European Radiology (2023) 33:6759–6770
https://doi.org/10.1007/s00330-023-09672-3

European Radiology
EUROPEAN SOCIETY OF RADIOLOGY

IMAGING INFORMATICS AND ARTIFICIAL INTELLIGENCE

Development and validation of a deep learning model for prediction of intracranial aneurysm rupture risk based on multi-omics factor

Mirzat Turhon[1,2] · Mengxing Li[1,2] · Huibin Kang[1,2] · Jiliang Huang[1,2] · Fujunhui Zhang[1,2] · Ying Zhang[1,2] · Yisen Zhang[1,2] · Aierpati Maimaiti[3] · Dilmurat Gheyret[3] · Aximujiang Axier[3] · Miamaitili Aisha[3] · Xinjian Yang[1,2] · Jian Liu[1,2]

Received: 21 September 2022 / Revised: 27 January 2023 / Accepted: 24 February 2023 / Published online: 26 April 2023

Abstract

Objective The clinical ability of radiomics to predict intracranial aneurysm rupture risk remains unexplored. This study aims to investigate the potential uses of radiomics and explore whether deep learning (DL) algorithms outperform traditional statistical methods in predicting aneurysm rupture risk.

Methods This retrospective study included 1740 patients with 1809 intracranial aneurysms confirmed by digital subtraction angiography at two hospitals in China from January 2014 to December 2018. We randomly divided the dataset (hospital 1) into training (80%) and internal validation (20%). External validation was performed using independent data collected from hospital 2. The prediction models were developed based on clinical, aneurysm morphological, and radiomics parameters by logistic regression (LR). Additionally, the DL model for predicting aneurysm rupture risk using integration parameters was developed and compared with other models.

Results The AUCs of LR models A (clinical), B (morphological), and C (radiomics) were 0.678, 0.708, and 0.738, respectively (all $p < 0.05$). The AUCs of the combined feature models D (clinical and morphological), E (clinical and radiomics), and F (clinical, morphological, and radiomics) were 0.771, 0.839, and 0.849, respectively. The DL model (AUC = 0.929) outperformed the machine learning (ML) (AUC = 0.878) and the LR models (AUC = 0.849). Also, the DL model has shown good performance in the external validation datasets (AUC: 0.876 vs 0.842 vs 0.823, respectively).

Conclusion Radiomics signatures play an important role in predicting aneurysm rupture risk. DL methods outperformed conventional statistical methods in prediction models for the rupture risk of unruptured intracranial aneurysms, integrating clinical, aneurysm morphological, and radiomics parameters.

Key Points

- *Radiomics parameters are associated with the rupture risk of intracranial aneurysms.*
- *The prediction model based on integrating parameters in the deep learning model was significantly better than a conventional model.*
- *The radiomics signature proposed in this study could guide clinicians in selecting appropriate patients for preventive treatment.*

Keywords Intracranial aneurysm · Stroke · Deep learning

Mirzat Turhon and Mengxing Li are joint first authors.

✉ Miamaitili Aisha
mmtaili@aliyun.com

✉ Xinjian Yang
yangxinjian@voiceoftiantan.org

✉ Jian Liu
jianliu_ns@163.com

[1] Department of Interventional Neuroradiology, Beijing Neurosurgical Institute, Capital Medical University, Beijing 100070, People's Republic of China

[2] Department of Neurosurgery, Beijing TianTan Hospital, Capital Medical University, Beijing, People's Republic of China

[3] Department of Neurosurgery, Xinjiang Medical University Affiliated First Hospital, Urumqi, Xinjiang 840017, People's Republic of China

Abbreviations

ACA	Anterior cerebral artery
AComA	Anterior communication artery
AI	Artificial intelligence
AR	Aspect ratio
AUC	Area under curve
CASAM	Computer-assisted semi-automated measurement
DL	Deep learning
DSA	Digital subtraction angiography
GBDT	Gradient-boosted decision trees
ICA	Internal carotid artery
LASSO	Least absolute shrinkage and selection operator
LR	Logistic regression
MCA	Middle cerebral artery
ML	Machine learning
NSI	Nonsphericity index
PC	Posterior circulation
PComA	Posterior communication artery
RF	Random forest
ROC	Receiver operating characteristic
SAH	Subarachnoid hemorrhage
SR	Size ratio
UI	Undulation index
UIA	Unruptured intracranial aneurysm

Introduction

Unruptured intracranial aneurysms (UIAs) are common life-threatening conditions, with an incidence of more than 3% worldwide and 7% in China [1, 2]. While the annual incidence of aneurysm rupture is typically less than 1%, ruptured aneurysms can lead to fatal events, such as subarachnoid hemorrhage (SAH), with high mortality and morbidity [3, 4]. As a result, UIA treatment is often performed as a preventive measure before ruptures occur. However, UIA treatments, including both neurosurgical clipping and endovascular coiling, are associated with a significant risk of complications [5]. A meta-analysis of 114 studies (108,263 UIAs) found that endovascular treatment (74 studies) was associated with a 4.96% risk of morbidity and 0.30% mortality, and neurosurgical treatment (54 studies) had an 8.34% risk of morbidity and 0.10% mortality [5]. Currently, the treatment decision for UIAs is the balance of the risk between aneurysm rupture and proactive procedural-related complications. Therefore, discriminating the rupture risk of UIAs is an important part of managing and treating patients.

In recent years, artificial intelligence (AI) has been used to detect, predict the stability of, and simulate operating on UIAs [6]. Radiomics is an emerging branch of AI that converts digital medical images (also known as features) containing information relevant to tumor pathophysiology into measurable and quantifiable data. Along with clinical and qualitative imaging data, these data can improve medical decision-making [7]. Liu et al applied pyradiomics-extracted shape features to predict the stability of aneurysms, obtaining satisfactory results (AUC = 0.853) [8]. There is growing interest in using machine learning and morphological and radiological properties to predict aneurysm ruptures. However, few studies have independently demonstrated and confirmed the effectiveness of deep learning (DL) models using radiomics parameters to predict the rupture risk of aneurysm ruptures.

In this paper, we sought to determine the ability of radiomics features to predict the rupture risk of UIAs. Furthermore, we developed a DL predicting model to assess the rupture risk of UIAs and explored whether they outperformed previously reported statistical methods.

Methods

Study design and participants

This study was approved by the local ethics committees, and each patient provided written informed consent before surgery. From January 2014 to December 2018, one hospital treated 2179 patients with intracranial aneurysms. The inclusion criteria were as follows: (1) at least one intracranial aneurysm confirmed by digital subtraction angiography (DSA) imaging; (2) older than 18 years; and (3) accessibility of morphological and radiomics data. The exclusion criteria were as follows: (1) diagnosis of traumatic, bacterial, or dissecting aneurysm; (2) presence of other intracranial vascular malformations, such as a cerebral arteriovenous malformation or cerebral arteriovenous fistula; and (3) absence of clinical data or 3D-DSA imaging.

The final internal cohort included 1391 patients with 1451 aneurysms. An additional independent dataset was used for external validation at another hospital (349 patients with 358 intracranial aneurysms). The flow chart of the inclusion and exclusion processes is shown in Fig. 1.

Acquisition of clinical, morphological, and radiomics features

The following clinical features were collected from the medical records of both hospitals: age, gender, smoking status, drinking status, hypertension, hyperlipemia, diabetes mellitus, coronary heart disease, family history, multiplicity, and the presence of symptoms.

Morphologic parameters were measured from 3D-DSA imaging using computer-assisted semi-automated measurement (CASAM), according to previously described methods

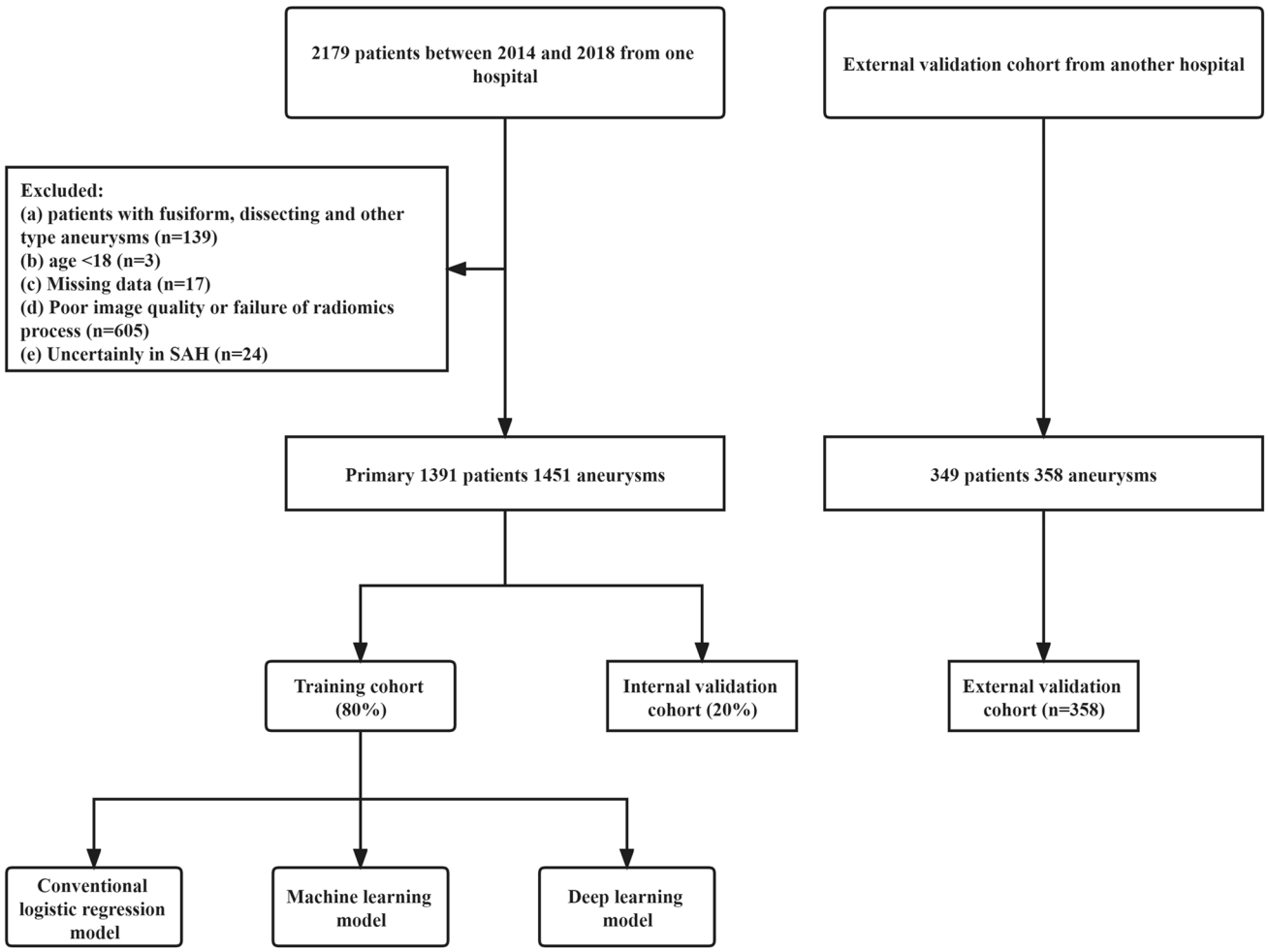

Fig. 1 Flowchart of the patients' inclusion and exclusion process

[9, 10]. CASAM allowed us to obtain 21 morphological and derived parameters such as the diameter, width, height, and neck diameter of the aneurysm. All CASAM procedures were performed by one experienced neuro-interventionist according to the perfect inter-class consistency of CASAM (Supplementary Method 1 and Supplementary Fig. 1). Additionally, a central review committee consisting of two neuroradiologists and a senior neuro-interventionist reviewed all of the imaging data and processing. Each evaluation was performed independently by two members, and in cases when the evaluation results were disputed, the team reached a unanimous decision after a discussion.

Aneurysms with irregular shapes were defined based on the presence of daughter sacs or lobules, according to a previous study [10]. The relative location of the aneurysm to the parent vessel was classified as either a sidewall or bifurcation type [11]. Aneurysm locations were classified as internal carotid artery (ICA), anterior cerebral artery (ACA), anterior communicating artery (AcomA), middle cerebral artery (MCA), and posterior circulation (PC, including vertebral artery, basilar artery, posterior cerebral artery, and anterior and posterior cerebellar arteries).

Regarding the morphological features, CASAM software was used for vascular reconstruction and aneurysm segmentation. The segmentation mask was then input into the PyRadiomics module in Python to extract the radiomics features. The 3D-shape feature is an important feature extraction in radiomics, mainly based on the description of the size and shape of the three-dimensional ROI image of the aneurysm. These features are independent of the gray-level intensity distribution in the ROI and are therefore only computed from non-derived images and masks [12]. Seventeen 3D morphological features were automatically extracted for each aneurysm. These included Flatness, Compactness 1 and 2, Elongation, Surface Area, Maxium2DDiameter (Slice, Column, and Row), SurfaceVolumeRatio, Maximum3DDiameter, Sphericity, and SphericalDisproportion. Flatness indicates the relationship between the largest and smallest principal components of the ROI shape. Elongation represents the relationship between the two principal components of the

ROI form. The remaining definitions of 3D-shape radiomics features are described in Supplementary Table 1.

Dimensionality reduction and radiomics feature selection

Feature extraction and dimensionality reduction were performed using Python (version 3.4) software. First, similar high-correlation features were removed using Spearman's correlation coefficient to construct radiomics features. Pairs of traits with a Spearman correlation coefficient greater than 0.9 were considered strongly correlated, and only one of the two traits were accepted. Second, the least absolute shrinkage and selection operator (LASSO) method was used to select the most robust discriminative radiomics features with sufficient reproducibility and association with rupture risk in the training cohort [13]. LASSO regression is a type of multivariate linear regression that efficiently reduces the number of features and regularization. The best-performing radiomics features were used for subsequent analysis.

Construction of models

The model construction workflow is displayed in Fig. 2. To determine the ability of radiomics parameters to predict the rupture risk of UIAs, we first constructed prediction models based on three signatures (clinical, morphological, and radiomics) for comparison. Model A was established based on clinical features. Model B is a baseline model constructed using conventional morphological parameters automatically measured directly by CASAM. Model C is a radiomics morphological feature model selected from LASSO regression. All variables were first tested by statistical analysis. The traditional logistic regression (LR) model selects only variables with statistical significance at the $p < 0.05$ level in the univariate analysis as input variables.

Second, we combined different features based on the three previous models to evaluate whether the combined model can further improve the performance of the predicting model. Model D was created based on clinical and morphological features, and model E was created based on clinical and radiomics features. Model F was constructed by combining clinical, morphological, and radiomics features.

Finally, we developed a DL model and explored whether it was better than the traditional LR model and ML algorithms at predicting the rupture risk of UIAs based on integrated features. To determine the most suitable machine learning algorithm, we tested six popular machine learning algorithms and selected the gradient-boosted decision trees (GBDT) algorithm as our machine learning classifier. The deep learning algorithm used in this study was the Transformer. We also used three ML

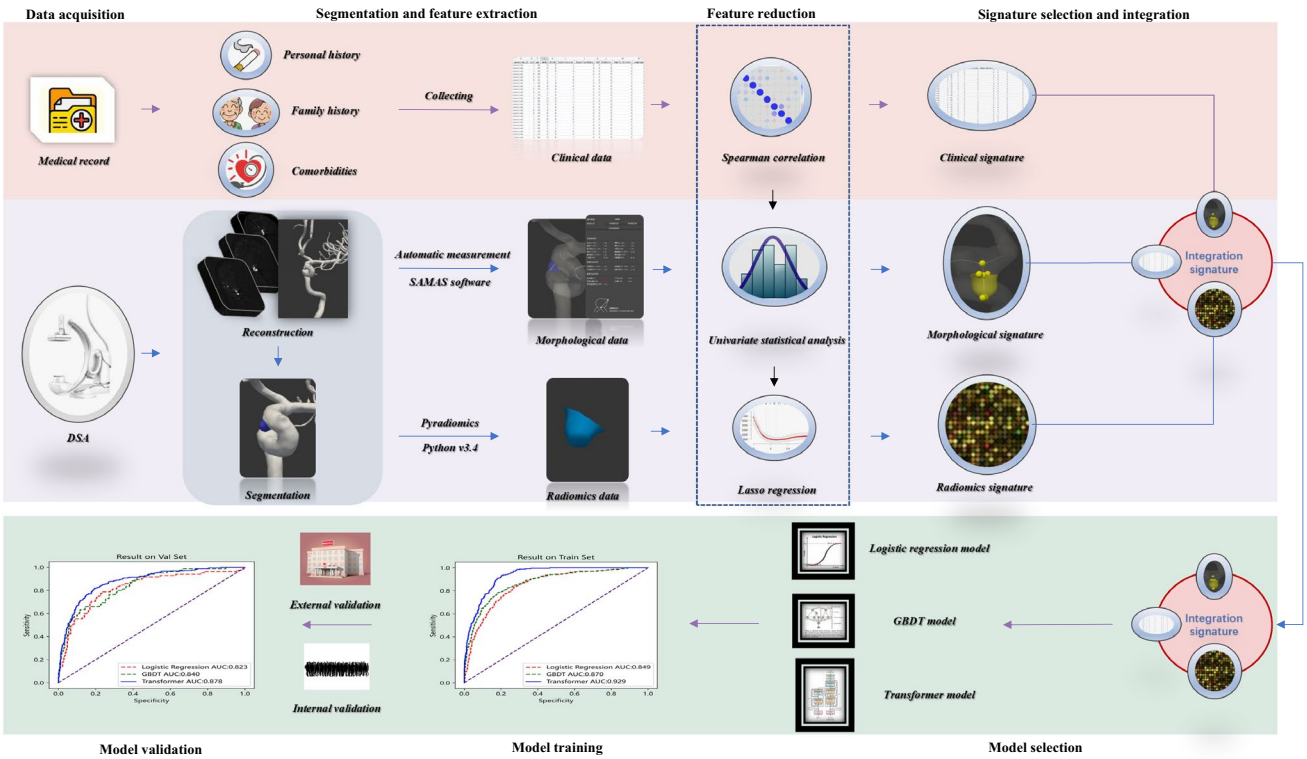

Fig. 2 Workflow of the study

models, including Random Forest (RF), Adaboost, and GBDT, to identify the five most important features relating to the rupture risk of UIAs in three signatures. A brief introduction to these algorithms is included in Supplementary Method 2.

Statistical analysis and model performance evaluation

The R Project for Statistical Computing (version 4.1.2, R Foundation for Statistical Computing, Vienna, Austria) was used to assess the ability of each parameter to distinguish the rupture state of aneurysms. For continuous variables, we first used the Shapiro-Wilk test to determine normality, followed by Student's *t* test or the Mann-Whitney *U* test. For categorical variables, Fisher's exact test or chi-square test was used to assess differences between groups. For normally distributed data, comparisons between the three groups were tested using analysis of variance (ANOVA). Abnormal data were analyzed using independent-sample nonparametric tests. A two-sided p value < 0.05 was considered statistically significant. We then generated receiver operating characteristic (ROC) curves and calculated the area under the curve (AUC) values to compare the effectiveness of the different models. The ROC curves between different methods were calculated using the method described by Delong et al compared to the asymptotically precise method previously described [14].

Results

Clinical, morphological, and radiomics parameters

A total of 1391 consecutive patients with 1451 intracranial aneurysms were included in our study. Of the 1451 aneurysms, we randomly divided the dataset into training (80%) and internal validation (20%) cohorts. The clinical, morphologic, and radiologic characteristics of ruptured (272) and non-ruptured (885) aneurysm patients in the training cohort are displayed in Table 1. In univariate analysis, there were significant differences ($p < 0.05$) between the ruptured and unruptured groups in terms of sex, age, smoking, alcohol abuse, hypertension, presence of a symptom, family history, multiplicity, location, aneurysm width, aneurysm transverse diameter, aneurysm volume, aneurysm neck diameter, flow angle, aneurysm length diameter-to-width ratio, aneurysm neck area, aneurysm transverse diameter-to-neck diameter ratio, the diameter of the parent artery, aspect ratio (AR), size ratio (SR), undulation index (UI), nonsphericity index (NSI), irregularity, bifurcation, mesh volume, voxel volume, surface area, surface area-to-volume ratio, sphericity, compactness 1, compactness 2, spherical disproportion, maximum 2D diameter (slice and column), minor and least axis length, elongation, and flatness.

We retrospectively collected data for a cohort of 349 patients with 358 aneurysms at another hospital for external validation. Table 2 shows the detailed clinical, morphologic, and radiomics characteristics in the training and inside-outside validation cohorts. There were significant differences ($p < 0.05$) between the training group and the internal and external validation groups with respect to clinical factor (mean age, $p = 0.030$) and radiomics parameters (surface area-to-volume ratio, $p = 0.049$; maximum 2D diameter slice, $p = 0.020$; major axis length, $p = 0.008$).

LR prediction model performance and validation

The performance of the LR prediction model is shown in Fig. 3. Figure 3a shows the ROC curves of three based prediction models in the training cohort. The AUCs of model A (clinical characteristic), B (morphological characteristics), and C (radiomics characteristics) were 0.708 (95% CI, 0.636 to 0.780), 0.678 (95% CI, 0.595 to 0.761), and 0.738 (95% CI, 0.666 to 0.810), respectively. Of these three models, the radiomics model had the best performance (all $p < 0.05$).

We further combined the above three signatures in different ways to obtain three combined models. Figure 3b shows the ROC curves of three combined prediction models in the training cohort. The AUCs of models D (clinical and morphological characteristic), E (clinical and radiomics characteristics), and F (clinical, morphological, and radiomics characteristics) were 0.771 (95% CI, 0.698 to 0.844), 0.839 (95% CI, 0.785 to 0.893), and 0.849 (95% CI, 0.808 to 0.890), respectively. Model F integrated all signatures and outperformed the other two, while the model including radiomics with the clinical signatures was better than the model with morphological signatures. We further validated these six models using internal and external validation sets to assess their generality and broad applicability (Supplementary Table 2 and Supplementary Fig. 2).

DL model performance and validation

Figure 4 shows the AUCs, accuracy, sensitivity, and specificity of the conventional LR, ML, and DL algorithms in the three datasets. In the training datasets (Fig. 4a), the Transformer model had the best performance, with an AUC value of 0.929 (95% CI, 0.893–0.965). The AUC value of the GBDT model was 0.878 (95% CI, 0.840–0.916). The LR model performed relatively poorly, with an AUC value of 0.849 (95% CI, 0.808–0.890).

The ROC curve and AUC values of each model in two validation datasets are shown in Fig. 4b, c. In the internal validation datasets, the AUCs of the LR, GBDT, and Transformer models were 0.812 (95% CI, 0.757 to 0.867), 0.851 (95% CI, 0.809 to 0.893), and 0.878 (95% CI, 0.834 to 0.922), respectively.

Table 1 Clinical, morphological, and radiomic characteristics of the ruptured and unruptured aneurysm patients in the training cohort

Variables	Unruptured group ($n=885$)	Ruptured group ($n=272$)	p value
Clinical characteristics			
Age	55.50 ± 9.40	54.35 ± 12.20	0.045
Female	609 (68.8%)	154 (56.6%)	< 0.001
Smoking	167 (18.9%)	91 (33.5%)	< 0.001
Alcohol abuse	154 (17.4%)	92 (33.8%)	< 0.001
Hypertension	429 (48.5%)	165 (60.7%)	< 0.001
Hyperlipidemia	85 (9.6%)	37 (13.6%)	0.060
Coronary heart disease	85 (9.6%)	3 (11.0%)	0.055
Diabetes mellitus	97 (11.0%)	34 (12.5%)	0.483
Presence of symptom	454 (51.3%)	213 (78.3%)	< 0.001
Family history	60 (6.8%)	40 (14.7%)	< 0.001
Multiplicity	370 (41.8%)	74 (27.2%)	< 0.001
Morphological characteristics			
Location			< 0.001
ICA	601 (67.9%)	118 (43.4%)	
MCA	102 (11.5%)	32 (11.8%)	
ACA	35 (4.0%)	17 (6.2%)	
PC	61 (6.9%)	14 (5.1%)	
AcomA	86 (9.7%)	91 (33.5%)	
Aneurysm length diameter	4.60 ± 2.38	4.81 ± 2.03	0.153
Aneurysm width	5.08 ± 2.15	4.66 ± 1.73	0.001
Aneurysm height	4.08 ± 2.22	4.10 ± 1.79	0.879
Aneurysm transverse diameter	5.28 ± 2.28	4.87 ± 1.85	0.003
Aneurysm maximum diameter	6.35 ± 2.73	6.34 ± 2.30	0.952
Aneurysm volume	82.79 ± 161.74	56.89 ± 78.55	< 0.001
Aneurysm neck diameter	4.67 ± 1.78	4.15 ± 1.39	< 0.001
Flow angle	100.19 ± 32.97	114.67 ± 30.03	< 0.001
Aneurysm length diameter-to-width ratio	0.90 ± 0.21	1.04 ± 0.25	< 0.001
Aneurysm neck area	19.60 ± 22.35	15.03 ± 11.53	< 0.001
Aneurysm transverse diameter-to-neck diameter ratio	1.14 ± 0.26	1.19 ± 0.29	0.011
Aneurysm angle	52.52 ± 20.64	50.78 ± 19.62	0.219
Proximal diameter of parent artery	3.18 ± 0.86	2.61 ± 0.77	< 0.001
Distal diameter of parent artery	2.93 ± 0.85	2.37 ± 0.70	< 0.001
Average diameter of parent artery	3.19 ± 0.81	2.61 ± 0.69	< 0.001
Length diameter of parent vessel	13.29 ± 3.92	12.48 ± 3.19	0.001
AR	1.01 ± 0.44	1.17 ± 0.44	< 0.001
SR	1.59 ± 0.99	2.00 ± 1.02	< 0.001
UI	0.16 ± 0.06	0.19 ± 0.06	< 0.001
NSI	0.36 ± 0.04	0.38 ± 0.04	< 0.001
VNR	4.56 ± 5.33	4.68 ± 4.56	0.716
Irregularity	158 (17.9%)	96 (35.3%)	< 0.001
Bifurcation	314 (35.5%)	184 (67.6%)	< 0.001
Radiomic characteristics			
Mesh volume	205.18 ± 627.71	137.19 ± 309.51	0.016
Voxel volume	205.68 ± 627.85	137.68 ± 309.75	0.016
Surface area	164.07 ± 218.46	131.60 ± 159.09	0.008
Surface area-to-volume ratio	1.71 ± 0.94	1.89 ± 0.98	0.006
Sphericity	0.79 ± 0.06	0.77 ± 0.06	< 0.001
Compactness 1	0.037 ± 0.004	0.036 ± 0.004	< 0.001
Compactness 2	0.50 ± 0.10	0.47 ± 0.10	< 0.001

Table 1 (continued)

Variables	Unruptured group ($n=885$)	Ruptured group ($n=272$)	p value
Spherical disproportion	1.27 ± 0.11	1.30 ± 0.12	< 0.001
Maximum 3D diameter	7.95 ± 4.20	7.63 ± 4.00	0.267
Maximum 2D diameter (slice)	7.03 ± 3.74	6.45 ± 3.38	0.016
Maximum 2D diameter (column)	7.10 ± 3.82	6.59 ± 3.50	0.040
Maximum 2D diameter (row)	7.15 ± 3.76	6.66 ± 3.51	0.057
Major axis length	6.68 ± 3.62	6.56 ± 3.57	0.632
Minor axis length	5.12 ± 2.72	4.46 ± 2.26	< 0.001
Least axis length	4.23 ± 2.36	3.68 ± 1.95	< 0.001
Elongation	0.78 ± 0.11	0.70 ± 0.14	< 0.001
Flatness	0.64 ± 0.12	0.58 ± 0.12	< 0.001

Abbreviations: *ICA* internal carotid artery, *MCA* middle cerebral artery, *ACA* anterior cerebral artery, *PC* posterior circulation, *AcomA* anterior communication artery, *AR* aspect ratio, *SR* size ratio, *UI* undulation index, *NSI* nonsphericity index, *VNR* volume-to-neck ratio

Similarly, the AUCs of the LR, GBDT, and Transformer models were 0.823 (95% CI, 0.774 to 0.872), 0.842 (95% CI, 0.798 to 0.886), and 0.876 (95% CI, 0.834 to 0.918) in the external validation datasets, respectively. The Transformer model performed better than the other two models in the validation cohorts. Additionally, the accuracy, sensitivity, and specificity of each model in the training, internal validation, and external validation datasets are displayed in Fig. 3d–f. Finally, Supplementary Table 3 shows the feature ranks of the top 5 variants derived by the RF, Adaboost, and GBDT algorithms.

Discussion

In this study, we developed a model for UIA rupture risk prediction using a DL algorithm based on the clinical, morphological, and radiomics features of 3D-DSA in a large Chinese cohort and validated it using internal and external datasets. The radiomics features of the traditional LR model outperformed models using clinical and morphological aneurysm features. Of the combined signature models, the comprehensive model (model F), which integrated all signatures, performed best, followed by model E, which added radiomics features to the clinical features. Therefore, we believe that radiomics parameters are one of the leading features that can predict the risk of aneurysm rupture. Additionally, we also found the DL model integrating clinical, morphological, and radiomics characteristics performed better than the LR and ML models in rupture risk assessments for UIAs. Of them, the DL model performed best, indicating that the DL algorithm can be used for optimal patient management.

Assessing the risk of UIA rupture is a controversial topic due to the heterogeneity of existing studies. Clinically, it is believed that several risk factors are closely associated with the rupture of UIAs [15]. For example, the PHASES score, which is the most commonly used scoring system to assess the risk of UIA rupture, was developed based on a pooled analysis of six prospective cohort studies using data from 8382 patients and 10,272 aneurysms [16]. In the PHASES score, geographic origin, age, presence or absence of hypertension, aneurysm size and location, and earlier SAH history were identified as relevant parameters for assessing the rupture risk of UIA. However, due to the incomplete collection of basic research data, the PHASES score does not account for other morphological factors such as irregular aneurysm shape and UI. Therefore, further improvements are needed before it can be widely used in clinical practice [17]. Similarly, our study showed that only the clinical characteristics model has poor prediction ability (AUC = 0.706). This suggests that clinical features alone are insufficient to predict UIA rupture risk.

In some studies, morphological characteristics were found to be an important factor for discriminating aneurysm stability [18, 19]. Size, UI, AR, flow angle, irregularity, height/width ratio, and locations are widely used morphological indexes for aneurysm stratification [20]. A similar result was obtained in our study and demonstrated that aneurysm length, diameter-to-width ratio, UI, irregularity, and location were the five most important morphological variables for identifying the rupture risk of UIAs. The B model (AUC = 0.685), which only used these morphological signatures, was not able to predict UIA rupture risk. However, adding traditional clinical markers to the morphological characteristics model significantly improved the discrimination power of the model (AUC = 0.786). These results further confirmed that morphological features were important features for predicting UIA rupture risk and can improve assessment performance when combined with clinical features.

Radiomics has been successfully used to better understand intracranial aneurysms [8, 19]. However, adding

Table 2 Clinical, morphological, and radiomics characteristics of the training and validation datasets

Variables	Training cohort ($n=1157$)	Internal validation cohort ($n=294$)	External validation cohort (=358)	p value
Clinical characteristics				
Age	55.23 ± 10.13	55.40 ± 9.73	56.84 ± 10.52	0.030
Female	763 (65.9%)	210 (71.4%)	239 (66.8%)	0.202
Smoking	258 (22.3%)	63 (21.4%)	90 (25.1%)	0.452
Alcohol abuse	246 (21.3%)	53 (18.0%)	88 (24.6%)	0.125
Hypertension	594 (51.3%)	149 (50.7%)	205 (57.3%)	0.119
Hyperlipidemia	122 (10.5%)	34 (116%)	33 (9.2%)	0.612
Coronary heart disease	88 (7.6%)	14 (4.8%)	21 (5.9%)	0.165
Diabetes mellitus	131 (11.3%)	22 (7.5%)	39 (10.9%)	0.159
Presence of symptom	667 (57.6%)	170 (57.8%)	210 (58.7%)	0.944
Family history	100 (8.6%)	27 (9.2%)	25 (7.0%)	0.534
Multiplicity	444 (38.4%)	105 (35.7%)	135 (37.7%)	0.856
Rupture	272 (23.5%)	80 (27.2%)	85 (23.7%)	0.408
Morphological characteristics				
Location				0.238
ICA	719 (62.2%)	192 (65.3%)	219 (61.2%)	
MCA	134 (11.5%)	27 (9.2%)	52 (14.5%)	
ACA	52 (4.5%)	12 (4.1%)	15 (4.2%)	
PC	75 (6.5%)	22 (7.5%)	30 (8.4%)	
AcomA	177 (15.3%)	41 (13.9%)	42 (11.7%)	
Aneurysm length diameter	4.65 ± 2.30	4.85 ± 2.50	4.59 ± 2.35	0.329
Aneurysm width	4.98 ± 2.06	5.12 ± 1.97	5.03 ± 2.29	0.564
Aneurysm height	4.08 ± 2.42	4.24 ± 2.26	4.04 ± 2.20	0.479
Aneurysm transverse diameter	5.18 ± 2.20	5.33 ± 2.08	5.26 ± 2.32	0.560
Aneurysm maximum diameter	6.35 ± 2.63	6.63 ± 2.82	6.31 ± 2.74	0.229
Aneurysm volume	76.70 ± 146.88	82.03 ± 153.42	83.79 ± 168.63	0.693
Aneurysm neck diameter	4.55 ± 1.71	4.68 ± 1.65	4.59 ± 1.80	0.503
Flow angle	103.60 ± 32.87	103.49 ± 35.35	103.16 ± 32.19	0.976
Aneurysm length diameter-to-width ratio	0.93 ± 0.23	0.94 ± 0.26	0.91 ± 0.20	0.204
Aneurysm neck area	18.53 ± 20.42	19.30 ± 14.46	19.03 ± 16.71	0.781
Aneurysm transverse diameter-to-neck diameter ratio	1.15 ± 0.26	1.15 ± 0.26	1.16 ± 0.30	0.799
Aneurysm angle	52.11 ± 20.41	51.78 ± 20.34	51.22 ± 20.89	0.767
Proximal diameter of parent artery	3.05 ± 0.87	3.04 ± 0.80	3.04 ± 0.90	0.996
Distal diameter of parent artery	2.80 ± 0.85	2.85 ± 0.79	2.83 ± 0.87	0.625
Average diameter of parent artery	3.06 ± 0.82	3.08 ± 0.73	3.06 ± 0.85	0.902
Length diameter of parent vessel	13.10 ± 3.77	13.30 ± 3.51	12.95 ± 3.54	0.489
AR	1.05 ± 0.45	1.07 ± 0.55	1.02 ± 0.44	0.514
SR	1.69 ± 1.01	1.75 ± 1.19	1.65 ± 0.93	0.451
UI	0.17 ± 0.06	0.16 ± 0.06	0.16 ± 0.06	0.457
NSI	0.37 ± 0.04	0.37 ± 0.04	0.36 ± 0.04	0.260
VNR	4.59 ± 5.16	4.97 ± 6.64	4.59 ± 5.10	0.549
Irregularity	254 (22.0%)	67 (22.8%)	74 (20.7%)	0.798
Bifurcation	498 (34.0%)	122 (41.5%)	140 (41.3%)	0.796
Radiomics characteristics				
Mesh volume	189.2. ± 569.75	225.49 ± 448.58	237.19 ± 794.33	0.342
Voxel volume	189.7. ± 569.89	226.24 ± 448.88	237.66 ± 794.64	0.342
Surface area	156.44 ± 206.44	177.34 ± 213.35	166.97 ± 276.39	0.319
Surface area-to-volume ratio	1.76 ± 0.95	1.64 ± 0.77	1.79 ± 0.91	0.049
Sphericity	0.79 ± 0.06	0.79 ± 0.05	0.78 ± 0.06	0.814

Table 2 (continued)

Variables	Training cohort ($n=1157$)	Internal validation cohort ($n=294$)	External validation cohort ($=358$)	p value
Compactness 1	0.037 ± 0.004	0.036 ± 0.004	0.037 ± 0.004	0.830
Compactness 2	0.50 ± 0.10	0.49 ± 0.10	0.50 ± 0.10	0.865
Spherical disproportion	1.28 ± 0.11	1.27 ± 0.08	1.28 ± 0.11	0.724
Maximum 3D diameter	7.87 ± 4.15	8.41 ± 4.58	7.84 ± 4.63	0.137
Maximum 2D diameter (slice)	6.89 ± 3.67	7.40 ± 4.07	6.90 ± 4.25	0.020
Maximum 2D diameter (column)	7.00 ± 3.75	7.42 ± 3.95	6.91 ± 3.95	0.059
Maximum 2D diameter (row)	7.03 ± 3.70	7.32 ± 3.91	7.03 ± 4.04	0.494
Major axis length	6.66 ± 3.61	7.14 ± 3.98	6.57 ± 3.93	0.008
Minor axis length	4.96 ± 2.63	5.23 ± 2.76	5.01 ± 3.04	0.322
Least axis length	4.10 ± 2.28	4.35 ± 2.48	4.14 ± 2.62	0.281
Elongation	0.76 ± 0.12	0.76 ± 0.13	0.77 ± 0.12	0.352
Flatness	0.63 ± 0.13	0.62 ± 0.13	0.63 ± 0.12	0.645

Abbreviations: *ICA* internal carotid artery, *MCA* middle cerebral artery, *ACA* anterior cerebral artery, *PC* posterior circulation, *AcomA* anterior communication artery, *AR* aspect ratio, *SR* size ratio, *UI* undulation index, *NSI* nonsphericity index, *VNR* volume-to neck ratio

radiomics to differential models assessing aneurysm rupture remains controversial. Calvin et al found that radiomics features extracted from angiography images provided no additional value when identifying rupture risk [21]. However, Liu et al used radiomics-derived morphological features to identify aneurysm stability and described flatness as the most important predictor [8]. Similarly, our study showed that the radiomics features were better at predicting the rupture risk of UIAs when only combined radiomics signatures than models of other features. Compared to clinical and morphological features, radiomics features showed the best prediction power (AUC, 0.708 vs. 0.678 vs. 0.738, all $p<0.05$). Furthermore, the combined model E, which added radiomics features to the clinical features, showed better performance relative to the combined model D, which added morphological features (AUC, 0.842 vs. 0.771, $p<0.05$). Therefore, our study suggests that radiomics plays an important role in predicting the rupture risk of UIAs and highlights the potential clinical utility of radiomics analysis.

Artificial intelligence algorithms are a novel method that could have advantages over traditional statistical methods when addressing certain clinical problems [22, 23]. Previous machine learning studies that have classified the stability of an intracranial aneurysm and assessed rupture risk have shown encouraging results [8, 24]. Zhu et al showed that ML models outperformed statistical LR methods and PHASES

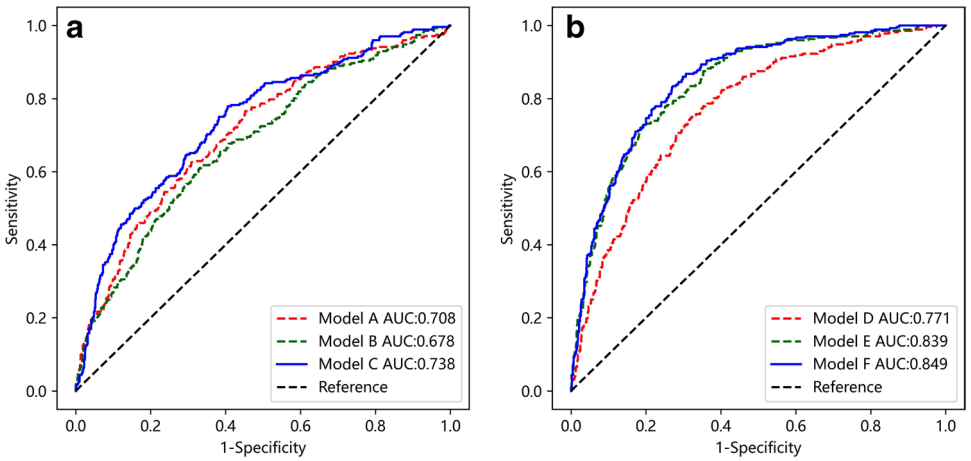

Fig. 3 The ROC curves of six prediction models using conventional LR. **a** The AUCs of the three basic models are based on clinical (model A), morphological (model B), and radiomics (model C) characteristics. The AUCs of models A (red curve), B (green curve), and C (blue curve) were 0.708, 0.678, and 0.738, respectively. **b** The AUCs of the three combined models. The AUC of model D (clinical and morphological) is 0.771 (red curve). The AUC of model E (clinical and radiomics) is 0.839 (green curve). The AUC of model F (clinical, morphological, and radiomics) is 0.849 (blue curve)

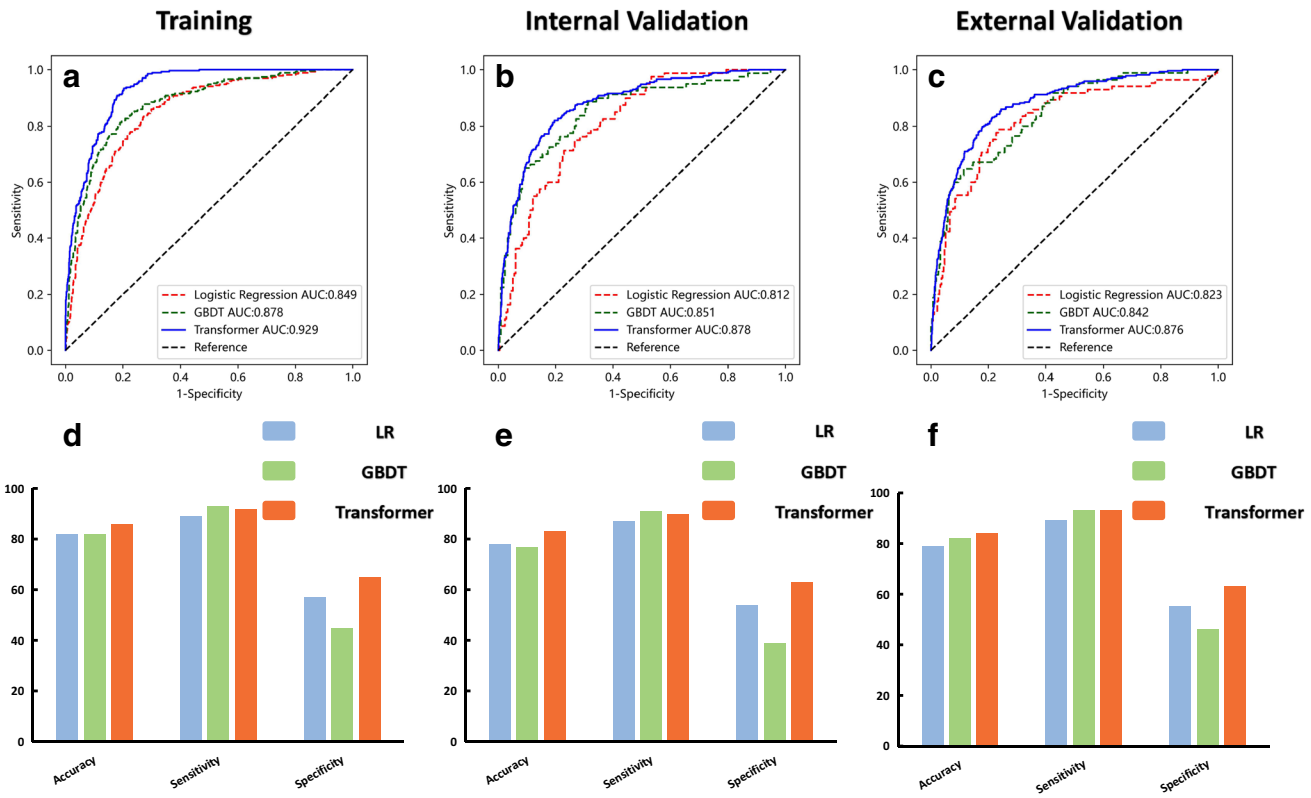

Fig. 4 The ROC curves, sensitivity, specificity, and accuracy of conventional LR, GBDT, and Transformer algorithms. **a**–**c** The AUCs of the LR, GBDT, and Transformer algorithms in the training, internal validation, and external validation datasets. **d**–**f** The sensitivity, specificity, and accuracy of the LR, GBDT, and Transformer algorithms in the training, internal validation, and external validation datasets. The accuracies of the LR, GBDT, and Transformer algorithms in the training, internal validation, and external validation datasets are 86%, 83%, and 84% (blue bar); 87%, 84%, and 83% (green bar); and 89%, 82%, and 81% (orange bar), respectively; the sensitivity of the LR, GBDT, and Transformer algorithms in the training, internal validation, and external validation datasets are 91.6%, 90.4%, and 92.9% (blue bar); 91.0%, 89.0%, and 87.8% (green bar); and 92.6%, 90.6%, and 89.8% (orange bar), respectively; The specificities of the LR, GBDT, and Transformer algorithms in the training, internal validation, and external validation datasets are 65.3%, 63.1%, and 62.7% (blue bar); 72.2%, 70.5%, and 71.3% (green bar); and 75.8%, 73.5%, and 73.3% (orange bar), respectively

scores in IA stability assessment [20]. Kim et al showed that the application of a convolutional neural network (CNN) based on 3D-DSA imaging was able to distinguish the ruptured state of small aneurysms [25]. However, their primary limitation is the lack of external validation of the models, which improves model confidence and is needed before the models can be used in clinical settings. In this study, an external validation dataset from another hospital was used to test the three models combining clinical, morphologic, and radiomics characteristics, and the AUCs of the three datasets ranged from 0.823 to 0.876. Therefore, the results from external validation data help increase the credibility of our model and further solidify its application in clinical practice, and the positive results further demonstrate the robustness of the DL model.

To our knowledge, no study has assessed the performance of the DL algorithm using an aneurysm rupture risk model based on clinical, morphological, and radiological parameters. A preliminary study by Ou et al found that combining morphological and radiomics features further improved discrimination performance [26]. Another study found that combining radiomics with traditional clinical and morphological models can provide additional benefits for staging ruptured middle cerebral aneurysms (MCA) [27]. However, all of the models developed in these studies were created using simple statistical methods such as LR and LASSO regression. In our study, the Transformer model was significantly better at predicting the rupture risk of UIAs as the ML algorithm and conventional LR based on clinical, morphological, and radiomics characteristics (AUC, 0.929 vs. 0.878 and 0.849, all $p < 0.001$, respectively). Therefore, our results indicate that DL algorithms have the potential to identify rupture risk and contribute to better clinical management of UIAs.

This study has several strengths. First, we visualized all intracranial aneurysms using 3D-DSA as a reference standard based on a large sample and validated the model with internal and external validation datasets. Second, radiomics

signatures were used to predict ruptured aneurysms. Our results further confirm the important role of radiomics in prediction modeling. Third, our comparison of the three feature models and another three combined signature models using LR outlines the role of each factor when predicting the rupture risk models. Fourth, our results show that DL algorithms based on clinical, morphologic, and radiologic features are significantly better at identifying the risk of UIA rupture compared to the traditional LR method.

Limitations

This study has certain limitations that must be acknowledged. The first is selection bias, which is typical in retrospective cohort studies. About 20% of patients (Hunt-Hess grade 5) with ruptured aneurysms cannot reach the hospital for treatment or surgery in time. The loss of data on these patients caused data bias and underestimated the risk of rupture. Second, although a previous study reported that intracranial aneurysm morphology does not change dramatically between pre- and post-rupture, the event of rupture itself may have affected the morphology of intracranial aneurysm; this may generate a possible bias on our results. However, a follow-up of all UIAs would be impractical and unethical because high-risk UIAs (irregular, daughter sac, and multilobed structure) have a high rupture risk; once ruptured, the outcome is very poor. Therefore, a large, prospective, multicenter study is needed to further validate our results.

Conclusions

This study shows that radiomics parameters play an important role in assessing the rupture risk of UIAs. In addition, integrated functions, including radiomics parameters, can improve the prediction of rupture risk in UIA. The DL algorithm outperforms the traditional LR and ML models when assessing UIA rupture risk in tests using internal and external validation datasets. The DL model has significant potential to support clinical decision-making, and its future application in clinical practice can enable the individualized and optimal management of patients with IAs.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00330-023-09672-3.

Funding This study has received funding from the National Natural Science Foundation of China (grant numbers: 82272092, 82072036), Beijing Municipal Administration of Hospitals Incubating Program (grant number: PX2022022), and Research Projects of National Health Commission Capacity Building and Continuing Education Center in 2021 (grant number: GWJJ2021100103).

Declarations

Guarantor The scientific guarantor of this publication is Jian Liu of Beijing Neurosurgical Institute and Beijing Tiantan Hospital.

Conflict of 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.

Statistics and biometry One of the authors has significant statistical expertise.

Informed consent Written informed consent was obtained from all patients in this study.

Ethical approval and consent to participate The study has been approved by the Ethics Committee of Beijing Tiantan Hospital (Approval No. KY 2020–134-01). This clinical trial will be conducted in accordance with the principles laid down by the 18th World Medical Assembly (Helsinki, 1964) and all applicable amendments laid down by the World Medical Assemblies and the International Conference on Harmonization Guidelines for Good Clinical Practice.

Methodology
- Retrospective
- Cross-sectional study
- Multicenter study

References

1. Li MH, Chen SW, Li YD et al (2013) Prevalence of unruptured cerebral aneurysms in Chinese adults aged 35 to 75 years: a cross-sectional study. Ann Intern Med 159:514–521
2. Hackenberg KAM, Hänggi D, Etminan N (2018) Unruptured intracranial aneurysms. Stroke 49(9):2268–2275
3. Morita A, Kirino T, Hashi K et al (2012) The natural course of unruptured cerebral aneurysms in a Japanese cohort. New Engl J Med 366:2474–2482
4. Korja M, Kivisaari R, Rezai Jahromi B et al (2017) Natural history of ruptured but untreated intracranial aneurysms. Stroke 48:1081–1084
5. Algra AM, Lindgren A, Vergouwen MDI et al (2019) Procedural clinical complications, case-fatality risks, and risk factors in endovascular and neurosurgical treatment of unruptured intracranial aneurysms: a systematic review and meta-analysis. JAMA Neurol 76(3):282–293
6. Shi Z, Hu B, Schoepf UJ et al (2020) Artificial intelligence in the management of intracranial aneurysms: current status and future perspectives. AJNR Am J Neuroradiol 41(3):373–379
7. Lambin P, Rios-Velazquez E, Leijenaar R et al (2012) Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer 48:441–446
8. Liu Q, Jiang P, Jiang Y et al (2019) Discrimination of aneurysm stability using a machine learning model based on PyRadiomics-derived morphological features. Stroke 50(9):2314–2321
9. Zhai X, Geng J, Zhu C et al (2021) Risk factors for pericallosal artery aneurysm rupture based on morphological computer-assisted semiautomated measurement and hemodynamic analysis. Front Neurosci 18(15):759806
10. Geng J, Hu P, Ji Z et al (2020) Accuracy, and reliability of the computer-assisted semi-automated morphological analysis of intracranial aneurysms: an experimental study with digital

phantoms and clinical aneurysm cases. Int J Comput Assist Radiol Surg 15(10):1749–1759
11. Zhang Y, Tian Z, Jing L et al (2016) Bifurcation type and larger low shear area are associated with rupture status of very small intracranial aneurysms. Front Neurol 24(7):169
12. van Griethuysen JJM, Fedorov A, Parmar C et al (2017) Computational radiomics system to decode the radiographic phenotype. Cancer Res 77(21): e104-e107
13. Wang W, Peng Y, Feng X et al (2021) Development and validation of a computed tomography-based radiomics signature to discriminate response to neoadjuvant chemotherapy for locally advanced gastric cancer. JAMA Network Open 4(8): e2121143
14. DeLong ER, DeLong DM, Clarke-Pearson DL (1988) Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 44(3):837–845
15. Thompson BG, Brown RD Jr, Amin-Hanjani S et al (2015) American Heart Association Stroke Council, Council on Cardiovascular and Stroke Nursing, and Council on Epidemiology and Prevention; American Heart Association; American Stroke Association. Guidelines for the management of patients with unruptured intracranial aneurysms: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. Stroke 46(8):2368–400
16. Greving JP, Wermer MJ, Brown RD Jr et al (2014) Development of the PHASES scores for discrimination of risk of rupture of intracranial aneurysms: a pooled analysis of six prospective cohort studies. Lancet Neurol 13(1):59–66
17. Juvela S (2021) PHASES score and treatment scoring with cigarette smoking in the long-term discrimination of rupturing of unruptured intracranial aneurysms. J Neurosurg 136(1):156–162
18. Wermer MJ, van der Schaaf IC, Algra A et al (2007) Risk of rupture of unruptured intracranial aneurysms in relation to patient and aneurysm characteristics: an updated meta-analysis. Stroke 38(4):1404–1410
19. Kashiwazaki D, Kuroda S; Sapporo SAH Study Group (2013) Size ratio can highly discriminate rupture risk in intracranial small (<5 mm) aneurysms. Stroke 44(8):2169-73
20. Zhu W, Li W, Tian Z et al (2020) Stability assessment of intracranial aneurysms using machine learning based on clinical and morphological features. Transl Stroke Res 11(6):1287–1295
21. Ludwig CG, Lauric A, Malek JA et al (2021) Performance of radiomics derived morphological features for discrimination of aneurysm rupture status. J Neurointerv Surg 13(8):755–761
22. Beam AL, Kohane IS (2018) Big data and machine learning in health care. JAMA 319(13):1317–1318
23. Zhang Y, Zhang B, Liang F et al (2019) Radiomics features on non-contrast-enhanced CT scans can precisely classify AVM-related hematomas from other spontaneous intraparenchymal hematoma types. Eur Radiol 29(4):2157–2165
24. Liu J, Chen Y, Lan L et al (2018) Discrimination of rupture risk in anterior communicating artery aneurysms with a feed-forward artificial neural network. Eur Radiol 28(8):3268–3275
25. Kim HC, Rhim JK, Ahn JH et al (2019) Machine learning application for rupture risk assessment in small-sized intracranial aneurysm. J Clin Med 8(5):683
26. Ou C, Chong W, Duan CZ et al (2021) A preliminary investigation of radiomics differences between ruptured and unruptured intracranial aneurysms. Eur Radiol 31(5):2716–2725
27. Zhu D, Chen Y, Zheng K et al (2021) Classifying ruptured middle cerebral artery aneurysms with a machine learning based, radiomics-morphological model: a multicentral study. Front Neurosci 11(15):721268