11/16) published within the last 2 y; 81.2% (13/16) of studies were retrospectively designed. Most of the studies were performed in the Chinese population (50%, 8/16), followed by Italy and England (12.5%, 2/16). A total of 2429 patients (men: 1612) were included, with a population size varying from 21 to 548 (median: 135) and a mean age of 65.73 ± 6.36 y; 31.3% (5/16) of studies had a sample size smaller than 100 [20,24,26,27,31].

Through radiomics or texture analysis, 6 studies aimed to identify plaque vulnerability [18,19,23,26,31,32], 5 studies aimed to predict vascular events such as transient ischemic attack, stroke or cardiovascular disease [21,24,28–30], 3 studies tended to classify symptomatic carotid plaques [22,25,33], 1 study tried to assess radiomics robustness [27] and 1 study attempted to evaluate in-stent restenosis [19]. Details are outlined in Table 1.

RQS and risk of bias

Table 2 and Figure 2 summarize the RQS results. The lowest point was −2, and the highest score was 22. The mean RQS score was 9.25 (standard deviation: 6.04), representing 25.7% of the possible maximum value of 36. About 56.3% (9/16) of studies were credited between 10 and 20 points, corresponding to 27.8%–55.6% of total points. Among the 16 items, most studies achieved image protocol quality, multivariable analysis with non-radiomics features, feature reduction, discrimination statistics, and comparison with a gold standard, while many studies failed to analyze cutoff values and potential clinical utility. In addition, no researchers conducted a phantom study, imaging at multiple time points or cost-effectiveness analysis or shared open data. Only one study each proposed a prospective design [25] and biological correlates [26]; four studies did not have validation group [21,23,24,31].

The risk of bias and applicability concerns assessment for the 16 studies were summarized in Figure 3. The reviewers' consistency was good (k = 0.84, 95% confidence interval [CI]: 0.79–0.90). Although the general quality of the 16 studies was high, some possible sources of

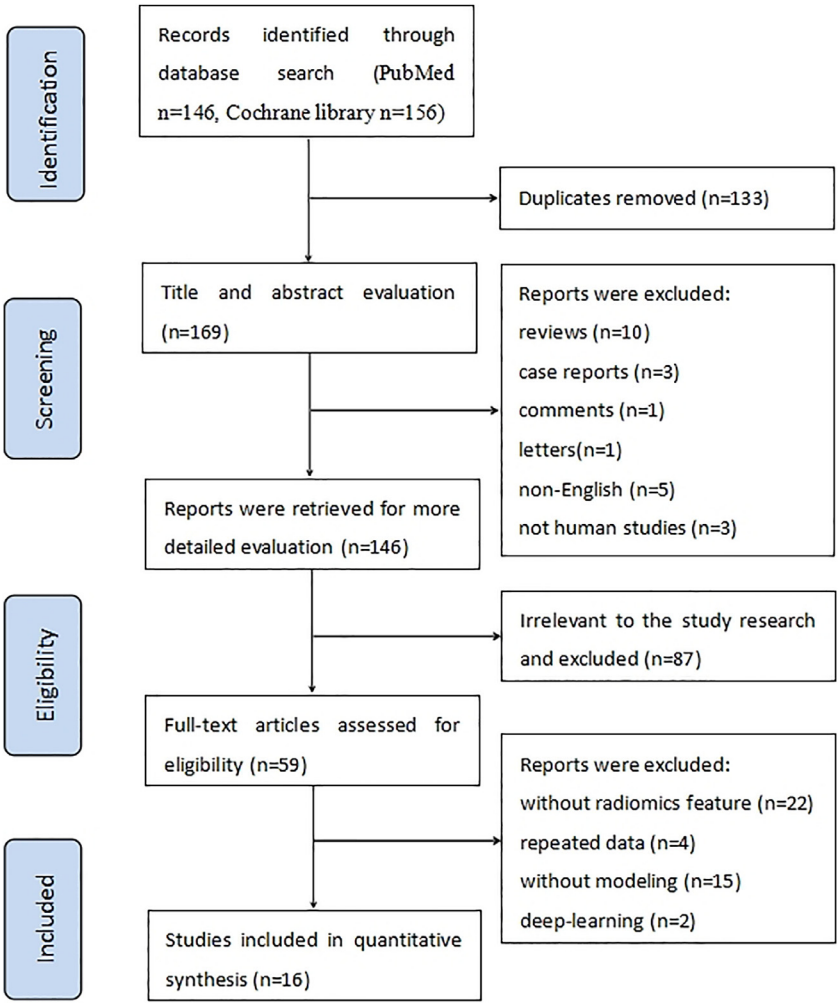

Figure 1. Flowchart of study selection according to Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) guidelines.

bias were identified with the QUADAS-2 tool: "flow and timing" was the dominant source of bias, followed by "patient selection." A total of 960 patients in 68.7% (11/16) of studies were not included in the final analysis [18–25,30–33], accounting for 28.3% of all recruited subjects. The interval between the index test and reference test was unclear in 31.2% (5/16) of studies [18,20,24,28,32]. It was unclear how patients were enrolled (consecutive or randomly or else) in 11 studies [18–21,26–31,33], and one study was a case–control design [31], both adding to the patient selection bias. About 43.8% (7/16) of studies did not report whether blindness was used [19,21,24,26–28,33]. The detailed results of RQS and QUADAS-2 assessment were available in Tables S2 and S3 (online only).

Imaging acquisition

The numbers of studies that used US, CTA, MRI, CT and [^{18}F]fluorodeoxyglucose positron emission tomography ([^{18}F]FDG PET) data to perform radiomics analysis for carotid atherosclerosis were 6, 6, 2, 1 and 1, respectively. All studies acquired images with a single modality to ensure the same technical parameters or methodology. To avoid data heterogeneity and bias, images in 7 studies were subjected to imaging normalization and resampled before region of interest (ROI) segmentation [18,20,25,27,29,32,33].

ROI segmentation and feature extraction

Nine different software were used for ROI segmentation, the most popular being ITK-SNAP and 3D-Slicer, and all but one of the studies used manual segmentation [29]. For those carotid arteries with more than one plaque, 50% (8/16) of studies chose the largest or most represented plaque to outline the ROI [19,22,24,25,30–33], whereas 18.8% (4/16) of studies tended to outline all the plaques [23,26,28,29], and the remaining 4 studies did not mention the details of ROI segmentation [18,20,21,27].

As for feature extraction, 7 different software were used; the most commonly used software was PyRadiomics, and the extracted features included first-order, shape, texture and wavelet features.

Feature selection and modeling

The number of extracted features ranged from 4 to 2107 in the included studies. As many of these features could be highly correlated with others, 13 studies used various means to effectively reduce feature reductant and increase robustness. And the least absolute shrinkage and selection operator (LASSO) algorithm was the most widely used in feature selection and reduction. The remaining 3 studies extracted only 4 [21], 6 [31] and 29 [28] features from their images, respectively.

After feature selection, various kinds of models were built. Through comparison of the performance of prediction models, 6 studies found the best model was a texture-based model [20,22,24,28,29,31], 5 studies combined traditional imaging features and radiomics features [19,26,27,32,33] and 5 studies also included clinical risk factors into a final combined prediction model [18,21,23,25,30]. Age, sex and history of hypertension, smoking, cardiac disease and high high-density lipoprotein (HDL), low-density lipoprotein (LDL)/HDL ratio, apolipoprotein B

Table 1
Study and patient characteristics

Author	Country	Study design	No. of patients (train vs. test)	Mean age	Men	Participant type	Purpose	Conclusions
Chen et al. 2022 [18]	China	R	115 (81:34)	51.38 ± 13.32	91	≥30% stenosis	Identification of plaque vulnerability	HRMRI texture features provide incremental value for carotid atherosclerotic risk assessment.
Cheng et al. 2022 [19]	China	R	221 (NA)	66.89 ± 8.07	186	Carotid endarterectomy	Prediction of in-stent restenosis	Radiomics and plaque features afforded the best predictive performance.
Cilla et al. 2022 [20]	Italy	R	30 (NA)	72.96	19	>70% internal carotid stenosis	Identification of plaque vulnerability	CTA-based radiomics and machine learning can discriminate plaque composition.
Colombi et al. 2021 [21]	Italy	R	172 (NA)	77	112	Carotid artery stenting	Prediction of unfavorable outcome after CAE	Kurtosis was an independent predictor of unfavorable outcome after CAS.
Dong et al. 2022 [22]	China	R	120 (NA)	66.68 ± 7.75	100	≥50% stenosis	Identification of symptomatic patients	Radiomics-based machine-learning analysis improves the discriminatory power of carotid CTA.
Doonan et al. 2016 [23]	Canada	R	160 (NA)	69.7 ± 9	71	Carotid endarterectomy	Evaluation the association of echodensity and textural features with plaque instability	Plaque echodensity and textural features are associated with histologic instability.
Ebrahimian et al. 2022 [24]	USA	R	85 (NA)	73 ± 10	56	Suspected or known ICA/CCA stenosis	Assessment of ICA/CCA stenosis and prediction of surgical outcome	DECT-based spectral radiomic features can differentiate patients with different luminal ICA/CCA stenosis grades.
Huang et al. 2022 [25]	China	P	548 (384:154)	62 ± 10	373	Carotid plaque	Identification of symptomatic plaques	Nomogram has a high diagnostic performance for identification of symptomatic carotid plaques.
Kafouris et al. 2021 [26]	France	R	21 (NA)	70.43 ± 7	18	High-grade carotid stenosis	Prediction of plaque vulnerability	Texture analysis can be applied in [^{18}F] FDG PET carotid imaging, providing valuable information for plaque characterization.
Le et al. 2021 [27]	UK	R	41 (NA)	63.47 ± 8.89	32	Carotid artery-related stroke or TIA	Assessment of radiomic robustness	A set of radiomic features are robust and have superior predictive performance for the classification of culprit versus non-culprit carotid arteries in patients with stroke and TIA.
Lo and Hung 2022 [28]	China	R	177 (NA)	61.5	89	Stroke	Diagnosis of ischemic stroke	CCD-based texture feature improves ischemic stroke diagnoses.
van Engelen et al. 2014 [29]	Netherlands	P	298 (NA)	70.45	110	Plaque area 40–600 mm^2	Prediction of vascular events	Plaque texture and volume changes are strongly predictive of vascular events.
Wang et al. 2022 [30]	China	R	105 (70:35)	63.4	73	CAD	Evaluation of the severity of CAD	Radiomics nomogram has potential for risk stratification of CAD before ICA.
Zaccagna et al. 2021 [31]	UK	R	24 (NA)	63 ± 10	14	Carotid atherosclerosis	Identification of plaque vulnerability	CT texture analysis can identify vulnerable patients in stroke and TIA.
Zhang et al. 2022 [32]	China	P	150 (105:45)	61.7 ± 10	120	Atherosclerotic plaque	Identification of plaque vulnerability	US texture feature can predict vulnerability of atherosclerotic plaque.
Zhang et al. 2021 [33]	China	R	162 (121:41)	66.8 ± 7.35	148	>30% stenosis	Identification of symptomatic plaques	MRI-based radiomics model can distinguish symptomatic from asymptomatic carotid plaques.

CAD, coronary artery disease; CAE, carotid endarterectomy; CAS, carotid artery stenting; CCD, carotid artery ultrasound; CTA, computed tomography angiography; DECT, dual-energy computed tomography; [^{18}F]FDG PET, fluorine-18-labeled fluorodeoxyglucose positron emission tomography; HRMRI, high-resolution magnetic resonance imaging; ICA, invasive coronary angiography; ICA/CCA, internal carotid artery/common carotid artery; NA, not available; P, prospective study; R, retrospective study; TIA, transient ischemia attack; US, ultrasonography.

and high-sensitivity C-reactive protein were independent clinical risk factors.

Performance of prediction models

The AUC represents the prediction power of models; the mean ± standard deviation in the training cohort was 0.88 ± 0.02 (range: 0.741–0.989), and 56.3% (9/16) of studies reported classification measures such as sensitivity, specificity and positive or negative predictive value. Seventy-five percent (12/16) of studies conducted internal validation through 5/10-fold or leave-one-out cross-validation, random resampling or bootstrap. However, only 37.5% (6/16) of those validation cohorts reported *C*-statistics [18,19,25,30,32,33], ranging from 0.83 to 0.986. Four studies drew calibration curves or performed the Hosmer–Lemeshow test to calibrate the models [18,24,30,32], three studies used decision curve analysis (DCA) to estimate the clinical utility of the models [18,25,30] and three studies calculated a cutoff value to classify vulnerable or symptomatic plaque [25,31,33]. Six studies drew a nomogram based on prediction models [18,19,25,30,32,33].

Discussion

This review selected 16 pieces of literatures for analysis and revealed a growing number of studies using radiomics to investigate high-risk carotid plaque, especially in the last 2 y. Based on a median population size of 135, the added value of radiomics in routine clinical imaging modalities has been extensively explored, with AUCs as high as 0.989 in the training set and 0.986 in the test set, highlighting the management and prediction of prognosis of patients with carotid plaque. Despite the promising and encouraging preliminary results, the overall quality of the included research is low, with a mean RQS score of 9.25 (<33.3% of the total points).

Table 2
Characteristics of the radiomics studies for carotid plaque and results of radiomics quality score

Study	Modality	Segmentation	Feature	Modeling method	No. of imaging features included in model	AUC in training set (95% CI)	Classification measures	Validation and method	RQS (point)
Chen et al. 2022 [18]	HRMRI	Manually	First-order, shape, texture, wavelet	LASSO, LR	8	0.929 (0.881−0.982)	ACC	Internal, 10-FCV	18
Cheng et al. 2022 [19]	CTA	Manually	First-order, shape, texture	COX, LASSO	6	0.88 (0.82−0.95)	SEN, SPE	Internal, 5-FCV	13
Cilla et al. 2022 [20]	CTA	Manually	First-order, shape, texture	LR, SVM, CART	2	0.987 (NA)	SEN, SPE, ACC, F-measure	Internal, 5-FCV	12
Colombi et al. 2021 [21]	CTA	Manually	First-order	LR	3	0.789 (0.73−0.847)	NA	NA	3
Dong et al. 2022 [22]	CTA	Manually	First-order, shape, texture	LR, SVM, XG-BOOST	20	0.858 (0.782−0.933)	NA	Internal, 5-FCV	12
Doonan et al. 2016 [23]	US	Manually	First-order, texture	LR	5	NA	NA	NA	3
Ebrahimian et al. 2022 [24]	DECT	Manually	First-order, shape, texture	LR	10	0.94 (NA)	SEN, SPE, ACC	NA	7
Huang et al. 2021 [25]	US	Manually	First-order, shape, texture	LASSO, LR	4	0.927 (0.90−0.956)	Cutoff, SEN, SPE, ACC, PPV, NPV	Internal, randomly	25
Kafouris et al. 2021 [26]	F18-FDG PET	Manually	First-order, texture	LR	1	0.97 (NA)	Youden index	Internal, 200 bootstrap	14
Le et al. 2021 [27]	CTA	Manually	First-order, texture	LASSO, SWV, decision tree, random forest	3	0.73 (NA)	SEN, SPE	Internal, 5-FCV	12
Lo and Hung 2022 [28]	CUS	Manually	First-order, shape, texture	LR, SVW,	11	0.94 (NA)	SEN, SPE, ACC, PPV, NPV	Internal, LOOCV	7
van Engelen et al. 2014 [29]	3D US	Semi-auto	Texture	Cox	8	0.78 (NA)	NA	Internal, 10-FCV	18
Wang et al. 2022 [30]	US	Manually	First-order, shape, texture, wavelet	LASSO, LR	11	0.741 (0.646−0.835)	SEN, SPE, ACC, Youden index	Internal, 10-FCV	16
Zaccagna et al. 2020 [31]	CT	Manually	First-order	NA	6	0.81 (NA)	Cutoff, SEN, SPE	NA	1
Zhang et al. 2022 [32]	US	Manually	First-order, texture, wavelet	LASSO, LR	8	0.88 (NA)	NA	Internal, 10-FCV	21
Zhang et al. 2020 [33]	MRI	Manually	First-order, shape, texture, wavelet	LASSO, LR	33	0.989 (NA)	Cutoff, SEN, SPE, ACC, NPV, PPV	Internal, 1000 bootstrap	17

ACC, accuracy; CART, classification and regression tree analysis; CTA, computed tomography angiography; DECT, dual-energy computed tomography; [^{18}F]FDG PET, fluorine-18-labeled fluorodeoxyglucose positron emission tomography; 5-FCV, 5-fold cross-validation; HRMRI, high-resolution magnetic resonance imaging; LASSO, least absolute shrinkage and selection operator; LOOCV, leave-one-out-validation; LR, logistic regression; NA, not available; NPV, negative predictive value; PPV, positive predictive value; RQS, radiomics quality score; SEN, sensitivity; SPE, specificity; SVM, support vector machine; US, ultrasonography.

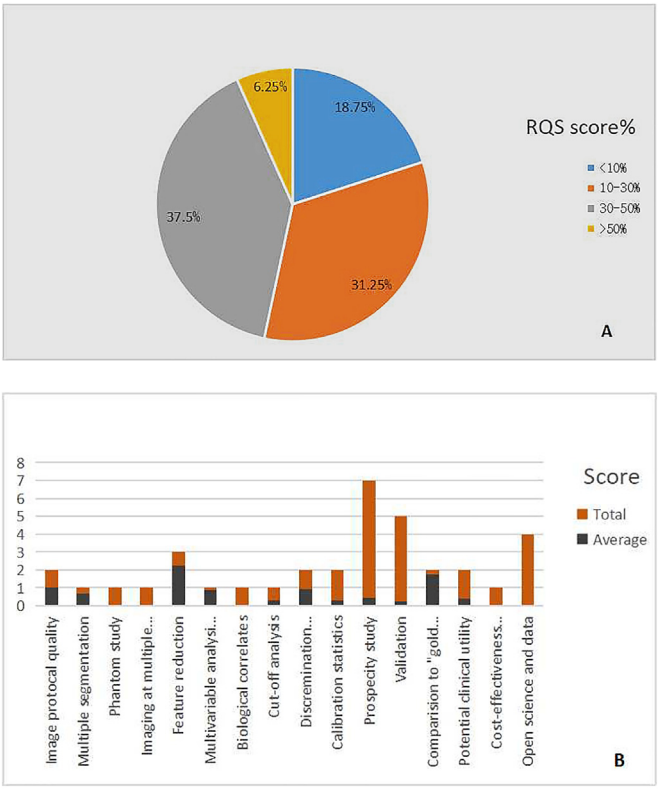

Figure 2. Methodological quality assessed by using the radiomics quality score (RQS) tool. (A) Proportion of studies with different RQS percentage scores. (B) Average scores of each RQS item (*red bars* represent the full points of each item, and *gray bars* represent actual points).

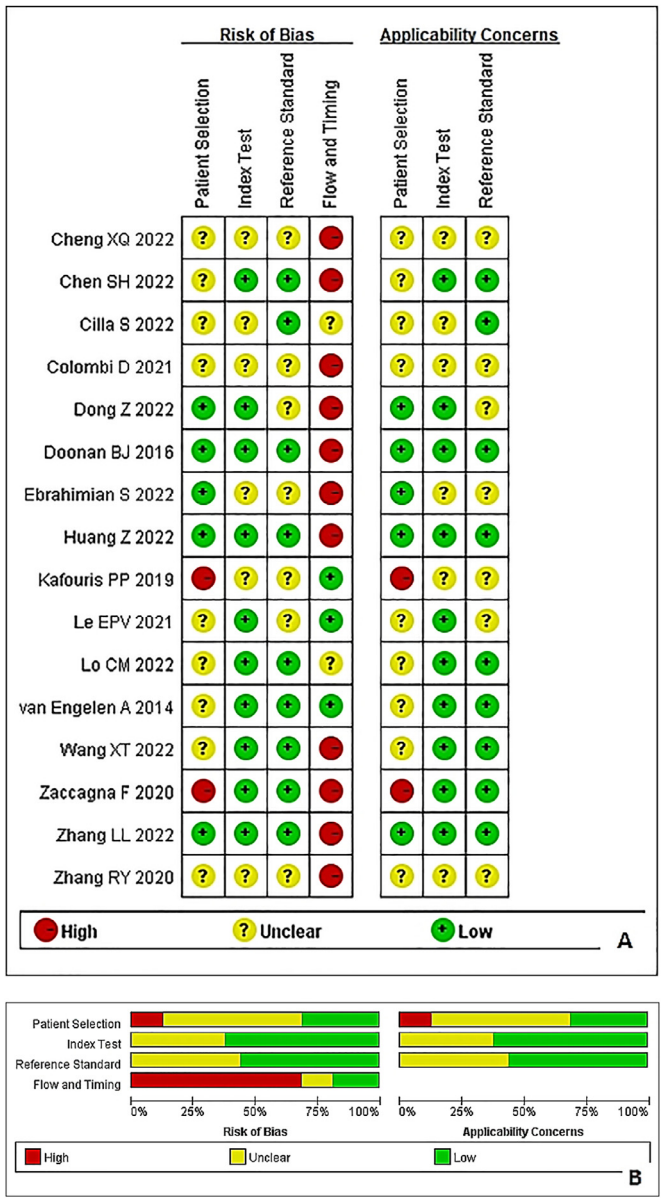

Figure 3. Risk of bias and applicability concerns of each included study (A) and the overall judgment (B) assessed by using the revised tool Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2).

As the bridge between medical imaging and personalized medicine, radiomics is quite a complex project, as the steps include data selection, medical imaging, feature extraction, algorithm operation and model construction. Each step of the workflow can be achieved by several strategies and approaches, which undoubtedly induces substantial methodological heterogeneity and bias among radiomics studies. Therefore, before translating the radiomics model to clinical application, standardization of the reporting norms is essential to guarantee the prediction models are reproducible and reliable. To normalize the process and evaluate the quality of radiomics reports, the founder of radiomics, Lambin, proposed the RQS tool in 2017. Recently, the quality assessment of radiomics research has currently been applied to oncology, the field in which radiomics was involved earliest, with an average RQS score from 3.41 to 13.5 (9.4%–37.5% of the total) [13,34,35]. Wakabayashi et al. [34] quantitatively reviewed the radiomics in hepatocellular carcinoma and included 23 studies that reached a median point of 8.35 ± 5.38. A lower score of 5.0 was assessed by Mühlbauer et al. [35] in a systematic review that contained 113 radiomics studies of renal cell carcinoma. In our study, an average RQS of 9.25 (range: −2 to 22) was calculated, suggesting the relatively low or moderate overall reporting quality of present radiomics models in carotid plaques, and the methodological variability of the studies is considerable.

The choice of imaging protocol, ROI, target event and reasonable study design is the beginning of radiomics analysis. Standardized and widely applied imaging protocols can eliminate unnecessary confounding variability, improving the reproducibility and comparability of studies across different populations, centers or regions. Before ROI segmentation, image pre-processing generally requires normalization to minimize gray value heterogeneity. In contrast, prior normalization of CTA/CT images seems unnecessary as gray values are already calibrated to Hounsfield units [27,28]. In the first domain of RQS, ROI segmentation is a critical step in determining which lesions to analyze in the image. An ideal carotid plaque segmentation method can accurately define the target lesion in the image with high reproducibility and low cost. Multiple segmentation, such as different clinicians, software, algorithms and breath cycles, is one way to reduce bias. In this review, most studies manually segmented the largest or most representative plaques from the axial and/or sagittal planes as a single 2-D slice by two physicians, leading to missing information and underestimation of intraplaque heterogeneity. Machine and deep learning can automatically target carotid plaques, further reducing bias [36]. However, the "phantom study" and "test–retest analysis" performed poorly in the included literature. These two items aimed to measure the uncertainties from organ motion or expansion, while a phantom study can detect possible feature variability of different vendors. A quantitative review of 77 oncology radiomics studies suggested that it is impractical to conduct two items for a clinical situation [37]. As radiomics biomarkers are supposed to be robust, some robust and reproducible studies using phantoms have been published in the oncology field [38–40], whereas relatively few data are available in cardiovascular imaging [41]. In our selected studies,

only Zhang et al. [33] extracted texture features from four different MRI sequences.

Like clinical prediction modeling, radiomics modeling involves three major aspects: feature/factor selection, modeling methodology and validation. Notably, both radiomics and non-radiomics features should be taken into consideration with the prediction target. In this review, Cilla et al. [20] and Lo and Hung [28] analyzed only differences in texture features, while other authors analyzed differences in clinical and traditional imaging between groups. Many modeling algorithms are available today, including LASSO, support vector machine, random forest, logistic/Cox regression and regression tree analysis. Extracting countless features in a limited number of people will result in redundancy. Documenting the various machine learning methods helps to eliminate unreliable factors and avoid overfitting. Additionally, validation is an integral part of a desirable radiomics model; a model without validation is of limited value. In RQS, the validation type performed accounts for as much as 10 of 36 points, and the highest score (5 points) belongs to a validation based on three or more external data sets. All 16 studies included in this review were single-center research studies, 12 of which were internally validated. Lack of external validation is a common phenomenon in radiomics analyses. Park et al.'s review [37] found that 81.8% of studies missed validation on external data sets, while Wang et al.'s review found this proportion to be even higher, at 90.1% [42].

Discrimination, calibration and cutoff analysis are metrics used to evaluate model performance, accounting for 2, 2 and 1 of the RQS, respectively. Discrimination refers to the degree to which a model differentiates risk with and without events, while calibration refers to the degree to which predicted risks are compared with observed outcomes. AUC, receiver operating characteristic (ROC) curve or *C*-statistic corresponding to *p* value, 95% confidence interval and calibration curves/plots with their statistical significance are usually used to report discrimination and calibration statistics, respectively. AUC is the equivalent to *C*-statistic in diagnostic prediction models [43]. A cutoff value of the ROC curve is used to divide the patients into low- and high-risk groups for a specific disease, which is also one of the requirements for replicating the results of previous studies. The average AUC was 0.88 ± 0.02 in the training cohort, indicating the good predictive power of these radiomics models. In the test cohorts, however, only 6 studies reported *C*-statistic, and only a quarter of studies presented cutoff values and calibration curves. DCA reports the current and potential application of the model in a clinical setting; however, only 3 studies performed a decision curve [18,25,30].

Prospective studies with sufficient sample sizes are a prerequisite to the reliability of experimental results, which accounted for the highest score of 7 points in RQS, whereas a retrospectively designed study introduces selection bias. As one of the most poorly performed items, the biological correlate reveals the possible associations between textures and phenotype or gene–protein expression patterns. Evaluation of biological variables and imaging findings often presented in oncology, and the term *radiogenomics* is widely used in the literature. Microscopically, high expression of C-reactive protein (CRP), interleukin (IL)-6, IL-18 and tumor necrosis factor-α indicate severe plaque inflammation, suggesting plaque vulnerability [44–46]. Oikonomou et al. [47] found that texture features of the coronary artery plaques were related to the expression of COL1A1 and CD31, which are markers of fibrosis and vascularity, respectively. Of these included studies, only Kafouris et al. [26] investigated the correlation between [^{18}F]PET/CT-based radiomics features and CD31 and CD68, indicating that texture features can predict the expression of CD31. Preliminary evidence has revealed a meaningful correlation between inflammatory molecules and contrast-enhanced ultrasound features [48]. As mentioned previously, radiomics features extracted from imaging data are able to identify vulnerable plaques and risk stratification. Therefore, linking imaging features to underlying tissue biology markers allows for a more personalized assessment of CAS progression.

Other poorly performed items were open science and data and cost-effectiveness analysis. Code and data sharing enables the initiation of highly powered prospective studies and accelerates the development and validation of radiomics signatures derived from new and existing data. Cost-effectiveness analysis can assess the value of radiomics predictive models in health economics when applied clinically [49,50]. It is expected that a new model with comparable accuracy should not be more expensive than previously available predictors. Considering the status of radiomics in methodology and clinical validation, Wang et al. [42] thought that the evaluation of this item was less urgent.

Under the guidance of the RQS tool, prospectively designed studies with standardized CT, MRI, US and PET/CT protocols can yield additional valuable information on carotid plaque patients. Future development may also require the CEUS modality for its promising diagnostic capability, especially in intraplaque neovascularization and morphology evaluation. An ideal predictive model should include an adequate sample size, commonly calculated by the number of events per variable (EPV). To avoid overfitting, ranking prediction frequently recommended an EPV sample $\geq$10 for categorical predictor [51,52], 20 individuals per factor for continuous outcomes [53] and at least 100 events and 100 non-events in validation models [54]. In addition to the RQS tool, the TRIPOD statement and CHARMS checklists are recommended guidelines for radiomics prediction modeling [53,55]. Authors need to record the number of candidate participants, the final population included, the presence or absence of missing data and the handling of missing data to better understand potential bias and reduce the negative impacts of missing data on model development and validation [56]. Furthermore, the validation set generated by an internal data set is not only similar to the test group but also has a small sample size that provides little extra information beyond the apparent performance, which actually increases the risk of bias. Therefore, external validation from different times, locations or repeating an existing modeling process in the validation data can improve the reproducibility and applicability [57,58]. There is still much room for exploring the link between the radiomics features and the biological phenotype of carotid plaques.

Limitations

This review has some inherent limitations. First, the sample size included is relatively small, and the quality of reports varies. Second, there are no independent, externally validated studies from which to draw convincing conclusions on the effectiveness of radiomics models in predicting carotid plaque. In addition, the heterogeneity and bias from various methodologies, feature selections, algorithms, modalities and study cohorts of the included studies should be mentioned. Furthermore, we did not perform a meta-analysis of the performance metrics of the integrated predictive models, nor could we compare model performance because of different research purposes, high heterogeneity and large differences in predictor variables.

Conclusion

The current study suggests that radiomics reports of carotid plaque are still in their infancy, and many hurdles need to be overcome. However, radiomics models have provided additional information on patients with carotid plaque. For diagnosis, radiomics features may help classify vulnerable plaques and symptomatic patients with higher accuracy than traditional imaging features. For prognosis, radiomics models could effectively predict cardiovascular events or unfavorable outcomes before and after management. In the future, prospective radiomics trials with standardized workflow, adequate output, optimal sample size and external validation will yield reliable and reproducible imaging predictions that are valid for clinical implementation and progression.

Acknowledgments

This study has received funding from the National Natural Science Foundation of China under Grant No. 82271995.

Conflict of interest

The authors declare no competing interests.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ultrasmedbio.2023.06.008.

References

[1] Barrett KM, Brott TG. Stroke caused by extracranial disease. Circ Res 2017;120:496–501.
[2] Global Burden of Disease Study (GBDS) Collaborators. Global, regional, and national burden of stroke and its risk factors, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Neurol 2021;20:795–820.
[3] Aboyans V, Ricco JB, Bartelink MEL, Bjorck M, Brodmann M, Cohnert T, et al. 2017 ESC guidelines on the diagnosis and treatment of peripheral arterial diseases, in collaboration with the European Society for Vascular Surgery (ESVS): document covering atherosclerotic disease of extracranial carotid and vertebral, mesenteric, renal, upper and lower extremity arteries. Endorsed by: the European Stroke Organization (ESO), The Task Force for the Diagnosis and Treatment of Peripheral Arterial Diseases of the European Society of Cardiology (ESC) and of the European Society for Vascular Surgery (ESVS). Eur Heart J 2018;39:763–816.
[4] Li Q, Cai M, Wang H, Chen L. Diagnostic performance of contrast-enhanced ultrasound and high-resolution magnetic resonance imaging for carotid atherosclerotic plaques: a systematic review and meta-analysis. J Ultrasound Med 2023;42:739–49.
[5] Sultan SR, Bashmail FT, Alzahrani NA, Alharbi SI, Anbar R, Alkharaiji M. Contrast-enhanced ultrasound for the evaluation of symptomatic and asymptomatic carotid plaques: a systematic review and meta-analysis. Echocardiography 2022;39:1032–43.
[6] Huang R, DeMarco JK, Ota H, Macedo TA, Abdelmoneim SS, Huston 3rd J, et al. Prognostic value of intraplaque neovascularization detected by carotid contrast-enhanced ultrasound in patients undergoing stress echocardiography. J Am Soc Echocardiogr 2021;34:614–24.
[7] Boswell-Patterson CA, Hetu MF, Kearney A, Pang SC, Tse MY, Herr JE, et al. Vascularized carotid atherosclerotic plaque models for the validation of novel methods of quantifying intraplaque neovascularization. J Am Soc Echocardiogr 2021;34:1184–94.
[8] Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer 2012;48:441–6.
[9] Lohmann P, Franceschi E, Vollmuth P, Dhermain F, Weller M, Preusser M, et al. Radiomics in neuro-oncological clinical trials. Lancet Digit Health 2022;4:e841–9.
[10] Chen Q, Pan T, Wang YN, Schoepf UJ, Bidwell SL, Qiao H, et al. A coronary CT angiography radiomics model to identify vulnerable plaque and predict cardiovascular events. Radiology 2023:221693.
[11] Chen Q, Zhang L, Mo X, You J, Chen L, Fang J, et al. Current status and quality of radiomic studies for predicting immunotherapy response and outcome in patients with non-small cell lung cancer: a systematic review and meta-analysis. Eur J Nucl Med Mol Imaging 2021;49:345–60.
[12] Janssen BV, Verhoef S, Wesdorp NJ, Huiskens J, de Boer OJ, Marquering H, et al. Imaging-based machine-learning models to predict clinical outcomes and identify biomarkers in pancreatic cancer: a scoping review. Ann Surg 2022;275:560–7.
[13] Yuan E, Chen Y, Song B. Quality of radiomics for predicting microvascular invasion in hepatocellular carcinoma: a systematic review. Eur Radiol 2023;33:3467–77.
[14] Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71.
[15] Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol 2017;14:749–62.
[16] Wolff RF, Moons KGM, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. Ann Intern Med 2019;170:51–8.
[17] Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic. Ann Intern Med 2011;155:529–36.
[18] Chen S, Liu C, Chen X, Liu WV, Ma L, Zha Y. A radiomics approach to assess high risk carotid plaques: a non-invasive imaging biomarker, retrospective study. Front Neurol 2022;13:788652.
[19] Cheng X, Dong Z, Liu J, Li H, Zhou C, Zhang F, et al. Prediction of carotid in-stent restenosis by computed tomography angiography carotid plaque-based radiomics. J Clin Med 2022;11:3234.
[20] Cilla S, Macchia G, Lenkowicz J, Tran EH, Pierro A, Petrella L, et al. CT angiography-based radiomics as a tool for carotid plaque characterization: a pilot study. Radiol Med 2022;127:743–53.
[21] Colombi D, Bodini FC, Rossi B, Bossalini M, Risoli C, Morelli N, et al. Computed tomography texture analysis of carotid plaque as predictor of unfavorable outcome after carotid artery stenting: a preliminary study. Diagnostics (Basel) 2021;11:2214.
[22] Dong Z, Zhou C, Li H, Shi J, Liu J, Liu Q, et al. Radiomics versus conventional assessment to identify symptomatic participants at carotid computed tomography angiography. Cerebrovasc Dis 2022;51:647–54.
[23] Doonan RJ, Gorgui J, Veinot JP, Lai C, Kyriacou E, Corriveau MM, et al. Plaque echodensity and textural features are associated with histologic carotid plaque instability. J Vasc Surg 2016;64 671–7.e8.
[24] Ebrahimian S, Homayounieh F, Singh R, Primak A, Kalra MK, Romero JM. Spectral segmentation and radiomic features predict carotid stenosis and ipsilateral ischemic burden from DECT angiography. Diagn Interv Radiol 2022;28:264–74.
[25] Huang Z, Cheng XQ, Liu HY, Bi XJ, Liu YN, Lv WZ, et al. Relation of carotid plaque features detected with ultrasonography-based radiomics to clinical symptoms. Transl Stroke Res 2022;13:970–82.
[26] Kafouris PP, Koutagiar IP, Georgakopoulos AT, Spyrou GM, Visvikis D, Anagnostopoulos CD. Fluorine-18 fluorodeoxyglucose positron emission tomography-based textural features for prediction of event prone carotid atherosclerotic plaques. J Nucl Cardiol 2021;28:1861–71.
[27] Le EPV, Rundo L, Tarkin JM, Evans NR, Chowdhury MM, Coughlin PA, et al. Assessing robustness of carotid artery CT angiography radiomics in the identification of culprit lesions in cerebrovascular events. Sci Rep 2021;11:3499.
[28] Lo CM, Hung PH. Computer-aided diagnosis of ischemic stroke using multi-dimensional image features in carotid color Doppler. Comput Biol Med 2022;147:105779.
[29] van Engelen A, Wannarong T, Parraga G, Niessen WJ, Fenster A, Spence JD, et al. Three-dimensional carotid ultrasound plaque texture predicts vascular events. Stroke 2014;45:2695–701.
[30] Wang X, Luo P, Du H, Li S, Wang Y, Guo X, et al. Ultrasound radiomics nomogram integrating three-dimensional features based on carotid plaques to evaluate coronary artery disease. Diagnostics (Basel) 2022;12:256.
[31] Zaccagna F, Ganeshan B, Arca M, Rengo M, Napoli A, Rundo L, et al. CT texture-based radiomics analysis of carotid arteries identifies vulnerable patients: a preliminary outcome study. Neuroradiology 2021;63:1043–52.
[32] Zhang L, Lyu Q, Ding Y, Hu C, Hui P. Texture analysis based on vascular ultrasound to identify the vulnerable carotid plaques. Front Neurosci 2022;16:885209.
[33] Zhang R, Zhang Q, Ji A, Lv P, Zhang J, Fu C, et al. Identification of high-risk carotid plaque with MRI-based radiomics and machine learning. Eur Radiol 2021;31:3116–26.
[34] Wakabayashi T, Ouhmich F, Gonzalez-Cabrera C, Felli E, Saviano A, Agnus V, et al. Radiomics in hepatocellular carcinoma: a quantitative review. Hepatol Int 2019;13:546–59.
[35] Muhlbauer J, Egen L, Kowalewski KF, Grilli M, Walach MT, Westhoff N, et al. Radiomics in renal cell carcinoma—a systematic review and meta-analysis. Cancers (Basel) 2021;13:1348.
[36] Guang Y, He W, Ning B, Zhang H, Yin C, Zhao M, et al. Deep learning-based carotid plaque vulnerability classification with multicentre contrast-enhanced ultrasound video: a comparative diagnostic study. BMJ Open 2021;11:e047528.
[37] Park JE, Kim D, Kim HS, Park SY, Kim JY, Cho SJ, et al. Quality of science and reporting of radiomics in oncologic studies: room for improvement according to radiomics quality score and TRIPOD statement. Eur Radiol 2020;30:523–36.
[38] Hatt M, Tixier F, Cheze Le Rest C, Pradier O, Visvikis D. Robustness of intratumour [^{18}F]FDG PET uptake heterogeneity quantification for therapy response prediction in oesophageal carcinoma. Eur J Nucl Med Mol Imaging 2013;40:1662–71.
[39] Schwier M, van Griethuysen J, Vangel MG, Pieper S, Peled S, Tempany C, et al. Repeatability of multiparametric prostate MRI radiomics features. Sci Rep 2019;9:9441.
[40] Buch K, Kuno H, Qureshi MM, Li B, Sakai O. Quantitative variations in texture analysis features dependent on MRI scanning parameters: a phantom model. J Appl Clin Med Phys 2018;19:253–64.
[41] Edalat-Javid M, Shiri I, Hajianfar G, Abdollahi H, Arabi H, Oveisi N, et al. Cardiac SPECT radiomic features repeatability and reproducibility: a multi-scanner phantom study. J Nucl Cardiol 2021;28:2730–44.
[42] Wang Q, Li C, Zhang J, Hu X, Fan Y, Ma K, et al. Radiomics models for predicting microvascular invasion in hepatocellular carcinoma: a systematic review and radiomics quality score assessment. Cancers (Basel) 2021;13:5864.
[43] Debray TP, Damen JA, Snell KI, Ensor J, Hooft L, Reitsma JB, et al. A guide to systematic review and meta-analysis of prediction model performance. BMJ 2017;356:i6460.
[44] Li D, Li B, Yang L, Wang J. Human cytomegalovirus infection is correlated with atherosclerotic plaque vulnerability in carotid artery. J Gene Med 2020;22:e3236.
[45] Tomas L, Edsfeldt A, Mollet IG, Perisic Matic L, Prehn C, Adamski J, et al. Altered metabolism distinguishes high-risk from stable carotid atherosclerotic plaques. Eur Heart J 2018;39:2301–10.
[46] Zhang X, Zhang T, Wu C, Zhou Y, Chen Z, Xu R. The association between inflammatory biomarkers and carotid artery plaque in normal-weight and metabolically healthy Chinese adults: a cross-sectional study. Hypertens Res 2023;46:330–8.
[47] Oikonomou EK, Williams MC, Kotanidis CP, Desai MY, Marwan M, Antonopoulos AS, 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.
[48] Ammirati E, Moroni F, Magnoni M, Di Terlizzi S, Villa C, Sizzano F, et al. Circulating $CD14^{+}$ and $CD14^{high}CD16^{-}$ classical monocytes are reduced in patients with signs of plaque neovascularization in the carotid artery. Atherosclerosis 2016;255:171–8.
[49] Sanghera S, Barton P, Bhattacharya S, Horne AW, Roberts TE, PRE-EMPT Research Group. Pharmaceutical treatments to prevent recurrence of endometriosis following surgery: a model-based economic evaluation. BMJ Open 2016;6:e010580.

[50] Wu B, Yang Z, Tobe RG, Wang Y. Medical therapy for preventing recurrent endometriosis after conservative surgery: a cost-effectiveness analysis. BJOG 2018;125:469–77.
[51] Courvoisier DS, Combescure C, Agoritsas T, Gayet-Ageron A, Perneger TV. Performance of logistic regression modeling: beyond the number of events per variable, the role of data structure. J Clin Epidemiol 2011;64:993–1000.
[52] Vittinghoff E, McCulloch CE. Relaxing the rule of ten events per variable in logistic and Cox regression. Am J Epidemiol 2007;165:710–8.
[53] Moons KG, de Groot JA, Bouwmeester W, Vergouwe Y, Mallett S, Altman DG, et al. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. PLoS Med 2014;11:e1001744.
[54] Vergouwe Y, Steyerberg EW, Eijkemans MJ, Habbema JD. Substantial effective sample sizes were required for external validation studies of predictive logistic regression models. J Clin Epidemiol 2005;58:475–83.
[55] Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. Ann Intern Med 2015;162:W1–73.
[56] Burton A, Altman DG. Missing covariate data within cancer prognostic studies: a review of current reporting and proposed guidelines. Br J Cancer 2004;91:4–8.
[57] Jiang W, Wang J, Shen X, Lu W, Wang Y, Li W, et al. Establishment and validation of a risk prediction model for early diabetic kidney disease based on a systematic review and meta-analysis of 20 cohorts. Diabetes Care 2020;43:925–33.
[58] Poorthuis MHF, Halliday A, Massa MS, Sherliker P, Clack R, Morris DR, et al. Validation of risk prediction models to detect asymptomatic carotid stenosis. J Am Heart Assoc 2020;9:e014766.