

Get Clarity On Generics

Cost-Effective CT & MRI Contrast Agents

 FRESENIUS
KABI

[WATCH VIDEO](#)

AJNR

This information is current as
of August 20, 2025.

Artificial Intelligence–Assisted Evaluation of the Spatial Relationship between Brain Arteriovenous Malformations and the Corticospinal Tract to Predict Postsurgical Motor Defects

Y. Jiao, J. Zhang, X. Yang, T. Zhan, Z. Wu, Y. Li, S. Zhao,
H. Li, J. Weng, R. Huo, J. Wang, H. Xu, Y. Sun, S. Wang
and Y. Cao

AJNR Am J Neuroradiol 2023, 44 (1) 17-25

doi: <https://doi.org/10.3174/ajnr.A7735>

<http://www.ajnr.org/content/44/1/17>

Artificial Intelligence–Assisted Evaluation of the Spatial Relationship between Brain Arteriovenous Malformations and the Corticospinal Tract to Predict Postsurgical Motor Defects

Y. Jiao, J. Zhang, X. Yang, T. Zhan, Z. Wu, Y. Li, S. Zhao, H. Li, J. Weng, R. Huo, J. Wang, H. Xu, Y. Sun, S. Wang, and Y. Cao



ABSTRACT

BACKGROUND AND PURPOSE: Preoperative evaluation of brain AVMs is crucial for the selection of surgical candidates. Our goal was to use artificial intelligence to predict postsurgical motor defects in patients with brain AVMs involving motor-related areas.

MATERIALS AND METHODS: Eighty-three patients who underwent microsurgical resection of brain AVMs involving motor-related areas were retrospectively reviewed. Four artificial intelligence–based indicators were calculated with artificial intelligence on TOF-MRA and DTI, including $FN_{5mm/50mm}$ (the proportion of fiber numbers within 5–50mm from the lesion border), $FN_{10mm/50mm}$ (the same but within 10–50mm), $FP_{5mm/50mm}$ (the proportion of fiber voxel points within 5–50mm from the lesion border), and $FP_{10mm/50mm}$ (the same but within 10–50mm). The association between the variables and long-term postsurgical motor defects was analyzed using univariate and multivariate analyses. Least absolute shrinkage and selection operator regression with the Pearson correlation coefficient was used to select the optimal features to develop the machine learning model to predict postsurgical motor defects. The area under the curve was calculated to evaluate the predictive performance.

RESULTS: In patients with and without postsurgical motor defects, the mean $FN_{5mm/50mm}$, $FN_{10mm/50mm}$, $FP_{5mm/50mm}$, and $FP_{10mm/50mm}$ were 0.24 (SD, 0.24) and 0.03 (SD, 0.06), 0.37 (SD, 0.27) and 0.06 (SD, 0.08), 0.06 (SD, 0.10) and 0.01 (SD, 0.02), and 0.10 (SD, 0.12) and 0.02 (SD, 0.05), respectively. Univariate and multivariate logistic analyses identified $FN_{10mm/50mm}$ as an independent risk factor for long-term postsurgical motor defects ($P = .002$). $FN_{10mm/50mm}$ achieved a mean area under the curve of 0.86 (SD, 0.08). The mean area under the curve of the machine learning model consisting of $FN_{10mm/50mm}$, diffuseness, and the Spetzler-Martin score was 0.88 (SD, 0.07).

CONCLUSIONS: The artificial intelligence–based indicator, $FN_{10mm/50mm}$, can reflect the lesion-fiber spatial relationship and act as a dominant predictor for postsurgical motor defects in patients with brain AVMs involving motor-related areas.

ABBREVIATIONS: AI = artificial intelligence; AUC = area under the curve; BAVM = brain AVM; CST = corticospinal tract; EuDX = Euler Delta Crossing; FN = fiber number; FP = proportion of fiber voxel points; HDVL = hemorrhagic presentation, diffuseness, deep venous drainage, and lesion-to-eloquence distance; LASSO = least absolute shrinkage and selection operator; LCD = lesion-to-corticospinal tract distance; LR = logistic regression; M-AVM = brain AVM involving motor-related areas; MDs = motor defects; ML = machine learning; ROC = receiver operating characteristic; S-M = Spetzler-Martin grading; XGBoost = Extreme Gradient Boosting

Brain AVMs (BAVMs) are congenital vascular dysplasias of the CNS that cause cerebral hemorrhage.¹ Radical resection by microsurgery to eliminate abnormal vascular masses is an

important curative treatment for BAVMs.² However, surgical management of BAVMs, particularly involving eloquent areas, may lead to serious neurologic defects and a poor prognosis.^{3,4} Preoperative evaluation for BAVMs to identify proper surgical candidates is crucial for reducing the risk of postsurgical defects.⁵

Received March 18, 2022; accepted after revision November 7.

From the Department of Neurosurgery (Y.J., J.Z., S.Z., H.L., J. Weng, R.H., J. Wang, H.X., Y.S., S.W., Y.C.), Beijing Tiantan Hospital, Capital Medical University, Beijing, China; and China National Clinical Research Center for Neurological Diseases (Y.J., J.Z., X.Y., T.Z., Z.W., Y.L., S.Z., H.L., J. Weng, R.H., J. Wang, H.X., Y.S., S.W., Y.C.), Beijing, China.

This study was supported by the National Natural Science Foundation of China (grant No. 81901175; principal investigator, Yuming Jiao), the National Key Research and Development Program of China during the 14th Five-Year Plan Period (grant No. 2021YFC2501102; principal investigator, Yong Cao), and the Beijing Advanced Innovation Center for Big Data-Based Precision Medicine (grant No. PXM2020_014226_000066).

Please address correspondence to Yong Cao, MD, Department of Neurosurgery, Beijing Tiantan Hospital, Capital Medical University, 119 South Fourth Ring Rd West, Fengtai District, Beijing, 100071, China; e-mail: caoyong@bjtt.org

Indicates open access to non-subscribers at www.ajnr.org

Indicates article with online supplemental data.

<http://dx.doi.org/10.3174/ajnr.A7735>

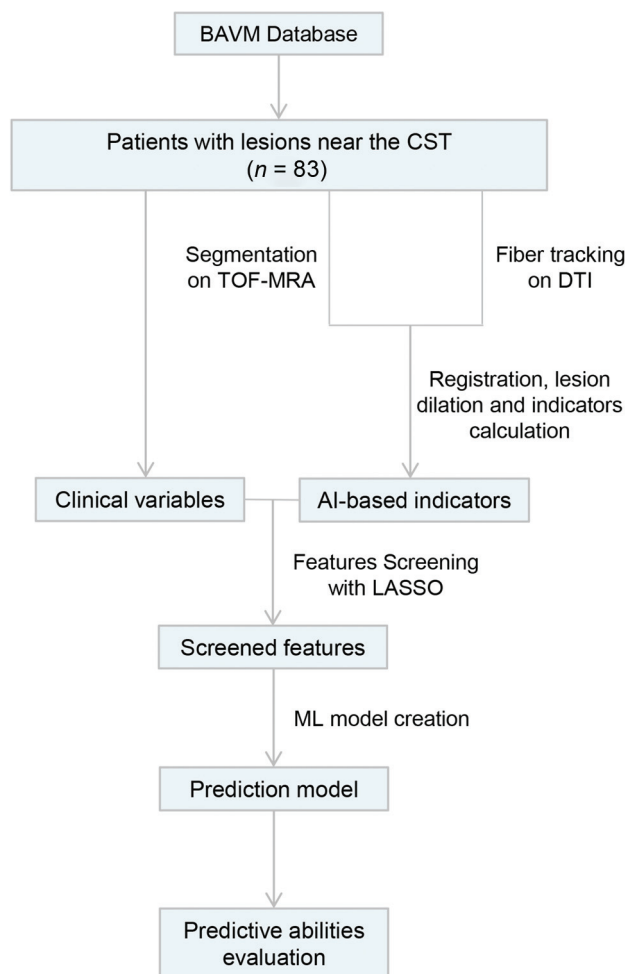


FIG 1. Flow diagram of patient selection, AI-based indicator calculation, ML model development, and predictive ability evaluation. AI-based indicators include $FN_{5mm/50mm}$, $FN_{10mm/50mm}$, $FP_{5mm/50mm}$, and $FP_{10mm/50mm}$. Clinical variables include age, sex, side, lobe, size, deep perforating artery supply, deep venous drainage, hemorrhage, seizure, and S-M grading score. LASSO indicates least absolute shrinkage and selection operator.

In our previous studies using fMRI and DTI, lesion-to-corticospinal tract distance (LCD) was used to evaluate the relationship between white fibers and lesions, which has proved to be an important predictive factor for postsurgical motor defects (MDs) in BAVMs involving motor-related areas (M-AVMs).^{6,7}

However, because BAVM lesions and the corticospinal tract (CST) are both stereoscopic objects, LCD obviously has some limitations in describing the spatial relationship between them. First, LCD is measured on the basis of a 2D plane and is a measurement that is too simple and crude to describe the complicated 3D spatial relationship between CST and BAVM lesions.^{8,9} Second, LCD is measured manually, which may lead to relatively inaccurate and subjective measurements.¹⁰ Finally, as a basic step of LCD calculation, fiber tracking skills may have confounding bias, which might vary with the operators' different levels of clinical experience or in different clinical centers.¹¹

Artificial intelligence (AI) technology, with its powerful processing and analysis ability, is emerging to develop a predictive model for clinical prognosis.¹² AI-based tools have the capability

of overcoming the shortcomings of manual evaluation, such as time-consuming workflows and substantial clinical experience requirements.¹³ Meanwhile, AI also has great potential in calculating the sophisticated 3D spatial relationship between stereoscopic objects.^{14,15} Whether exploiting AI technology to evaluate BAVMs has a better ability to predict postoperative motor function in M-AVMs remains unknown.

In this study, 83 patients who underwent microsurgical resection of M-AVMs were retrospectively reviewed. Four AI technology-based indicators were calculated followed by a machine learning (ML) model to predict the long-term postsurgical MDs of patients. Our study indicated that AI technology is an excellent method for predicting postoperative motor function in patients with M-AVMs.

MATERIALS AND METHODS

Patients

All patients were retrospectively reviewed from our BAVM database of a prospective clinical trial (ClinicalTrials.gov Identifier: NCT02868008) conducted between July 2015 and December 2020. Inclusion criteria were the following: 1) The lesions were near the CST (LCDs were <10 mm based on DTI tractography), 2) patients underwent microsurgical resection of the BAVMs, 3) patients were followed up for 6 months with complete information of motor function, and 4) patients' TOF-MRA and DTI data were acquired before microsurgery. The exclusion criteria included the following: 1) poor-quality neuroimaging, 2) atypical BAVM images after stereotactic radiosurgery or interventional therapy, and 3) incomplete imaging or follow-up data. The standard of postsurgical MDs was defined as patients not recovering to the level of being able to carry out all their usual activities (mRS > 2) 6 months after microsurgery as previously reported.^{5,16} Finally, 83 patients with BAVM lesions near the CST who underwent microsurgery were enrolled. The study was reviewed and approved by the ethics committee of the Beijing Tiantan Hospital Affiliated with Capital Medical University (approval No. KY2016-031-01). Before the study, the patients were informed about the design of the study, and each participant provided his or her informed consent. The design and all procedures adhered to the latest version of the Helsinki Declaration. The flow diagram of patient selection, AI-based indicator calculation, model development, and predictive ability evaluation is shown in Fig 1.

Baseline Data

The clinical factors including patients' age, sex, hemorrhage, and seizure information were collected from the prospectively collected database and the electronic medical records system by a neurosurgeon (J. Zhang). The variables such as BAVM size, lobe, diffuseness, deep venous drainage, deep perforating artery supply, Spetzler-Martin (S-M) score, and hemorrhagic presentation, diffuseness, deep venous drainage, and lesion-to-eloquence distance (HDVL) score were determined with consensus by 2 experienced neurosurgeons (H. Li and J. Weng) from preoperative angiograms, traditional MR imaging scans, and TOF-MRA images.⁵⁻⁷ As described in our previous study, the LCD was defined as the nearest distance from the BAVM lesions to the CST. The LCD_{Manual} serving only as an inclusion criterion (LCD_{Manual}

<10 mm), was measured with consensus by 2 experienced neurosurgeons (H. Li and Y. Jiao) on the iPlan 3.0 workstation (Brainlab). Sections for measurement were manually selected where the tracked CST appeared to be nearest to the margin of the lesions on the TOF-MRA images fused with fMRI and DTI.¹⁷ Meanwhile, the lesion-to-corticospinal tract distance measured by the AI method (LCD_{AI}) was used for developing a predictive model. Distances of any 2 points between the surface of the segmented lesions and tracked CSTs were measured by matrix operations. The minimum value among the distances was the nearest LCD_{AI}.

MR Image Acquisition

The MR images were obtained using a 3T MR imaging scanner (Magnetom Trio; Siemens). The sagittal T1 anatomic image acquired was a gradient-echo sequence: TR = 2300 ms, TE = 2.98 ms, section thickness = 1 mm, slices = 176, FOV = 256 mm, flip angle = 9°, matrix = 64 × 64, voxel size = 1 × 1 × 1 mm³, and bandwidth = 240 kHz. Axial TOF-MRA was performed using a 3D TOF gradient-echo acquisition sequence: TR = 22 ms, TE = 3.86 ms, section thickness = 1 mm, slices = 36 × 4, FOV = 220 × 220 mm², flip angle = 120°, and matrix = 512 × 512. DTI was performed using the DWI-EPI technique: TR = 6100 ms, TE 93 = ms, section thickness = 3 mm, slices = 45, FOV = 230 × 230 mm², and matrix = 128 × 128 with a motion-probing gradient in 30 orientations.^{6,7}

Lesion Segmentation, Fiber Tracking, Registration, and Lesion Dilation

Lesion Segmentation Assisted by AI. In the segmentation phase, we used the U-Net model proposed in our previous study to delineate BAVM lesions automatically. Before we trained the model, all AVM lesions on the TOF-MRA images were manually labeled with lesion masks by 3 neuroradiologists (H. Li, J. Wang, and J. Zhang) as the ground truth references according to their signal, shape, course, texture, and so forth. The manually labeled images were used to train and test the U-Net model. In the process of developing the U-Net segmentation model, we set the number of training epochs as 300 and evaluate the model after each epoch with a Dice score. The criterion used to stop the iteration was no more than a 0.5% increase of the Dice score in 50 consecutive epochs. The results of AI segmentation were checked and verified by 2 neurosurgeons (H. Li and J. Wang). Finally, the Dice score of the model reached 0.80. In addition, the goal of our study was to build a fully AI-based automatic prediction model for convenient clinical application: inputting the raw imaging data into the model without the time-consuming manual segmentation by neurosurgeons and directly going to the prediction of the prognosis. Therefore, in the present study, we used the U-Net model for segmentation rather than using manual segmentation. The details of the segmentation method were described in our previous work.¹⁸

Fiber Tracking. Fiber tracking was based on DTI, which could reflect the anisotropy of the diffusion motion of water molecules. The Constant Solid Angle model and the Euler Delta Crossing (EuDX) algorithm were used to track the fibers on DTI with the Python Library Dipy (<https://github.com/dipy>).^{19,20} The Constant

Solid Angle model was used to obtain the orientation distribution function of each voxel on DTI scans. The orientation distribution function was used to express the probability of the fiber distribution in finite directions per voxel.²¹ The EuDX algorithm used the obtained orientation distribution function to connect the directions of these voxels with the highest probabilities to form a complete fiber tract. The results of AI fiber tracking were checked and verified by neurosurgeons (H. Li and J. Wang).

Registration Assisted by AI. In this study, the registration method included affine registration and deformable registration; the objectives of using these methods were to find a coordinate transformation and to bring the 2 images as close as possible. The difference between affine registration and deformable registration was that the affine registration performed a linear transform operation and each voxel of the image had the same transformation, so the 2 images had the same size and position after affine registration.²²

Affine registration can be expressed as

$$\begin{bmatrix} x_A \\ y_A \\ z_A \\ 1 \end{bmatrix} = T_A \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} A_{1,1} & A_{1,2} & A_{1,3} & A_{1,4} \\ A_{2,1} & A_{2,2} & A_{2,3} & A_{2,4} \\ A_{3,1} & A_{3,2} & A_{3,3} & A_{3,4} \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}.$$

T_A is a 4 × 4 matrix that represents linear transformation in affine registration. (X, y, z) is the coordinate of a voxel in an image. (x_A, y_A, z_A) is the coordinate of a voxel in an image after affine registration.

On the other hand, deformable registration performed nonlinear transformations for which each voxel of the image had its own transformation direction, so the 2 images had the same shape and texture after deformable registration.²³ The whole process is represented as follows:

$$\begin{bmatrix} x_F \\ y_F \\ z_F \end{bmatrix} = \begin{bmatrix} \Delta X \\ \Delta Y \\ \Delta Z \end{bmatrix} + \begin{bmatrix} x_A \\ y_A \\ z_A \end{bmatrix}.$$

$\Delta X, \Delta Y, \Delta Z$ indicates the offset in each dimension of each voxel. (x_F, y_F, z_F) is the coordinate of a pixel in an image after deformable registration. The nonlinear transformation operation was performed on the coordinates (x_A, y_A, z_A) after affine registration.

To obtain T_A and ($\Delta X, \Delta Y, \Delta Z$), we used U-Net neural networks to build a model for the 2 registration processes.²⁴ In the affine registration model, the decoder was designed as a fully connected neural network, and the output was a 3 × 4 matrix. In the deformable registration model, the output was a 3-channel matrix, and each channel was represented as ($\Delta X, \Delta Y, \Delta Z$). To achieve end-to-end training, we added interpolation operations into the neural network model for which the output of the network was a registered image. In addition, our study could train the affine registration and deformable registration in a cascade by which the images input into the model could undergo both affine and deformable registration.

The registration method was used to complete the functional segmentation of the medical images. The ICBM-152 (<https://>

nist.mni.mcgill.ca/icbm-152-nonlinear-atlases-2009/), which has detailed anatomic structural segmentation, was used as an atlas.^{25,26} We first extracted all the data for fibers of patients in DTI scans through the Constant Solid Angle model and the EuDX algorithm. We performed affine registration and deformable registration operations on the patients' DTI scans with the atlas, and the patients' anatomic structural segmentation data were obtained. Then, we found ROIs, the anterior half of the lower pons on the ipsilateral side and precentral gyrus, related to motor function in the patients' anatomic structural segmentation.²⁷ The fibers between the 2 ROIs were extracted from the patients' DTI scans; these were the CST of the patient. When we combined the results of segmentation and fiber tracking on the patients' DTI scans, the fibers related to motor function could be distinguished from all the fibers. Finally, the CST tracked on DTI scans was registered to TOF-MRA scans with BAVM segmentation by affine registration so that the AI-based indicators reflecting the spatial relationship of the CST and lesion could be calculated (Fig 2).

Lesion Dilation. To evaluate the spatial relationship, we performed a dilation operation on the lesion mask segmented by the U-Net model to obtain the AI-based indicators. The dilation operation was implemented using affine transformation. The centroid position (x_c, y_c, z_c) of the lesion mask was calculated, and the position was used as the center of the affine transformation. Then, all the parameters related to the times of dilation in the affine transformation matrix were calculated. We recorded the size of each voxel in an image as x_v, y_v, z_v , which represented the size of each voxel in 3 dimensions. If the size of the area where the lesion was located was (h, w, d) and we wanted to dilate the lesion by β mm, the dilated lesion was in an $(h + 2*\beta/x_v, w + 2*\beta/y_v, d + 2*\beta/z_v)$ -sized area—that is, the voxels $\beta/x_v, \beta/y_v$, and β/z_v were expanded at both ends of the region. Therefore, the lesions were dilated $((h + 2*\beta/x_v)/h, (w + 2*\beta/y_v)/w, (d + 2*\beta/z_v)/d)$ times in each dimension.

The affine transformation matrix was expressed as

$$D = \begin{bmatrix} \frac{h + 2*\beta/x_v}{h} & 0 & 0 & 0 \\ 0 & \frac{w + 2*\beta/y_v}{w} & 0 & 0 \\ 0 & 0 & \frac{d + 2*\beta/z_v}{d} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

The dilation operation can be expressed as

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = D \begin{bmatrix} x - x_c \\ y - y_c \\ z - z_c \\ 1 \end{bmatrix} + \begin{bmatrix} x_c \\ y_c \\ z_c \\ 0 \end{bmatrix}.$$

AI-Based Indicator Calculations in Patients with M-AVMs

In this study, 4 AI-based indicators were proposed to predict postsurgical MDs, including fiber number (FN)_{5mm/50mm}, FN_{10mm/50mm}, and proportion of fiber voxel points (FP)_{5mm/50mm} and FP_{10mm/50mm}. FN_{5mm/50mm} and FN_{10mm/50mm} indicated the

proportion of fiber numbers from 5 to 50mm and from 10 to 50mm from the lesion border, respectively. FP_{5mm/50mm} and FP_{10mm/50mm} indicate the proportion of fiber voxel points from 5 to 50mm and from 10 to 50mm from the lesion border, respectively. The 4 indicators may reflect the lesion-to-fiber spatial relationship and the degree of potential CST damage caused by BAVM lesions. These indicators were automatically calculated by AI algorithm-assisted methods based on the results of lesion segmentation, fiber tracking, registration, and lesion dilation. In the process of indicator calculations, we used the method of lesion dilation to illustrate the distances of 5 mm, 10 mm, and 50mm from the borders of the lesion (Fig 3).

Development of a Machine Learning Model

In our study, we used the least absolute shrinkage and selection operator (LASSO) regression with the Pearson correlation coefficient to select the optimal features to develop the ML model among clinical features (including age, sex, lobe, size, deep perforating artery supply, deep venous drainage, diffuseness, hemorrhage, seizure, S-M score, LCD_{AB}, and AI-based features (including FN_{5mm/50mm}, FN_{10mm/50mm}, FP_{5mm/50mm}, and FP_{10mm/50mm}).

The aim of this process was to minimize the prediction error, determined by the following equation:

$$\operatorname{argmin}_{\beta} \|Y - X\beta\|^2 \text{ subject to } \|\beta\| = \sum_{j=0}^d \beta_j \leq t,$$

where β_j is the regression coefficient of variables selected and t is the L1-norm of the regression coefficient that controls the degree of penalty.²⁸ The impact of the penalty function was that coefficient estimates with a small contribution to the model were forced to be exactly zero.

On the basis of LASSO regression, a nomogram was built to provide a quantitative tool for clinical use. Additionally, different ML algorithms were used to build the model to find the best algorithm, including logistic regression (LR), random forest, Extreme Gradient Boosting (XGBoost) (<https://www.geeksforgeeks.org/xgboost/>), and support vector machine.

In this study, nested cross-validation comprising an outer cross-validation loop and an inner cross-validation loop was used for model optimization and evaluation. In the outer cross-validation loop, the data set was split into 10 equally sized folds based on k-fold stratified cross-validation. For each cross-validation iteration, 9 folds of data were used as the training set and 1 was used as the testing set. The training set was used as input in the inner cross-validation loop and was split into 5 folds (4 were used for training and 1 was used for validation) for feature selection and hyperparameter optimization.²⁹ The nested cross-validation was randomly repeated 50 times, and 50 different experimental results were obtained to ensure the robustness and stability of the model. The mean areas under the curve (AUCs) of the models built by different algorithms were compared to find the best performing algorithm to build the ML model.

Statistical Analysis

Statistical analyses were performed with SPSS, Version 20.0.0 (IBM). Nomograms, calibration plots, and decision curves were

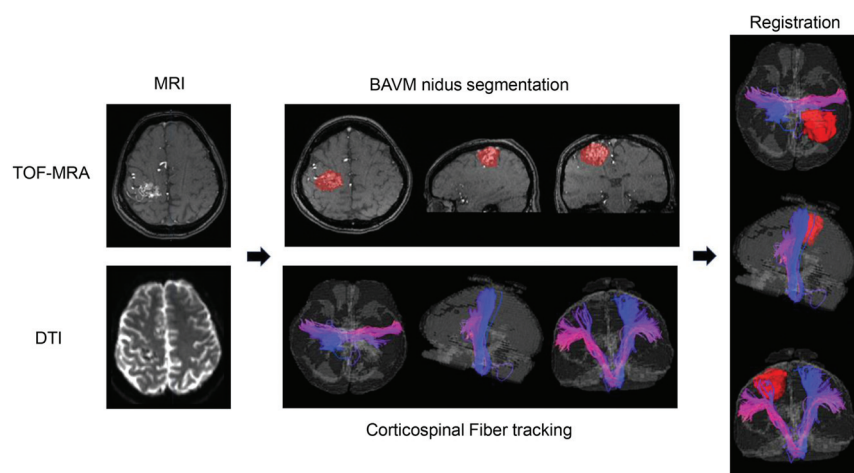


FIG 2. The workflow of segmentation, fiber tracking, and registration. Segmentation and fiber tracking were performed by the algorithm independently on TOF-MRA and DTI scans. The CSTs tracked on DTI scans were registered to TOF-MRA with BAVM segmentation to calculate the AI-based indicators.

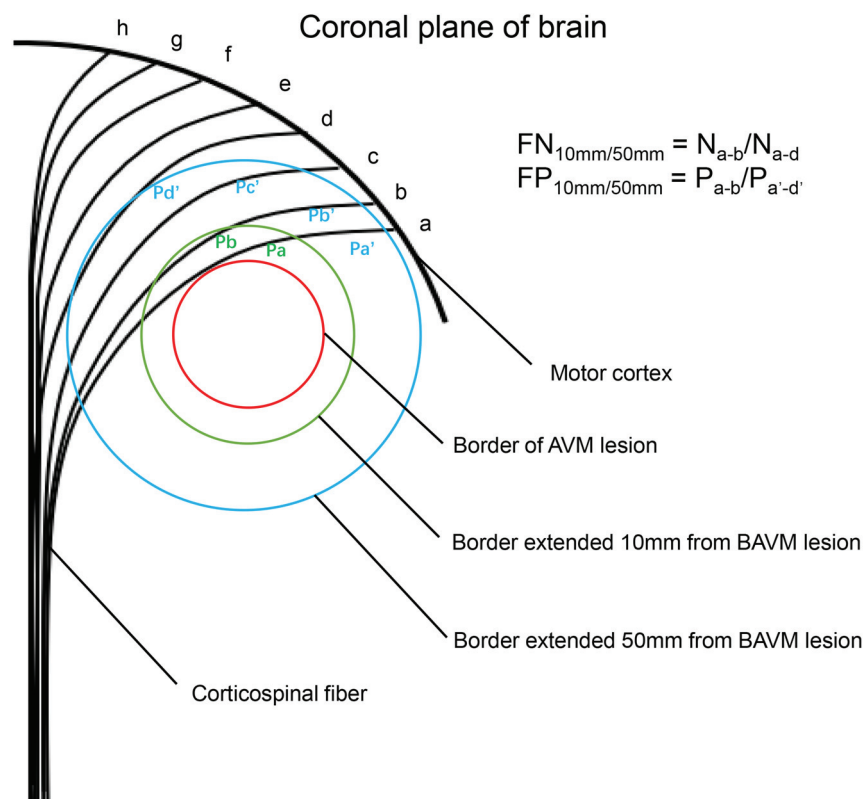


FIG 3. Schematic diagram of AI-based indicator calculation. The letters a–h indicate CSTs. The red, green, and blue circles indicate the borders of the BAVM lesion and the border with radii extended by the dilation algorithm to 10 and 50 mm from the lesion, respectively. Letters a–b indicate CSTs within 10 mm of the border of the BAVM lesion. Letters a–d indicate CSTs within 50 mm from the border of the BAVM lesion. N_{a-b} and N_{a-d} indicate the number of fibers between a and b and between a and d, respectively. P_{a-b} and $P_{a'-d'}$ indicate the sum of voxel points of fibers between a and b and between a and d, respectively.

generated using R statistical and computing software, Version R-4.0.5 (<http://www.r-project.org>). Receiver operating characteristic (ROC) plots were generated using Python, Version 3.6.0. The association between the variables and postsurgical MDs was

analyzed using univariate and multivariate analyses. The AUCs of different ML algorithms were compared to select the optimum algorithm with the highest AUC to build the ML model. The AUCs were calculated to compare the predictive ability of the AI-based indicator and ML model with reported predictive methods (LCD_{AI}, S-M score, and HDVL score).^{7,30} The DeLong test was performed to compare the AUC between FN_{10mm/50mm} and other predictive methods. The accuracy, specificity, sensitivity, precision, recall, and F1 score of different predictive methods were calculated to evaluate their predictive abilities. A calibration plot was used to graphically represent the agreement between the probability predicted by the nomogram and the actual probability of postsurgical MDs. The Brier Score of the calibration plot was calculated to evaluate the predictive ability of the nomogram. All statistical tests were 2-sided with a significance level of $P < .05$.

RESULTS

Clinical Characteristics

A total of 83 surgically treated patients with M-AVM were included in this study. Among all the patients, there were 53 (62.4%) male and 30 (37.6%) female patients, with a mean age of 30.2 (SD, 11.5) years. The mean lesion diameter was 35.1 (SD, 11.4) mm. A deep perforating artery supply was present in 22 patients (26.5%), and deep venous drainage was present in 7 (8.4%) patients. Twenty-five (30.1%) patients had diffuse lesions, and 58 patients (69.9%) had compact lesions. Thirty-three (39.8%) and 41 (49.4%) patients had preoperative hemorrhages and seizures, respectively. The mean LCD_{AI} was 4.3 (SD, 5.1) mm. More general information on the patients is shown in Table 1. After a mean follow-up of 6 months, 26 (31.3%) patients had long-term postsurgical MDs.

Calculation of AI-Based Indicators

In patients with and without postsurgical MDs, the mean FN_{5mm/50mm} was 0.24 (SD, 0.24) and 0.03 (SD, 0.06), the mean FN_{10mm/50mm} was 0.37 (SD, 0.27) and 0.06 (SD, 0.08), the mean FP_{5mm/50mm} was 0.06 (SD, 0.10) and 0.01 (SD, 0.02), and the mean FP_{10mm/50mm} was 0.10 (SD, 0.12) and

Table 1: Baseline characteristics of patients with BAVM

	Total Patients (n = 83)
Age (mean) (yr)	30.2 (SD, 11.5)
Sex (No.) (%)	
Male	53 (62.4)
Female	30 (37.6)
Lobe (No.) (%)	
Frontal	37 (44.6)
Temporal	18 (21.7)
Parietal	18 (21.7)
Occipital	8 (9.6)
Insula	2 (2.4)
Size (mean) (mm)	35.1 (SD, 11.4)
Deep perforating artery supply (No.) (%)	22 (26.5)
Deep venous drainage (No.) (%)	7 (8.4)
Diffuse nidus (No.) (%)	25 (30.1)
LCD _{AI} (mean) (mm)	4.3 (SD, 5.1)
Hemorrhage (No.) (%)	33 (39.8)
Seizure (No.) (%)	41 (49.4)
S-M score (No.) (%)	2.4 (SD, 0.9)

Table 2: Values of the calculated AI-based indicators

	Postsurgical MDs	
	Yes	No
FN _{5mm/50mm} (mean)	0.24 (SD, 0.24)	0.03 (SD, 0.06)
FN _{10mm/50mm} (mean)	0.37 (SD, 0.27)	0.06 (SD, 0.08)
FP _{5mm/50mm} (mean)	0.06 (SD, 0.10)	0.01 (SD, 0.02)
FP _{10mm/50mm} (mean)	0.10 (SD, 0.12)	0.02 (SD, 0.05)

0.02 (SD, 0.05), respectively (Table 2). The values of the AI-based indicators are detailed in the Online Supplemental Data.

Risk Factors for Postsurgical MDs

According to the univariate analysis, significant differences in size ($P = .006$), deep perforating artery supply ($P = .031$), deep venous drainage ($P = .032$), diffuse nidus ($P = .009$), LCD_{AI} ($P = .011$), S-M score ($P = .007$), FN_{5mm/50mm} ($P = .001$), FN_{10mm/50mm} ($P < .001$), FP_{5mm/50mm} ($P = .003$), and FP_{10mm/50mm} ($P < .001$) were found between the patient cohorts with and without postoperative MDs. No significant difference was found in the other factors. According to the multivariate analysis, FN_{10mm/50mm} ($P = .002$) was significantly associated with an increased risk of MDs. No significant difference was found in the other factors (Online Supplemental Data).

Development of the ML Model and Nomogram

FN_{10mm/50mm}, the S-M score, and diffuseness were selected by LASSO regression to establish the ML model. The Pearson coefficient between any 2 of these factors was < 0.35 (Fig 4A). The AUCs of LR, random forest, XGBoost, and support vector machine algorithms in prediction were 0.88 (SD, 0.07), 0.86 (SD, 0.07), 0.86 (SD, 0.07), and 0.77 (SD, 0.12), respectively. The LR with the highest AUC of 0.88 (SD, 0.07) was selected as the algorithm for the ML model (Fig 4B). The ML model resulted in Log $[p(x)/1-p(x)] = -3.516 + (10.046 \times \text{FN}_{10\text{mm}/50\text{mm}}) + (0.279 \times \text{S-M score}) + (1.203 \times \text{Diffuseness})$, where $p(x)$ was the probability of postsurgical MDs. The AUC of the ML model achieved 0.88 (SD, 0.07) with a specificity of 0.74 at a high sensitivity of 0.92 for prediction (Online Supplemental Data).

A nomogram was created combining FN_{10mm/50mm}, the S-M score, and diffuseness based on LASSO regression (Online Supplemental Data). The graphic preliminary score for each of the 3 factors was summed to generate the total score, which indicates the probability of postsurgical MDs. The calibration curve with a Brier Score of 0.131 showed excellent agreement between the predicted probabilities of the nomogram and the observed probabilities for postsurgical MDs (Online Supplemental Data).

Comparison of Different Predictive Methods

The AUC of FN_{10mm/50mm} was 0.86 (SD, 0.08), which was statistically higher than that of LCD_{AI} (0.75 [SD, 0.08], $P = .032$), the S-M score (0.61 [SD, 0.07], $P = .011$), and the HDVL score (0.78 [SD, 0.06], $P = .045$) (Fig 4C). Furthermore, the ML model had a mean AUC of 0.88 (SD, 0.07), slightly improved by 0.02 compared with FN_{10mm/50mm} (Fig 4D). The difference was not significant ($P = .48$). Details of the AUC, accuracy, sensitivity, specificity, precision, recall, and F1 score of different predictive methods are shown in the Online Supplemental Data.

DISCUSSION

To the best of our knowledge, there is no satisfactory method that is able to accurately reflect the 3D spatial relationship between BAVM lesions and the CST to predict postsurgical motor function. In this study, we proposed a novel indicator, FN_{10mm/50mm}, acquired by AI, which has excellent performance for predicting postsurgical MDs. Meanwhile, an ML model consisting of FN_{10mm/50mm}, the S-M score, and diffuseness was developed, having better prediction performance.

Currently, AI is widely used to develop a predictive model for clinical prognosis because of its irreplaceable advantages.¹² First, AI has the ability to analyze diverse data types (eg, demographic data, imaging data, and doctors' free-text notes) and incorporate them into predictions for prognosis.¹³ Second, AI can alleviate the subjectivity and need for expertise in the interpretation of medical images and clinical evaluation. Oermann et al³¹ used an ML model to predict outcomes after radiosurgery for BAVMs and achieved good performance. Third, AI is able to reconstruct the complex geometry of stereoscopic objects captured through sophisticated imaging instruments and calculate the quantitative indicators reflecting their spatial relationship that go beyond those measured by human readers.^{14,15} In this study, we used the FN_{10mm/50mm} automatically calculated by AI to indicate the potential for CST damage due to surgery, which could reflect the spatial relationship of the 3D orientation of the CST and the border of the BAVM lesion. Furthermore, an ML model consisting of FN_{10mm/50mm}, diffuseness, and S-M grading was also developed for prognosis prediction and achieved a better performance.

The distances of 5 and 10 mm were chosen as variables on the basis of our previous studies.^{6,17} In our previous study, we found that for the BAVMs involving the eloquent areas, the LCDs of 5 and 10 mm were the cutoff values for predicting postoperative dysfunction.^{6,17} It suggested that the fibers within 5 or 10 mm from the border of AVM lesions may be most likely to be injured during the resection of AVMs.^{6,17} Meanwhile, according to the anatomic evaluation, for the recruited AVMs adjacent to CSTs

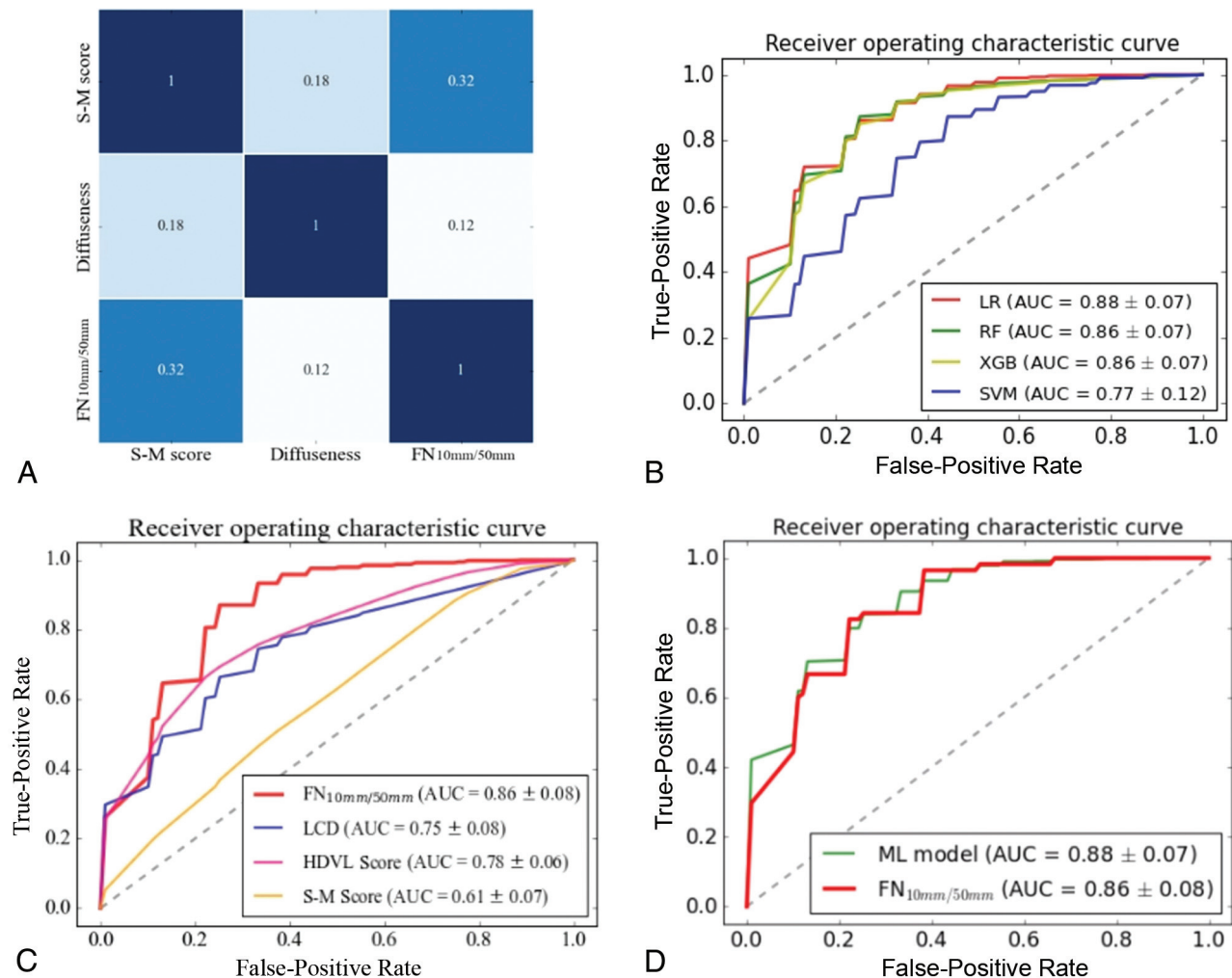


FIG 4. Results of feature screening of the ML model and different predictive methods in predicting postsurgical MDs. A, Result of feature screening with LASSO. The FN_{10mm/50mm}, diffuseness, and S-M score were selected for ML model building. The values in every square indicate the Pearson coefficient between every 2 features. B, ROC curves of different algorithms for the ML model in prediction. C, ROC curves of FN_{10mm/50mm}, LCD_{Al}, S-M score, and HDVL score in prediction. D, ROC curves of FN_{10mm/50mm} and the ML model in prediction. RF indicates random forest; SVM, support vector machine; XGB, XGBoost.

(the shortest distance between AVM lesions and CSTs of ≤ 10 mm), a range of 50mm from the boundary of AVM lesions may include all CST fibers. Therefore, we used 5 mm/50mm and 10 mm/50mm to reflect the potentially injured fiber proportion of CSTs in surgery. In addition, consistent with our previous study, we used the mRS score at 6 months after microsurgery to define the postoperative MDs (mRS > 2).^{5,32} Patients' visual or language deficits were also taken into consideration when grading the mRS.³³

Most interesting, our results showed that the proportion of potentially affected fiber numbers was more effective than the proportion of potentially affected voxel points in predicting long-term postsurgical MDs. This finding suggested that once the CSTs were injured at 1 point, they affected the whole length of the CSTs. We speculated that this situation might be related to the characteristics of nerve axons. According to previous reports, the interruption of axon anatomic continuity will destroy the function of the whole nerve fiber, resulting in partial or total loss of its functions due to the degeneration of nerve fibers distal

to the lesion and eventual death of axotomized neurons.³⁴ This outcome means that neural fibers have less plasticity.³⁵ Therefore, neurosurgeons should be more cautious and alert when crossing a section perpendicular to CSTs in an operation to reduce the possibility of postsurgical MDs.

Our study presents novelty in the algorithm: First, we proposed a method of automatically tracking WM fibers on DTI.³⁶ DTI has been indicated as the only noninvasive way to display WM tracts in vivo and has a unique advantage in identifying and estimating neural fibers at the subcortical level.^{37,38} However, in clinical work, the differences in fiber tracking techniques in diverse clinical centers and the experiences of diverse clinicians may lead to a divergence in the accuracy of fiber tracking among different clinicians.⁸⁻¹¹ In this study, the Constant Solid Angle model and the EuDX algorithm were used to track the CST automatically; this approach could reduce the bias of identification and improve the accuracy and efficiency of fiber tracking in different clinical centers.^{21,39} Second, for our registration method, compared with traditional registration methods, we used a neural

network to perform both affine and deformable registration in a cascade. The deep learning-based registration has higher accuracy and shorter execution time than conventional registration methods.^{40,41} Third, the lesion-dilation operation in our study was implemented using affine transformation, which can accurately control the expansion of the lesions at any distance in all directions. Finally, the LR algorithm with L2 regularization was used for model establishment. The algorithm also unified the LASSO characteristics and ridge characteristics to generate sparse weights to eliminate irrelevant features, possibly preventing overfitting and improving the generalization of the model.⁴²

Our study had several limitations. First, it was conducted in a single center with a limited number of patients due to a relatively small number of this kind of surgery performed. Second, due to its retrospective nature, it may be difficult to avoid information bias, selection bias, and confounding factors. Third, the automated segmentations and fiber tracking may have some errors because of the unclear boundary of AVM lesions themselves, rough surface of the manually section-by-section labeled lesions, and so forth. The model will be trained with more cases to minimize these errors in our following study. Finally, although internal cross-validation was performed, further external validation studies including more patients and from other centers are needed.

CONCLUSIONS

Surgical treatment for M-AVMs is challenging for neurosurgeons. An accurate prediction of the possibility of postsurgical MDs would help guide neurosurgeons in the selection of surgical candidates. In this study, a new AI-based indicator $FN_{10mm/50mm}$ and a corresponding ML model are proposed that both benefit from the assistance of AI techniques. This approach will have potential in the development of predictive software and to assist doctors with different levels of clinical experience in various clinical centers, providing more precise consultation to patients with M-AVM.

ACKNOWLEDGMENTS

We would like to acknowledge the support of Dr Feng Su from the Academy for Advanced Interdisciplinary Studies at Peking University in the study design and article revision. We also thank the professional statistician Anxin Wang from the China National Clinical Research Center for Neurologic Diseases for statistical assistance.

Disclosure forms provided by the authors are available with the full text and PDF of this article at www.ajnr.org.

REFERENCES

- Solomon RA, Connolly ES Jr. Arteriovenous malformations of the brain. *N Engl J Med* 2017;376:1859–66 [CrossRef Medline](#)
- Lawton MT. The role of AVM microsurgery in the aftermath of a randomized trial of unruptured brain arteriovenous malformations. *AJNR Am J Neuroradiol* 2015;36:617–19 [CrossRef Medline](#)
- Wang M, Jiao Y, Zeng C, et al. Chinese Cerebrovascular Neurosurgery Society and Chinese Interventional & Hybrid Operation Society, of Chinese Stroke Association clinical practice guidelines for management of brain arteriovenous malformations in eloquent areas. *Front Neurol* 2021;12:651663 [CrossRef Medline](#)
- Davidson AS, Morgan MK. How safe is arteriovenous malformation surgery? A prospective, observational study of surgery as first-line treatment for brain arteriovenous malformations. *Neurosurgery* 2010;66:498–504 [CrossRef Medline](#)
- Jiao Y, Li H, Fu W, et al. Classification of brain arteriovenous malformations located in motor-related areas based on location and anterior choroidal artery feeding. *Stroke Vasc Neurol* 2021;6:441–48 [CrossRef Medline](#)
- Lin F, Zhao B, Wu J, et al. Risk factors for worsened muscle strength after the surgical treatment of arteriovenous malformations of the eloquent motor area. *J Neurosurg* 2016;125:289–98 [CrossRef Medline](#)
- Jiao Y, Lin F, Wu J, et al. A supplementary grading scale combining lesion-to-eloquence distance for predicting surgical outcomes of patients with brain arteriovenous malformations. *J Neurosurg* 2018;128:530–40 [CrossRef Medline](#)
- Maier-Hein KH, Neher PF, Houde JC, et al. The challenge of mapping the human connectome based on diffusion tractography. *Nat Commun* 2017;8:1349 [CrossRef Medline](#)
- Jbabdi S, Johansen-Berg H. Tractography: where do we go from here? *Brain Connect* 2011;1:169–83 [CrossRef Medline](#)
- Campbell JS, Pike GB. Potential and limitations of diffusion MRI tractography for the study of language. *Brain Lang* 2014;131:65–73 [CrossRef Medline](#)
- Rheault F, Poulin P, Valcourt Caron A, et al. Common misconceptions, hidden biases and modern challenges of dMRI tractography. *J Neural Eng* 2020;17:011001 [CrossRef Medline](#)
- Huang S, Yang J, Fong S, et al. Artificial intelligence in cancer diagnosis and prognosis: Opportunities and challenges. *Cancer Lett* 2020;471:61–71 [CrossRef Medline](#)
- Rudie JD, Rauschecker AM, Bryan RN, et al. Emerging applications of artificial intelligence in neuro-oncology. *Radiology* 2019;290:607–18 [CrossRef Medline](#)
- Wang M, Shen LQ, Pasquale LR, et al. An artificial intelligence approach to assess spatial patterns of retinal nerve fiber layer thickness maps in glaucoma. *Transl Vis Sci Technol* 2020;9:41 [CrossRef Medline](#)
- Gonzalez Izard S, Sanchez Torres R, Alonso Plaza O, et al. Nextmed: automatic imaging segmentation, 3D reconstruction, and 3D model visualization platform using augmented and virtual reality. *Sensors (Basel)* 2020;20:2962 [CrossRef Medline](#)
- Banks JL, Marotta CA. Outcomes validity and reliability of the modified Rankin Scale: implications for stroke clinical trials—a literature review and synthesis. *Stroke* 2007;38:1091–96 [CrossRef Medline](#)
- Jiao Y, Lin F, Wu J, et al. Risk factors for neurological deficits after surgical treatment of brain arteriovenous malformations supplied by deep perforating arteries. *Neurosurg Rev* 2018;41:255–65 [CrossRef Medline](#)
- Jiao Y, Zhang JZ, Zhao Q, et al. Machine learning-enabled determination of diffuseness of brain arteriovenous malformations from magnetic resonance angiography. *Transl Stroke Res* 2022;13:939–48 [CrossRef Medline](#)
- Garyfallidis E, Brett M, Amirbekian B, et al; Dipy Contributors. Dipy, a library for the analysis of diffusion MRI data. *Front Neuroinform* 2014;8:8 [CrossRef Medline](#)
- Garyfallidis E, Brett M, Tsiaras V, et al. Identification of corresponding tracks in diffusion MRI tractographies. In: *Proceedings of the Annual Meeting of the International Society of Magnetic Resonance in Medicine*, Stockholm, Sweden. May 1–7, 2010
- Aganj I, Lenglet C, Sapiro G. ODF reconstruction in Q-ball imaging with solid angle consideration. *Proc IEEE Int Symp Biomed Imaging* 2009;2009:1398–401 [CrossRef Medline](#)
- Chee E, Wu ZJ. Airnet: self-supervised affine registration for 3D medical images using neural networks. 2018. <https://arxiv.org/abs/1810.02583v1>. Accessed January 10, 2021
- Shen Z, Vialard FX, Niethammer MJ. Region-specific diffeomorphic metric mapping. *Adv Neural Inf Process Syst* 2019;32:1098–110 [CrossRef Medline](#)

24. Lee B, Yamanakkanavar N, Choi JY. **Automatic segmentation of brain MRI using a novel patch-wise U-NET deep architecture.** *PLoS One* 2020;15:e0236493 [CrossRef Medline](#)
25. Fonov V, Evans AC, Botteron K, et al; Brain Development Cooperative Group. **Unbiased average age-appropriate atlases for pediatric studies.** *Neuroimage* 2011;54:313–27 [CrossRef Medline](#)
26. Fonov VS, Evans AG, McKinsty RC, et al. **Unbiased nonlinear average age-appropriate brain templates from birth to adulthood.** *Neuroimage* 2009;47:S102 [CrossRef](#)
27. Collins D, Zijdenbos A, Baaré W, et al. ANIMAL+INSECT: improved cortical structure segmentation. *Lecture Notes in Computer Science* 1999;1613:210–23
28. Tibshirani RM. **Regression shrinkage and selection via the lasso.** *J R Stat Soc Series B Stat Methodol* 1996;58:267–88 [CrossRef](#)
29. Krstajic D, Buturovic LJ, Leahy DE, et al. **Cross-validation pitfalls when selecting and assessing regression and classification models.** *J Cheminform* 2014;6:10 [CrossRef Medline](#)
30. Spetzler R, Martin N. **A proposed grading system for arteriovenous malformations.** *J Neurosurg* 1986;65:476–83 [CrossRef Medline](#)
31. Oermann EK, Rubinstein A, Ding D, et al. **Using a machine learning approach to predict outcomes after radiosurgery for cerebral arteriovenous malformations.** *Sci Rep* 2016;6:21161 [CrossRef Medline](#)
32. van Swieten J, Koudstaal P, Visser M, et al. **Interobserver agreement for the assessment of handicap in stroke patients.** *Stroke* 1988;19:604–07 [CrossRef Medline](#)
33. Mascitelli JR, Yoon S, Cole TS, et al. **Does eloquence subtype influence outcome following arteriovenous malformation surgery?** *J Neurosurg* 2018;131:876–83 [CrossRef Medline](#)
34. Navarro X, Vivó M, Valero-Cabré A. **Neural plasticity after peripheral nerve injury and regeneration.** *Prog Neurobiol* 2007;82:163–201 [CrossRef Medline](#)
35. Lundborg G. **A 25-year perspective of peripheral nerve surgery: evolving neuroscientific concepts and clinical significance.** *J Hand Surg Am* 2000;25:391–414 [CrossRef Medline](#)
36. Jiao Y, Lin F, Wu J, et al. **Plasticity in language cortex and white matter tracts after resection of dominant inferior parietal lobule arteriovenous malformations: a combined fMRI and DTI study.** *J Neurosurg* 2020;134:953–60 [CrossRef Medline](#)
37. Ciccarelli O, Catani M, Johansen-Berg H, et al. **Diffusion-based tractography in neurological disorders: concepts, applications, and future developments.** *Lancet Neurol* 2008;7:715–27 [CrossRef Medline](#)
38. Lin F, Wu J, Zhao B, et al. **Preoperative functional findings and surgical outcomes in patients with motor cortical arteriovenous malformation.** *World Neurosurg* 2016;85:273–81 [CrossRef Medline](#)
39. Bihan DL, Mangin J, Poupon C, et al. **Diffusion tensor imaging: concepts and applications.** *J Magn Reson Imaging* 2001;13:534–46 [CrossRef Medline](#)
40. de Vos BD, Berendsen FF, Viergever MA, et al. **A deep learning framework for unsupervised affine and deformable image registration.** *Med Image Anal* 2019;52:128–43 [CrossRef Medline](#)
41. Zhu Z, Cao Y, Qin C, et al. **Joint affine and deformable three-dimensional networks for brain MRI registration.** *Med Phys* 2021;48:1182–96 [CrossRef Medline](#)
42. Owen A. **A robust hybrid of lasso and ridge regression.** *Prediction and Discovery* 2007;443:59–72 [CrossRef](#)