International Journal of General Medicine

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

Predictions of Small Intracranial Aneurysms' Rupture Risk with Ensemble Machine Learning Model (Super Learner): A Retrospective Study in Two Tertiary Hospitals in China

Xiaolong Hu[1,2,*], Shifei Ye[2,*], Dayong Qi[2], Suya Li[2], Xiaoyu Tang[3], Yibin Fang[1,2,4]

[1]Tongji University School of Medicine, Tongji University Affiliated Shanghai 4th People's Hospital, Shanghai, People's Republic of China; [2]Department of Neurovascular Disease, Tongji University Affiliated Shanghai 4th People's Hospital, Shanghai, People's Republic of China; [3]Department of Neurosurgery, Suzhou Municipal hospital affiliated to Nanjing Medical university, Suzhou, Jiangsu Province, People's Republic of China; [4]Translational Research Institute of Brain and Brain-Like Intelligence, Shanghai Fourth People's Hospital, School of Medicine, Tongji University, Shanghai, People's Republic of China

*These authors contributed equally to this work

Correspondence: Xiaoyu Tang, Department of neurosurgery, Suzhou municipal hospital affiliated to Nanjing medical university, Suzhou, Jiangsu Province, People's Republic of China, Email xiaoyutang@njmu.edu.cn; Yibin Fang, Tongji University School of Medicine, Tongji University Affiliated Shanghai 4th People's Hospital, Shanghai, People's Republic of China, Email fangyibin@tongji.edu.cn

Purpose: This research aims to investigate the morphological, clinical and hemodynamic parameters influencing intracranial aneurysm rupture, develop a ensemble machine learning model (Super Learner) to predict its rupture risk.

Methods: This retrospective study analyzed aneurysm patients from two hospitals. Five base learners—decision tree (DT), k-nearest neighbor (KNN), random forest (RF), support vector machine (SVM) and extreme gradient boosting (XGBoost)—were constructed based on demographic, hemodynamic and geometric parameters. A Super Learner model was subsequently constructed using ensemble learning algorithms, with all models validated on an independent external dataset.

Results: The dataset comprised 97 patients in the training cohort, 42 in the internal validation cohort, and 86 in the external validation cohort. In the external validation cohort, the area under the curve (AUC) values: Super learner 0.94 (0.89–1.00), Random Forest 0.83 (0.76–0.91), XGBoost 0.93 (0.87–0.99), Support Vector Machine 0.82 (0.73–0.92), Decision Tree 0.84 (0.76–0.93), and k-Nearest Neighbors 0.51 (0.38–0.63).

Conclusion: The Super Learner model outperforms individual models in both performance and stability for predicting intracranial aneurysm rupture risk. It has been robustly validated on an external dataset, demonstrating strong generalizability.

Keywords: super learner, machine learning, aneurysm, predictive model

Introduction

Intracranial aneurysms (IAs) affect approximately 2–3% of the general population.[1] When ruptured, they can result in subarachnoid hemorrhage (SAH), a devastating condition associated with high rates of mortality[2] and long-term disability.[3] Timely identification of aneurysms at high risk of rupture is therefore critical for guiding treatment decisions and improving patient outcomes.

Aneurysm size is one of the most widely studied predictors of rupture risk,[4,5] with smaller aneurysms generally considered less likely to rupture. For instance, rupture rates as low as 0%, <0.5%, and 1% have been reported for aneurysms measuring 1 mm, 3 mm, and 7 mm, respectively.[6] However, clinical experience and retrospective studies have demonstrated that some small aneurysms do rupture, often with catastrophic consequences.[7,8] This highlights a key limitation of relying on aneurysm size alone for risk stratification.

Received: 21 May 2025
Accepted: 20 August 2025
Published: 3 November 2025

To address this limitation, several clinical risk assessment tools—such as the PHASES score—have been developed,[9] incorporating factors like aneurysm size, location, and patient history. While useful, these tools lack the granularity needed for individualized risk prediction and often fail to account for complex morphological and hemodynamic features.

Advancements in computational fluid dynamics (CFD) have enabled the quantification of hemodynamic parameters within aneurysms, such as wall shear stress and flow instability.[10,11] These features offer valuable insight into the biological processes underlying aneurysm rupture, but their integration into clinical practice remains limited, partly due to a lack of validated predictive models.

In recent years, machine learning (ML) has emerged as a powerful tool for medical prediction tasks, capable of uncovering nonlinear patterns and high-dimensional relationships from complex datasets.[12,13] Prior studies have demonstrated the potential of ML-based models to outperform conventional statistical approaches in predicting aneurysm rupture.[14] However, most existing ML models rely on single algorithms, which may suffer from inconsistent performance across datasets and lack robustness in clinical application.

To address this gap, we propose a Super Learner ensemble model that integrates multiple machine learning algorithms using cross-validation to optimize prediction performance.[15–17] This approach aims to leverage the strengths of various base learners while mitigating individual model limitations. Furthermore, we emphasize external validation and model interpretability to ensure both robustness and clinical applicability.

In this study, we developed and externally validated a Super Learner-based model for predicting rupture risk in intracranial aneurysms, incorporating clinical, morphological, and hemodynamic features. The model's interpretability was further assessed to enhance its translational potential in clinical decision-making.

Methods

Patients Selection

This retrospective study was approved by the Ethics Committee of Shanghai Fourth People's Hospital and the Institutional Review Board of Suzhou Municipal Hospital. Since the study used anonymized medical records and did not involve direct patient interaction, the need for informed consent was waived by the ethics committees. All data were handled confidentially, and the study complied with the ethical principles of the Declaration of Helsinki. This study retrospectively collected data from all intracranial aneurysm patients treated at the stroke centers of two hospitals between December 2021 and December 2023.

Study participants were selected according to predefined inclusion criteria and excluded based on specified exclusion parameters. Inclusion criteria: (1) The presence of intracranial aneurysms was confirmed through either digital subtraction angiography (DSA) or computed tomography angiography (CTA); (2) a saccular aneurysm. Exclusion criteria: (1) Traumatic, infectious, thrombotic, or dissecting aneurysms; (2) Saccular aneurysms associated with arteriovenous malformations. (3) Incomplete or poor-quality imaging that prevents aneurysm reconstruction and hemodynamic analysis. (4) H max > 7 mm. The study flowchart is shown in Figure 1. In cases of multiple aneurysms, the lesion with the maximum diameter was selected as the index aneurysm for classification purposes. Rupture status of intracranial aneurysms was determined retrospectively from inpatient medical records. Aneurysms were classified as ruptured if imaging (DSA or CTA) confirmed the presence of an intracranial aneurysm accompanied by subarachnoid or intracerebral hemorrhage. Small aneurysms were defined as those with Hmax ≤ 7 mm.

CFD Analysis

Patient-specific vascular geometries were generated through 3D reconstruction of clinical DICOM datasets acquired via rotational angiography (3D-DSA). Subsequent mesh optimization, including anatomical refinement and surface regularization, was performed using Geomagic Wrap 2015 (3D Systems, NC, USA) as illustrated in Figure 2. Minor branches were removed during this process, while aneurysms, parent vessels, and key branching vessels were preserved.

The reconstructed model was used for hemodynamic simulations. The reconstructed geometry was processed in ANSYS ICEM CFD 16.2 (ANSYS Inc., PA) for computational mesh generation, employing a uniform element size of 0.16 mm throughout the domain.

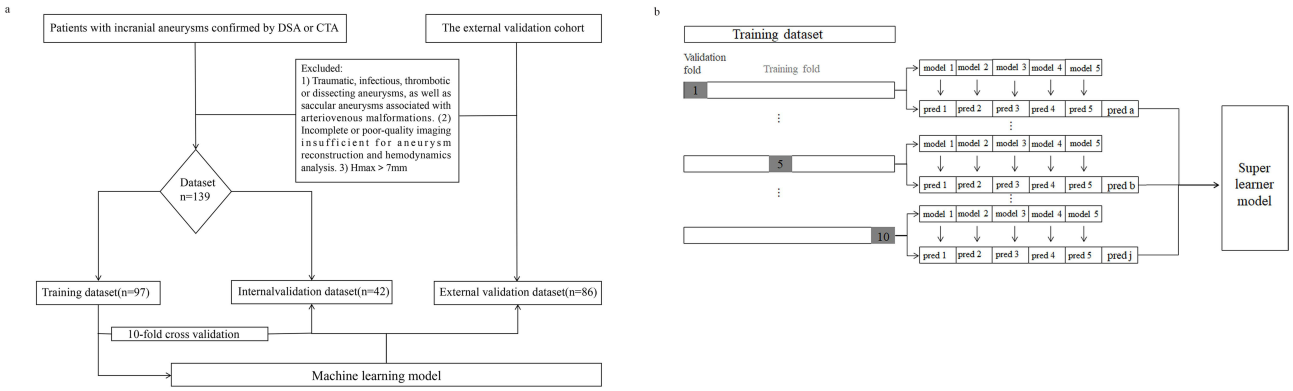

Figure 1 The flowchart of this study (**a**). Super learner algorithm construction illustrated with 10-fold cross-validation and 5 base models (SVM, KNN, DT, XGBOOST, RF) (**b**).

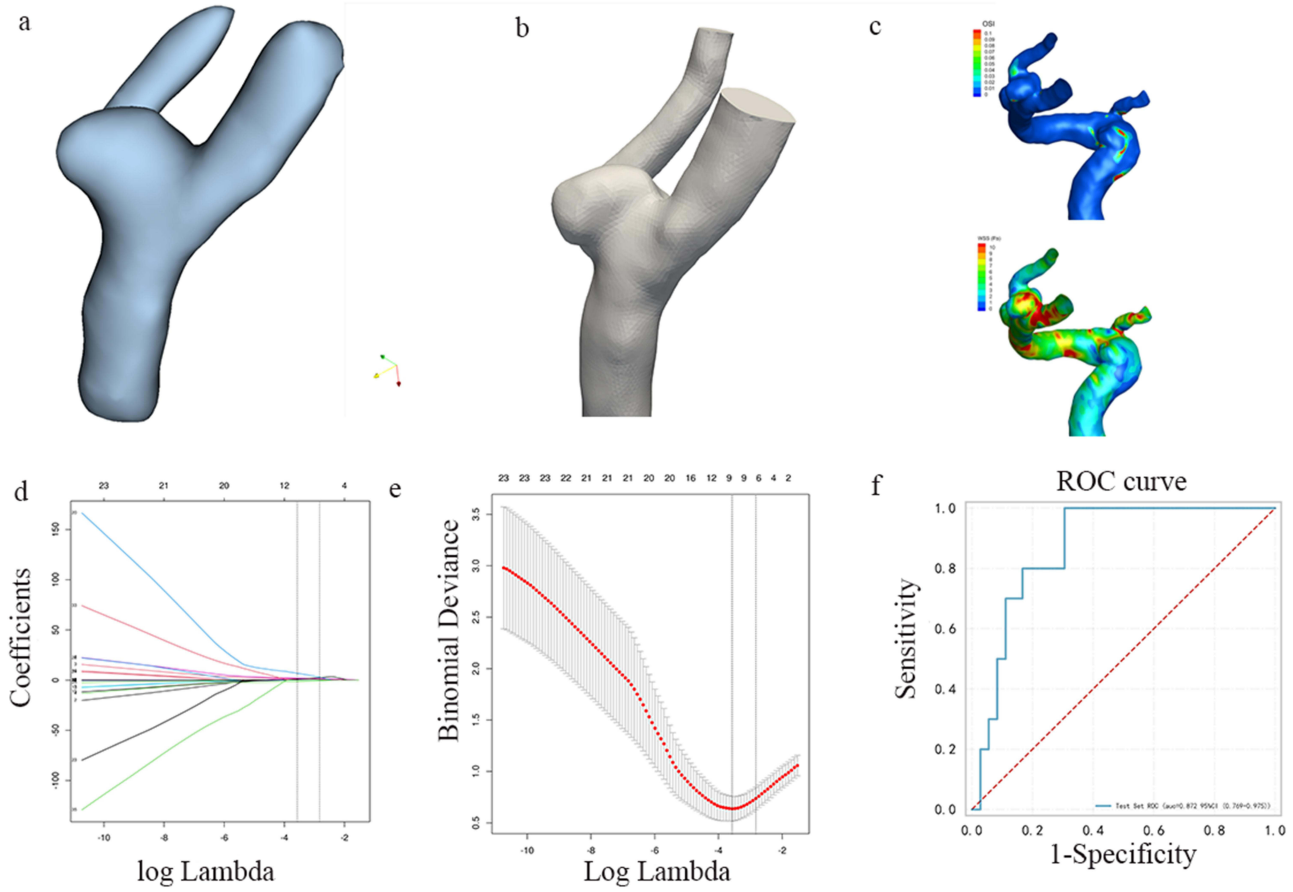

Figure 2 The flow chart of the model's construction and validation (**a**) The 3D aneurysm's model (**b**) The computational fluid dynamics model of the aneurysm (**c**) The streamline chart of aneurysm (**d**) The cross validation of features, CV (**e**) Least Absolute Shrinkage and Selection Operator algorithm was used to select features. (**f**) Receiver operating characteristic curve of machine learning models.

Numerical simulations were performed in ANSYS CFX 2019 (ANSYS Inc., PA) by using the incompressible Navier-Stokes equations, with the computational setup illustrated in Figure 2C. The blood flow was simulated as an incompressible Newtonian fluid (ρ=1056 kg/m^3, μ=0.0035 Pa·s) under laminar regime conditions. Vascular walls were treated as rigid boundaries enforcing no-slip conditions. The inlet boundary condition specified a pulsatile flow rate of 4.6 mL/s,

https://doi.org/10.2147/IJGM.S533558

while outlet flow distribution followed Murray's hemodynamic principle.[18] The twelve quantitative hemodynamic parameters analyzed in this study are detailed in the supplementary materials.

Morphological Parameters Measurement

Diagnostic evaluation included digital subtraction angiography for all study participants, with three-dimensional vascular reconstructions created from rotational angiography data using Syngo Workplace post-processing software (Siemens Healthineers, Germany). The post-processing algorithm converts DICOM-formatted axial DSA images into three-dimensional vascular models exported in stereolithography (STL) file format for computational analysis. Two board-certified neuroradiologists independently quantified the aneurysm and adjacent vascular geometry using dedicated workstation measurement tools. Both neuroradiologists were blinded to the patients' clinical manifestations and CT findings. Sixteen morphological parameters were analyzed, including inflow angle, tilt angle, vascular angle, HMAX, DMIDDLE, DNECK, DVESSEL, H, S, V, AR, UI, SR, EI, NSI, BNR, and irregular shape. The definitions and measurements of these parameters are provided in Supplementary figure 1.

Construction and Validation of Machine Learning Model

Data from the Cerebrovascular Disease Center of Shanghai Fourth People's Hospital was randomly split into a training set (n=97) and an internal validation cohort (n=42) in a 7:3 ratio. The Least Absolute Shrinkage and Selection Operator (LASSO) algorithm, combined with 10-fold cross-validation, was used to select modeling variables in the training cohort (Figure 1). Collinearity analysis was performed on the selected variables. Variables with a collinearity value >5 were removed as appropriate based on practical considerations (see supplementary materials for details). Based on the final selected modeling variables, five machine learning models and one Super Learner model were built based on the training cohort data. The models included SVM, RF, XGBoost, KNN, DT, and the Super Learner model. The models' effectiveness was validated using the internal validation cohort and an external validation cohort (N=86) (Figure 1). The models' performance was evaluated by the receiver operating characteristic (ROC) curve, area under the curve (AUC), sensitivity, specificity, accuracy, F1 score, positive predictive value (PPV), and negative predictive value (NPV). To ensure model diversity and reduce redundancy among base learners, we evaluated each algorithm's contribution to the ensemble. Based on cross-validated weights, models with zero coefficients (e.g, XGBoost, KNN) were removed from the final Super Learner ensemble, while models with distinct decision mechanisms and meaningful contribution (eg, random forest and decision tree) were retained.

Missing Data Handling

No missing data were encountered in this study. Imaging data were subject to strict quality control, and cases with incomplete or poor-quality scans were excluded prior to analysis. Clinical variables were retrospectively collected from comprehensive medical records, ensuring completeness. As such, no imputation or exclusion for missing values was required during model development.

Statistical Analysis

Statistical analyses were conducted using IBM SPSS Statistics software (v25.0, Armonk, NY) for descriptive statistics and R programming language (v3.4.3, R Foundation) for advanced predictive modeling.

Data distribution normality was evaluated using the Kolmogorov–Smirnov goodness-of-fit test. Continuous variables following normal distribution were summarized using parametric measures (mean ± SD), whereas non-parametric descriptors (median [IQR]) were employed for skewed distributions.

Nominal variables were summarized using absolute counts and relative frequencies (%). Between-group differences in categorical data were assessed using either Pearson's χ^2-test or likelihood ratio tests, selected based on expected cell frequencies. For parametric data, intergroup comparisons employed independent samples t-tests (two groups) or one-way ANOVA (≥3 groups) with post-hoc Tukey's testing. If ANOVA revealed significant differences, post hoc analysis was conducted for pairwise comparisons. Non-normally distributed data were analyzed using the independent sample non-parametric test. All statistical tests were two-sided, with a significance level set at $p < 0.05$. Multicollinearity

among the modeling variables was evaluated using the Variance Inflation Factor (VIF), with additional information provided in the supplementary materials. The Super Learner model in this study was developed using an ensemble learning algorithm that integrated multiple weighted base models (SVM, RF, XGBoost, DT, and KNN). Cross-validation was applied to select the optimal ensemble model. The hyperparameters of the five base models are shown in Supplementary Table 1. Super Learner is an ensemble-based algorithm that improves predictive performance by optimally weighting multiple base models. It determines the best model combination strategy through cross-validation. In this study, the Super Learner model incorporated ten-fold cross-validation and five base models (SVM, KNN, DT, XGBoost, and RF). The training dataset was randomly divided into 10 folds using K-fold cross-validation (K = 10). Each base model was trained on K−1 folds and tested on the remaining fold (Figure 1). The predictions from all base models on the test folds were collected, forming a new dataset in which each sample corresponded to multiple base model predictions. The prediction errors of each base model were then calculated. A non-negative least squares algorithm was used to learn the optimal combination weights while ensuring non-negative constraints. The optimal weighted averaging scheme was determined to minimize prediction error (see supplementary Table 2 for details on model parameters and weights).

Super Learner is inherently a black-box model, making direct interpretation challenging. To improve interpretability, we employed the following approach. Each quantitative modeling variable was standardized, and based on the predicted probability of aneurysm rupture in the validation dataset, patients were stratified into four risk groups using interquartile ranges: low risk (0–Q1), low-to-moderate risk (Q1–Q2), moderate-to-high risk (Q2–Q3), and high risk (Q3–Q4). The standardized mean of each quantitative variable was calculated within each risk group to illustrate its trend across different risk levels. For categorical variables, the proportion within each risk group was computed to examine their distribution across different risk levels. This approach provides insight into how individual variables contribute to aneurysm rupture risk prediction and enhances the interpretability of the Super Learner model.

Results

Demographics, Aneurysm, Hemodynamics Parameters

A total of 225 patients were included in this study and were divided into a training cohort (n=97), an internal validation cohort (n=42), and an external validation cohort (n=86). Table 1 presents the baseline characteristics of patients in the training cohort, comparing the ruptured and unruptured aneurysm groups. Significant differences were observed in Hypertension, Irregular shape, H max, H, S, V, AR, UI, SR, EI, NSI, WSSmin, OSI, LSA, and OSI max between the two groups (all $P < 0.05$). Table 2 provides detailed characteristics of patients, aneurysms, and hemodynamic parameters across the training, internal validation, and external validation cohorts. The training cohort consisted of 97 patients (ruptured: n=21, unruptured: n=76), the internal validation cohort included 42 patients (ruptured: n=13, unruptured: n=29), and the external validation cohort comprised 86 patients (ruptured: n=36, unruptured: n=50). Significant differences were observed in certain aneurysm, patient, and hemodynamic parameters among the three cohorts ($P < 0.05$).

Construction and Validation of mL Model

Figure 3a illustrates the ROC curves of six models (Super learner, Random Forest, XGBOOST, SVM, DT, KNN) in the internal validation cohort. The AUC values for Super learner, RF, XGBOOST, SVM, DT, and KNN were 0.86 (0.72–1), 0.73 (0.59–0.88), 0.89 (0.76–1), 0.89 (0.76–1), 0.89 (0.77–1), and 0.53 (0.32–0.74), respectively. Figure 3b presents the ROC curves of these 6 models in the external validation cohort. The AUC values for Super learner, RF, XGBOOST, SVM, DT, and KNN were 0.94 (0.89–1), 0.83 (0.76–0.91), 0.93 (0.87–0.99), 0.82 (0.73–0.92), 0.84 (0.76–0.93), and 0.51 (0.38–0.63), respectively. The Super learner model demonstrated overall superior performance compared to the other five base models. Table 3 summarizes the AUC (95% CI), sensitivity, specificity, accuracy, F1 score, PPV, and NPV for each model in the training, internal validation, and external validation cohorts.

Table 1 Characteristics of Demography, Aneurysms and Hemodynamic Between Ruptured and Unruptured Patients in the Training Cohort, No.(%)/M[IQR]

Characteristics	No.(%)/M[IQR]	Unruptured Group (n=76)	Ruptured Group (n=21)	P value
Patients characteristics				
Age(years)				0.973
<70	69(71.1%)	54(71.1%)	15(71.4%)	
≥70	28(28.9%)	22(28.9%)	6(28.6%)	
Gender				0.538
Male	36(37.1%)	27(35.5%)	9(42.9%)	
Female	61(62.9%)	49(64.5%)	12(57.1%)	
Hypertension	35(36.1%)	18(23.7%)	17(81%)	<0.001
Diabetes	50(51.5%)	43(56.6%)	7(33.3%)	0.059
Smoking	14(14.4%)	13(17.1%)	1(4.8%)	0.283
Drinking	14(14.4%)	13(17.1%)	1(4.8%)	0.283
Irregular shape	36(37.1%)	18(23.7%)	18(85.7%)	<0.001
Inflow angle	95.10±35.09	91.28±35.83	108.91±28.97	0.246
Tilt angle	86.36(53.97,121.33)	92.34(29.71,118.40)	82.44(64.27,127.14)	0.898
Vascular angle	17.85(8.41,35.10)	17.14(8.08,34.24)	24.34(16.33,38.18)	0.177
Hmax	2.82(2.22,4.16)	2.70(2.19,3.40)	4.28(2.97,5.76)	0.001
Dmiddle	3.56(2.99,4.48)	3.48(2.97,4.40)	4.06(3.29,4.47)	0.116
Dneck	3.55(2.95,4.41)	3.52(3.00,4.40)	3.79(2.96,4.20)	0.500
Dvessel	3.15±0.75	3.20±0.74	2.98±0.82	0.538
H	2.40(1.88,3.44)	2.30(1.87,2.88)	3.32(2.24,4.64)	0.008
S	24.63(16.85,44.24)	22.62(15.77,34.72)	54.01(23.11,61.79)	0.008
V	13.96(8.22,30.69)	12.06(7.44,23.78)	32.48(11.91,46.64)	0.013
AR	0.64(0.50,0.87)	0.59(0.49,0.74)	0.85(0.74,0.98)	0.005
UI	0.08(0.05,0.11)	0.07(0.05,0.10)	0.11(0.09,0.15)	0.003
SR	0.90(0.72,1.33)	0.81(0.72,1.11)	1.51(1.04,2.11)	0.001
EI	0.07(0.03,0.12)	0.05(0.03,0.10)	0.12(0.07,0.18)	<0.001
NSI	0.09(0.05,0.14)	0.07(0.04,0.13)	0.13(0.11,0.23)	<0.001
BNR	0.99(0.98,0.99)	0.98(0.98,0.99)	0.99(0.98,1.05)	0.052
TAWSS	7.91(4.61,15.07)	8.93(4.65,17.75)	6.30(4.10,8.70)	0.068
NWSS	0.87(0.62,1.23)	0.89(0.63,1.22)	0.76(0.60,1.18)	0.344
TAWSSv	9.62(6.03,14.67)	10.15(5.95,15.37)	9.01(6.34,12.82)	0.425
WSSmax	29.11(18.63,43.93)	29.08(19.10,43.66)	30.62(18.19,44.07)	0.902
WSSmin	1.00(0.45,1.76)	1.11(0.66,2.47)	0.58(0.15,0.97)	0.001

(*Continued*)

Table 1 (Continued).

Characteristics	No.(%)/M[IQR]	Unruptured Group (n=76)	Ruptured Group (n=21)	P value
HAS	41.24(20.42,73.63)	42.91(20.46,75.03)	31.71(22.20,59.44)	0.180
LSA	0(0,1.48)	0(0,1.09)	0.83(0.11,12.51)	0.004
OSI	0.01(0,0.01)	0.01(0,0.01)	0.01(0.01,0.01)	0.027
OSImax	0.13(0.07,0.22)	0.11(0.07,0.21)	0.20(0.17,0.24)	0.004
WSSG	10.02(5.97,19.99)	10.22(6.11,20.88)	9.03(6.41,12.13)	0.164
GON	0.06(0.04,0.09)	0.06(0.04,0.08)	0.07(0.05,0.09)	0.186
RRT	0.21(0.10,0.40)	0.18(0.09,0.37)	0.28(0.20,0.63)	0.011

Table 2 Characteristics of Patients, Aneurysms, and Hemodynamics in the Training and Validation Datasets

Characteristics	No.(%)/M[IQR]	Training Cohort (n=97)	Internal Validation (n=42)	External Validation Cohort (n=86)	P value
Patients characteristics					
Age(years)					0.035
<70	154(68.4%)	69(71.1%)	34(81%)	51(59.3%)	
≥70	71(31.6%)	28(28.9%)	8(19%)	35(40.7%)	
Gender					<0.01
Male	102(45.3%)	36(37.1%)	13(31%)	53(61.6%)	
Female	123(54.7%)	61(62.9%)	29(69%)	33(38.4%)	
Hypertension	86(38.2%)	35(36.1%)	15(35.7%)	36(41.9%)	0.676
Diabetes	112(49.8%)	50(51.5%)	18(42.9%)	44(51.2%)	0.609
Smoking	39(17.3%)	14(14.4%)	6(14.3%)	19(22.1%)	0.333
Drinking	34(15.1%)	14(14.4%)	4(9.5%)	16(18.6%)	0.392
Irregular shape	79(35.1%)	36(37.1%)	13(31%)	30(34.9%)	0.782
Inflow angle	99.94±32.52	95.1±35.09	95.39±34.06	107.63±27.20	0.020
Tilt angle	87.15(58.06,116.32)	89.36(58.36,121.09)	93.67(71.64,122.77)	75.34(51.21,107.82)	0.157
Vascular angle	21.13(10.24,36.34)	17.85(8.59,34.61)	13.44(5.21,35.64)	25.02(18.04,37.20)	0.002
Hmax	3.05(2.32,4.74)	2.82(2.22,4.09)	3.15(2.29,5.63)	3.70(2.70,5.10)	0.018
Dmiddle	3.57(2.86,4.51)	3.56(3.00,4.47)	3.70(2.86,4.54)	3.55(2.58,4.63)	0.678
Dneck	3.56(2.92,4.49)	3.55(2.96,4.40)	3.54(2.87,4.46)	3.61(2.93,4.66)	0.992
Dvessel	3.05±0.73	3.15±0.76	3.09±0.84	2.93±0.62	0.113
H	2.74(2.03,4.11)	2.40(1.89,3.39)	2.86(1.92,4.12)	3.07(2.29,4.58)	0.003
S	35.35(18.08,68.40)	24.63(16.88,43.60)	31.67(16.01,60.56)	58.99(29.10,96.90)	<0.01

(*Continued*)

Table 2 (Continued).

Characteristics	No.(%)/M[IQR]	Training Cohort (n=97)	Internal Validation (n=42)	External Validation Cohort (n=86)	P value
V	23.90(10.25,62.37)	13.96(8.28,29.99)	18.89(7.50,50.53)	55.66(21.80,103.98)	<0.01
AR	0.80(0.58,1.08)	0.64(0.51,0.86)	0.79(0.52,0.99)	0.97(0.77,1.19)	<0.01
UI	0.09(0.06,0.17)	0.08(0.05,0.11)	0.07(0.04,0.13)	0.17(0.09,0.27)	<0.01
SR	1.06(0.76,1.58)	0.90(0.73,1.30)	1.05(0.76,1.66)	1.28(0.90,1.73)	0.002
EI	0.11(0.05,0.19)	0.07(0.03,0.12)	0.08(0.05,0.12)	0.19(0.12,0.26)	<0.01
NSI	0.11(0.07,0.19)	0.09(0.06,0.14)	0.10(0.07,0.17)	0.15(0.09,0.21)	<0.01
BNR	0.99(0.97,1.06)	0.99(0.98,0.99)	0.99(0.98,1.04)	0.94(0.82,1.2)	0.044
TAWSS	6.30(4.10,10.35)	7.91(4.64,14.90)	7.57(4.74,10.01)	5.73(3.54,8.10)	0.001
NWSS	0.83(0.58,1.13)	0.87(0.63,1.22)	0.83(0.51,1.05)	0.81(0.58,1.10)	0.337
TAWSSv	7.76(5.86,11.97)	9.62(6.09,14.34)	8.85(5.96,14.56)	6.99(5.57,8.82)	<0.01
WSSmax	27.87(17.52,42.56)	29.11(18.89,43.80)	27.55(17.60,43.51)	24.65(16.24,36.28)	0.257
WSSmin	1.07(0.55,2.72)	1.00(0.46,1.70)	0.64(0.31,1.16)	1.45(0.83,7.32)	<0.01
HAS	43.06(20.34,71.52)	41.24(20.50,73.61)	37.50(12.10,59.09)	56.04(25.80,72.44)	<0.01
LSA	2.12(0,9.69)	0(0,1.37)	1.26(0,7.65)	7.42(3.74,14.39)	<0.01
OSI	0.01(0,0.02)	0.01(0,0.01)	0.01(0,0.01)	0.03(0.01,0.04)	<0.01
OSImax	0.19(0.09,0.26)	0.13(0.07,0.22)	0.18(0.09,0.24)	0.23(0.14,0.31)	<0.01
WSSG	11.42(6.07,19.80)	10.02(6.25,19.53)	8.49(5.43,13.35)	15.42(8.33,21.69)	0.033
GON	0.06(0.04,0.09)	0.06(0.04,0.09)	0.07(0.06,0.09)	0.07(0.03,0.11)	0.164
RRT	0.38(0.16,1.25)	0.21(0.10,0.40)	0.25(0.16,0.53)	1.30(0.52,1.96)	<0.01

Interpretable Analysis

Figure 4 shows the contribution of each modeling variable to the Super learner in the external validation set. Hypertension, irregular shape, and SR exhibit a monotonically increasing trend with rupture risk. The increase in these factors may significantly raise the risk of aneurysm rupture. WSS min is negatively correlated with aneurysms, and lower WSS min may be an important predictor of rupture.

Discussion

In this study, five basic machine learning models and a super learner were constructed based on demographics, aneurysm characteristics, and hemodynamic parameters to accurately predict the rupture status of intracranial aneurysms. The models were validated in an independent external validation cohort, showing stable performance. Additionally, interpretability analysis of the super learner confirmed that hypertension, irregular shape, larger SR values, and lower WSSmin increase the rupture risk of intracranial aneurysms.

Previous studies have shown that the size of an aneurysm[19,20] is a significant factor influencing aneurysm rupture. Although smaller aneurysms have a lower rupture risk,[21,22] they still carry a risk of rupture. Ruptured aneurysms are associated with high mortality and morbidity rates. Therefore, the small aneurysm rupture risk prediction model developed in this study holds significant clinical value.

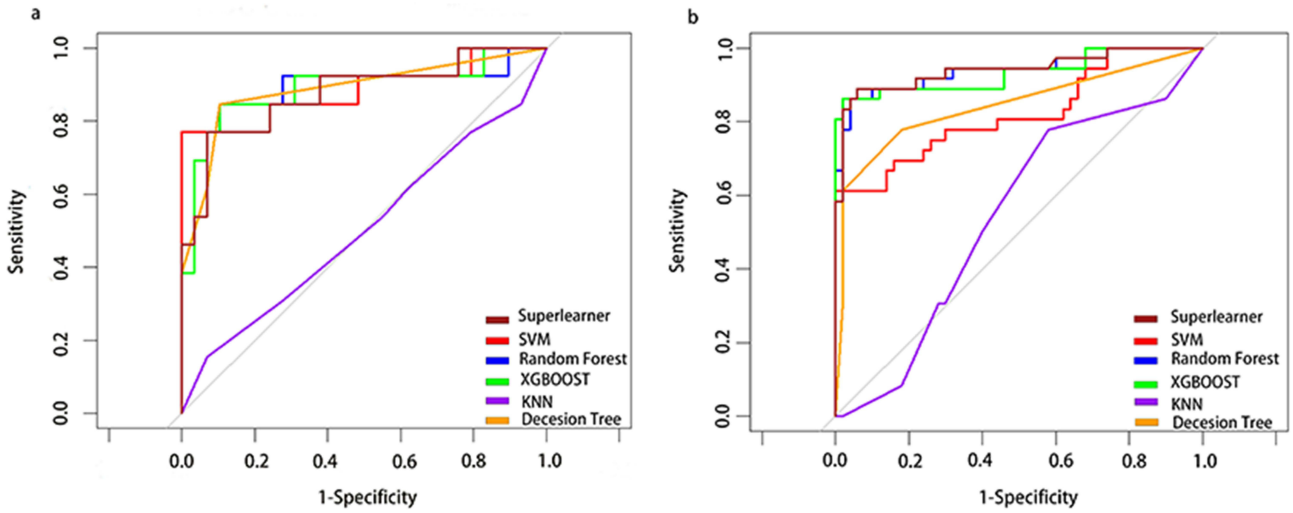

Figure 3 Receiver operating characteristic curves for Super learner in the internal (**a**) and external validation dataset (**b**).

Previous intracranial aneurysm rupture risk prediction models have largely been based on patient baseline data and morphological parameters of aneurysms,[14] which are static and do not reflect the true hemodynamic state within the aneurysm, thus presenting certain limitations.

This study used computational fluid dynamics to simulate the hemodynamic parameters of intracranial aneurysms in their true state, comprehensively analyzing the baseline data, morphology, and hemodynamics of ruptured aneurysms. It

Table 3 AUC, Sensitivity, Specificity, Accuracy, F1 Score, PPV and NPV of Super ML Model

Dataset	AUC(95% CI)	Sensitivity	Specificity	Accuracy	F1 Score	PPV	NPV
Super learner model							
Train	1(1–1)	1	1	1	1	1	1
Internal validation	0.86(0.72–1)	0.97	0.54	0.83	0.89	0.82	0.88
External validation	0.94(0.89–1)	1	0.61	0.84	0.88	0.78	1
Random Forest model							
Train	1(1–1)	1	1	1	1	1	1
Internal validation	0.73(0.59–0.88)	0.93	0.54	0.81	0.87	0.82	0.78
External validation	0.83(0.76–0.91)	0.93	0.54	0.81	0.87	0.82	0.78
XGBOOST model							
Train	1(1–1)	1	1	1	1	1	1
Internal validation	0.89(0.76–1)	0.97	0.46	0.81	0.88	0.80	0.85
External validation	0.93(0.87–0.99)	1	0.69	0.87	0.90	0.82	1
SVM model							
Train	0.99(0.99–1)	1	0.81	0.96	0.97	0.95	1
Internal validation	0.89(0.76–1)	1	0.69	0.90	0.94	0.88	1
External validation	0.82(0.73–0.92)	1	0.33	0.72	0.81	0.68	1

(*Continued*)

https://doi.org/10.2147/IJGM.S533558

Table 3 (Continued).

Dataset	AUC(95% CI)	Sensitivity	Specificity	Accuracy	F1 Score	PPV	NPV
Decision Tree							
Train	0.98(0.96–1)	0.96	0.86	0.94	0.96	0.96	0.86
Internal validation	0.89(0.77–1)	0.93	0.62	0.83	0.89	0.84	0.80
External validation	0.84(0.76–0.93)	0.98	0.61	0.83	0.87	0.78	0.96
Knn model							
Train	0.69(0.58–0.80)	1	0	0.78	0.88	0.78	1
Internal validation	0.53(0.32–0.74)	1	0	0.69	0.82	0.69	1
External validation	0.51(0.38–0.63)	1	0	0.58	0.74	0.58	1

found that hypertension, irregular morphology, larger SR values, and lower WSSmin may be significant factors influencing aneurysm rupture. This is consistent with previous research results,[23–26] which show that chronic hypertension leads to aneurysm expansion and a reduction in the wall strength, thereby increasing the rupture risk. Irregular

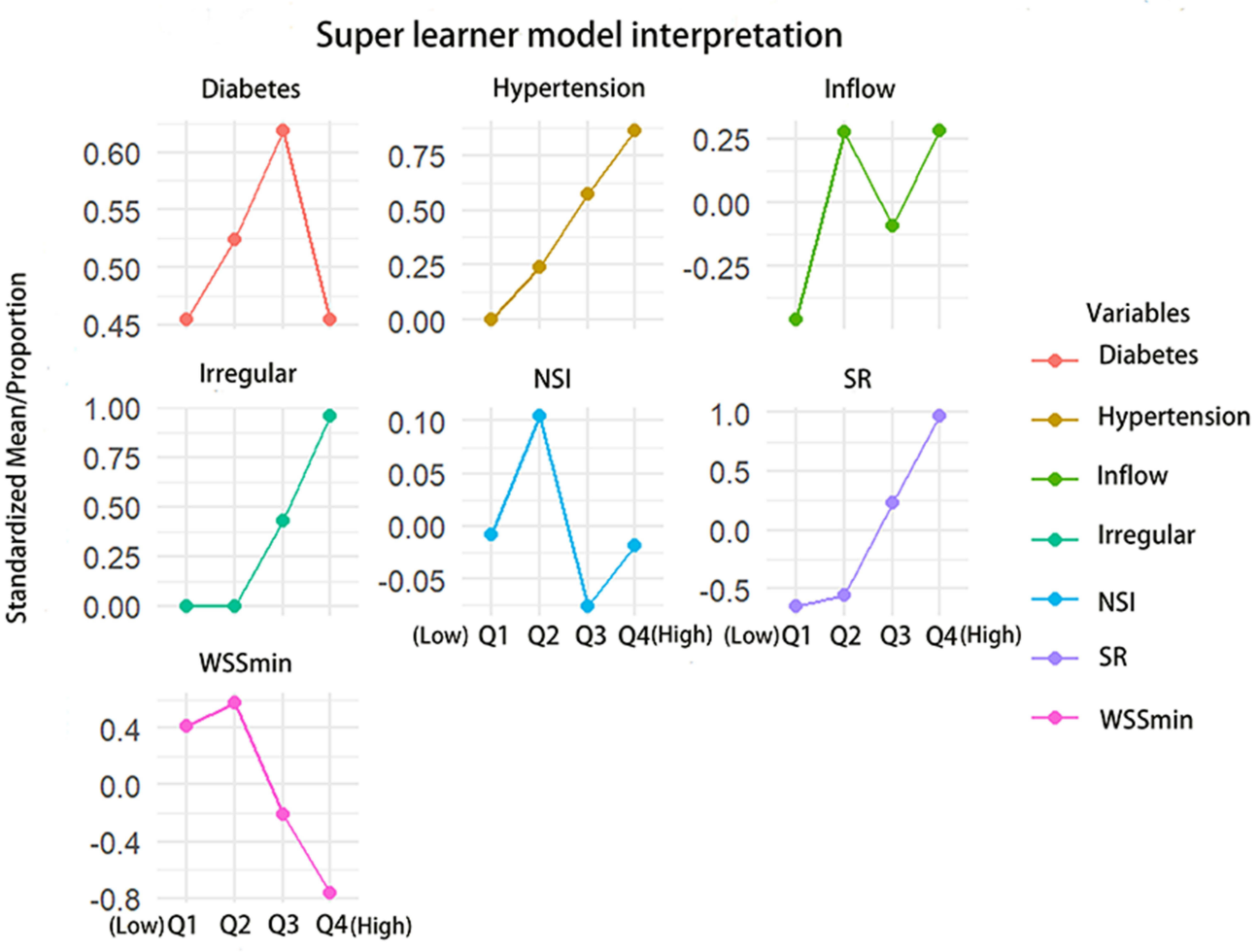

Figure 4 The interpretable analysis of Super learner model.

aneurysms are more prone to rupture, possibly due to turbulent flow within the aneurysm. The larger the SR value of the aneurysm, the larger it is relative to the neck, which is often associated with a slightly higher likelihood of intracranial aneurysm rupture. Extremely low WSSmin may lead to increased local inflammation, triggering aneurysm rupture.

Several prior studies have integrated morphology, hemodynamics, and patient characteristics to stratify the rupture risk of intracranial aneurysms. Notably, Detmer et al[27] developed a rupture probability model based on patient demographics, aneurysm location and morphology, and CFD-derived hemodynamic variables, achieving an AUC of approximately 0.85. Similarly, Rajabzadeh et al[28] proposed the Rupture Resemblance Score (RRS), a data-driven approach that identified high-risk aneurysms by comparing them to known ruptured phenotypes, demonstrating promising clinical utility.

Most existing machine learning studies on aneurysm rupture prediction have constructed multiple independent models,[29] assessing their performance using conventional metrics such as sensitivity, specificity, and AUC. While these approaches have shown reasonable accuracy, their predictive performance remains suboptimal. Owing to the distinct assumptions and algorithmic mechanisms underlying each model, no single method consistently achieves optimal performance across all evaluation domains.

In contrast, the Super Learner algorithm offers an ensemble-based framework that combines the strengths of diverse base learners through cross-validated weighting.[30] This approach improves model robustness and stability, especially in heterogeneous datasets, and reduces the risk of overfitting by minimizing reliance on any single poorly performing model.[31] Compared to traditional models like Detmer's or the RRS, the Super Learner in our study not only demonstrated superior discrimination (external AUC = 0.94), but also benefited from enhanced flexibility and generalizability without relying on rigid feature assumptions.

Additionally, our model incorporated interpretability analyses to reveal critical rupture-associated factors—including hypertension, irregular morphology, elevated SR, and reduced WSSmin—all of which align with previously reported risk features, thereby reinforcing the biological plausibility and reliability of our findings. Taken together, these results support the notion that ensemble-based machine learning approaches such as the Super Learner may provide a valuable complementary tool for aneurysm rupture risk stratification, especially when combined with detailed hemodynamic simulation and clinical data.

The Super Learner model achieved an AUC of 1.00 in the training set, which may suggest potential overfitting. However, several measures were implemented to mitigate this risk. First, stratified 10-fold cross-validation was used during training to improve model robustness. Second, ensemble learning methods such as Super Learner can inherently reduce variance and overfitting by combining diverse base learners. Most importantly, the model demonstrated strong and consistent performance in both internal and external validation cohorts, with an external AUC of 0.94, indicating good generalizability and suggesting that overfitting was limited.

One of the most promising clinical applications of our model lies in the management of small (<7 mm) unruptured intracranial aneurysms, where treatment decisions remain highly controversial. In many cases, clinicians rely heavily on experience and imaging features to estimate rupture risk. However, for small aneurysms lacking overt high-risk characteristics, this assessment can be highly subjective and uncertain. Our model offers a quantitative, individualized risk prediction that may assist in such ambiguous scenarios. When the model indicates a relatively high rupture probability despite the aneurysm's small size, clinicians may consider early intervention or intensified surveillance. Conversely, low-risk predictions may help avoid unnecessary procedures and reduce patient anxiety. Furthermore, as the model is applied prospectively, newly accumulated cases can be incorporated into ongoing training and refinement, enabling continuous performance improvement and enhanced clinical utility over time.

In this study, the Super learner accurately predicted the rupture status of intracranial aneurysms, with stable validation results, outperforming the other five base models in overall predictive performance and demonstrating higher reliability.

Limitation

Finally, this study has several limitations. First, the sample size is relatively small, with only 150 cases, which limits the generalizability of the findings. A larger sample size is needed to draw more robust conclusions. Second, the Super

learner is a black-box model, and further advancements in algorithms are needed to better understand its internal mechanisms.

Conclusion

This study developed a Super learner model based on demographics, aneurysm, and hemodynamic parameters, which accurately predicts the rupture status of intracranial aneurysms and outperforms the other five base models. Additionally, it validated that hypertension, irregular shape, larger SR values, and lower WSSmin increase the risk of intracranial aneurysm rupture.

Acknowledgments

This work was supported by The study was supported by National Natural Science Foundation of China (No. 82071279), Scientific funding of Shanghai fourth people's hospital (No. sykyqd06701;Y-YKYTS-2024-1003). Hongkou Leading Academic Discipline Project (No. HKLCFC202402).

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Disclosure

Xiaolong Hu and Shifei Ye are co-first authors for this work. The authors report no conflicts of interest in this work.

References

1. Vlak MH, Algra A, Brandenburg R, et al. Prevalence of unruptured intracranial aneurysms, with emphasis on sex, age, comorbidity, country, and time period: a systematic review and meta-analysis. *Lancet Neurol*. 2011;10(7):626–636. doi:10.1016/S1474-4422(11)70109-0
2. Wiebers DO, Torner JC, Meissner I. Impact of unruptured intracranial aneurysms on public health in the United States. *Stroke*. 1992;23 (10):1416–1419. doi:10.1161/01.STR.23.10.1416
3. Kelly PJ, Stein J, Shafqat S, et al. Functional recovery after rehabilitation for cerebellar stroke. *Stroke*. 2001;32(2):530–534. doi:10.1161/01.STR.32.2.530
4. Wiebers DO, Whisnant JP, Huston J, et al. Unruptured intracranial aneurysms: natural history, clinical outcome, and risks of surgical and endovascular treatment. *Lancet*. 2003;362(9378):103–110. doi:10.1016/S0140-6736(03)13860-3
5. Waqas M, Chin F, Rajabzadeh-Oghaz H, et al. Size of ruptured intracranial aneurysms: a systematic review and meta-analysis. *Acta Neurochir*. 2020;162(6):1353–1362. doi:10.1007/s00701-020-04291-z
6. Malhotra A, Wu X, Forman HP, et al. Growth and rupture risk of small unruptured intracranial aneurysms: a systematic review [published correction appears in Ann Intern Med. 2018 Dec 4;169(11):824. *Ann Intern Med*. 2017;167(1):26–33. doi:10.7326/M17-0246
7. Alleyne CH Jr. Aneurysmal subarachnoid hemorrhage: have outcomes really improved? *Neurology*. 2010;74(19):1486–1487. doi:10.1212/WNL.0b013e3181e0ef1a
8. UCAS Japan Investigators, Morita A, Kirino T, et al. The natural course of unruptured cerebral aneurysms in a Japanese cohort. *N Engl J Med*. 2012;366(26):2474–2482.
9. Greving JP, Wermer MJ, Brown RD Jr, et al. Development of the PHASES score for prediction of risk of rupture of intracranial aneurysms: a pooled analysis of six prospective cohort studies. *Lancet Neurol*. 2014;13(1):59–66. doi:10.1016/S1474-4422(13)70263-1
10. Saqr KM, Rashad S, Tupin S, et al. What does computational fluid dynamics tell us about intracranial aneurysms? A meta-analysis and critical review. *J Cereb Blood Flow Metab*. 2020;40(5):1021–1039. doi:10.1177/0271678X19854640
11. Hu X, Deng P, Ma M, et al. How does the recurrence-related morphology characteristics of the Pcom aneurysms correlated with hemodynamics? *Front Neurol*. 2023;14:1236757. doi:10.3389/fneur.2023.1236757
12. Senders JT, Arnaout O, Karhade AV, et al. Natural and artificial intelligence in neurosurgery: a systematic review. *Neurosurgery*. 2018;83 (2):181–192. doi:10.1093/neuros/nyx384
13. Senders JT, Staples PC, Karhade AV, et al. Machine learning and neurosurgical outcome prediction: a systematic review. *World Neurosurg*. 2018;109:476–486. doi:10.1016/j.wneu.2017.09.149
14. Liu Q, Jiang P, Jiang Y, et al. Prediction of aneurysm stability using a machine learning model based on pyradiomics-derived morphological features. *Stroke*. 2019;50(9):2314–2321. doi:10.1161/STROKEAHA.119.025777
15. Valdes G, Interian Y, Gennatas E, et al. The conditional super learner. *IEEE Trans Pattern Anal Mach Intell*. 2022;44(12):10236–10243. doi:10.1109/TPAMI.2021.3131976
16. Li Y, Wang B, Ma F, et al. Using the super-learner to predict the chemical acute toxicity on rats. *J Hazard Mater*. 2024;480:136311. doi:10.1016/j.jhazmat.2024.136311

17. Rose S. Mortality risk score prediction in an elderly population using machine learning. *Am J Epidemiol*. 2013;177(5):443–452. doi:10.1093/aje/kws241
18. Xiang J, Natarajan SK, Tremmel M, et al. Hemodynamic-morphologic discriminants for intracranial aneurysm rupture. *Stroke*. 2011;42(1):144–152. doi:10.1161/STROKEAHA.110.592923
19. Brown RD Jr, Broderick JP. Unruptured intracranial aneurysms: epidemiology, natural history, management options, and familial screening. *Lancet Neurol*. 2014;13(4):393–404. doi:10.1016/S1474-4422(14)70015-8
20. Jagadeesan BD, Delgado Almandoz JE, Kadkhodayan Y, et al. Size and anatomic location of ruptured intracranial aneurysms in patients with single and multiple aneurysms: a retrospective study from a single center. *J Neurointerv Surg*. 2014;6(3):169–174. doi:10.1136/neurintsurg-2012-010623
21. Weir B, Disney L, Karrison T. Sizes of ruptured and unruptured aneurysms in relation to their sites and the ages of patients. *J Neurosurg*. 2002;96(1):64–70. doi:10.3171/jns.2002.96.1.0064
22. Bijlenga P, Ebeling C, Jaegersberg M, et al. Risk of rupture of small anterior communicating artery aneurysms is similar to posterior circulation aneurysms. *Stroke*. 2013;44(11):3018–3026. doi:10.1161/STROKEAHA.113.001667
23. Spittell JA Jr. Hypertension and arterial aneurysm. *J Am Coll Cardiol*. 1983;1(2 Pt 1):533–540. doi:10.1016/S0735-1097(83)80085-0
24. Lindgren AE, Koivisto T, Björkman J, et al. Irregular shape of intracranial aneurysm indicates rupture risk irrespective of size in a population-based cohort. *Stroke*. 2016;47(5):1219–1226. doi:10.1161/STROKEAHA.115.012404
25. Zheng Y, Zhou B, Wang X, et al. Size, aspect ratio and anatomic location of ruptured intracranial aneurysms: consecutive series of 415 patients from a prospective, multicenter, observational study. *Cell Transplant*. 2019;28(6):739–746. doi:10.1177/0963689718817227
26. Meng H, Tutino VM, Xiang J, et al. High WSS or low WSS? Complex interactions of hemodynamics with intracranial aneurysm initiation, growth, and rupture: toward a unifying hypothesis. *AJNR Am J Neuroradiol*. 2014;35(7):1254–1262. doi:10.3174/ajnr.A3558
27. Detmer FJ, Chung BJ, Mut F, et al. Development and internal validation of an aneurysm rupture probability model based on patient characteristics and aneurysm location, morphology, and hemodynamics. *Int J Comput Assist Radiol Surg*. 2018;13(11):1767–1779. doi:10.1007/s11548-018-1837-0
28. Rajabzadeh-Oghaz H, Waqas M, Veeturi SS, et al. A data-driven model to identify high-risk aneurysms and guide management decisions: the Rupture resemblance score. *J Neurosurg*. 2020;135(1):9–16. doi:10.3171/2020.5.JNS193264
29. Chen G, Lu M, Shi Z, et al. Development and validation of machine learning prediction model based on computed tomography angiography-derived hemodynamics for rupture status of intracranial aneurysms: a Chinese multicenter study. *Eur Radiol*. 2020;30:5170–5182. doi:10.1007/s00330-020-06886-7
30. Van der Laan MJ, Rose S. *Targeted Learning: Causal Inference for Observational and Experimental Data[M]*. New York: Springer; 2011.
31. Charu V, Liang JW, Mannalithara A, et al. Benchmarking clinical risk prediction algorithms with ensemble machine learning for the noninvasive diagnosis of liver fibrosis in NAFLD. *Hepatology*. 2024;80(5):1184–1195. doi:10.1097/HEP.0000000000000908