# STATISTICAL ANALYSIS PLAN

**Official Title:** Development and Validation of a Hemorrhage Risk Prediction Model for Moyamoya Disease Based on 5T MRI and Metabolomic Features: A Prospective Cohort Study

**NCT Number:** [To be inserted upon registration]

**Document Date:** August 19, 2026

**Protocol Basis:** Chinese study protocol Version 2.0 dated November 17, 2025

**Ethics Approval Date of Controlling Protocol:** January 13, 2026

**Sponsor:** Beijing Hospital, Beijing, China

**Funding:** Qihang Special Project of the Central High-Level Hospital Clinical Research Project, Beijing Hospital (Grant No.: BJ-2025-222)

**Principal Investigator:** Shaosen Zhang, MD, Attending Neurosurgeon, Department of Neurosurgery, Beijing Hospital

**Document Control:** This Statistical Analysis Plan was finalized on August 19, 2026 for ClinicalTrials.gov on the basis of the Chinese study protocol Version 2.0 dated November 17, 2025. The controlling protocol was approved by the Beijing Hospital Ethics Committee on January 13, 2026 (Approval No. 2025BJYYEC-KY400-02).

# 1. Purpose and Scope

This Statistical Analysis Plan (SAP) describes the pre-specified statistical and analytical methods for the prospective cohort study "Development and Validation of a Hemorrhage Risk Prediction Model for Moyamoya Disease Based on 5T MRI and Metabolomic Features." It operationalizes the endpoint definitions, sample size considerations, data processing, modeling, validation, and performance evaluation procedures described in the study protocol Version 2.0 (November 17, 2025) and applies to all analyses performed within the study.

# 2. Study Overview

This is a single-center, prospective, observational cohort study conducted at Beijing Hospital. Patients with moyamoya disease (MMD) or moyamoya syndrome confirmed by digital subtraction angiography (DSA) are consecutively enrolled and classified, according to their initial presentation, into a hemorrhagic-type group and an ischemic-type group. All participants undergo preoperative 5.0T MRI, routine multimodal imaging, and preoperative blood and urine sampling. Participants are followed for 12 months after enrollment; new-onset intracranial hemorrhage, analyzed per cerebral hemisphere, is the endpoint event. The cohort (approximately 500 participants) is divided at a 7:3 ratio into a training cohort and a validation cohort for model development and validation.

# 3. Endpoints and Definitions

## 3.1 Primary clinical endpoint

**New-onset intracranial hemorrhage within 12 months of enrollment**, analyzed per cerebral hemisphere. Hemorrhage events are adjudicated on the basis of the participant's clinical symptoms and confirmatory head CT or MRI. For each event, the time of occurrence, location, and hematoma size are recorded. Hemispheres that receive revascularization surgery during follow-up are handled according to the protocol: the reason, side, and timing of surgery are recorded in detail, and the subject's follow-up is censored accordingly.

## 3.2 Primary model performance endpoint

**Area under the receiver operating characteristic curve (AUROC)** of the integrated hemorrhage risk prediction model, evaluated in the validation cohort; internal validation is performed with 5 × 5-fold cross-validation on the training cohort, using the best AUROC as the grid-search tuning endpoint.

## 3.3 Secondary endpoints

1. Novel metabolomic biomarkers associated with hemorrhage risk (target: 2–3 biomarkers), cross-validated in superficial temporal artery specimens by immunohistochemistry.
2. 5T MRI markers: quantitative collateral vessel features, hemodynamic parameters (wall shear stress [WSS], wall shear stress gradient [WSSG], oscillatory shear index [OSI], relative residence time [RTT]), vessel wall features, and their associations with hemorrhage events.

3. Deep learning segmentation performance: loss value and Dice coefficient of the VT-UNet collateral vessel segmentation model.
4. Model auxiliary performance metrics: accuracy, recall, precision, and F1 score.
5. Hemorrhage risk stratification system derived from the validated model and core biomarkers.

## 4. Sample Size Considerations

The sample size is based on the primary clinical endpoint. Given the known annual hemorrhage rate of MMD of 17%, and assuming 5 predictor variables (potential markers) with 10 events (hemorrhage cases) required per variable, $5 \times 10 = 50$ events are required. The required sample size is $N = 50 / 17\% \approx 294$. Adjusting for an anticipated 10% loss to follow-up, $N = 294 / (1 - 0.10) \approx 327$. With a 7:3 split into training and validation cohorts, the total sample size is $N = 327 / 0.7 \approx 466$, so approximately **500 participants** will be enrolled.

## 5. Analysis Populations and Units of Analysis

- **Enrolled population**: all participants who meet the eligibility criteria and provide informed consent.
- **Unit of analysis for the hemorrhage endpoint**: the cerebral hemisphere (each participant contributes up to two hemispheres).
- Participants or hemispheres meeting the protocol withdrawal or termination criteria (refusal of participation or follow-up, craniotomy of the observed hemisphere during follow-up, major organ dysfunction, poor compliance, or need for other surgical treatment) will be handled as specified in the protocol; hemorrhage events and revascularization procedures (reason, side, timing) are recorded during follow-up.

## 6. Data Preprocessing and Feature Selection

All preprocessing will be performed in the Python environment:

1. **Normalization by variable type**: categorical variables are converted to one-hot encoding; continuous variables are transformed to approximately normal distributions using the Box-Cox transformation.
2. **Standardization**: all data are standardized using z-score standardization (computed on the training set and applied consistently to the validation set).
3. **Dimensionality reduction**: L1 regularization is applied to the standardized data, with the best ROC-AUC as the primary tuning objective, to select the imaging and metabolomic features most predictive of hemorrhage.
4. Imaging data are standardized by min–max normalization before deep learning model training.
5. Radiomic features that are constant (e.g., all values 0 or 1) or otherwise non-informative are removed before modeling.

## 7. Descriptive and Comparative Statistics

- Continuous variables will be presented as mean ± standard deviation or median (interquartile range), according to distribution; categorical variables will be presented as counts and percentages.

- Two-group comparisons of continuous variables: independent-samples t-test or Mann-Whitney U test.
- Comparisons of means among more than two groups: one-way analysis of variance.
- Comparisons of rates/proportions: chi-square test or Fisher's exact test.
- Correlation analysis: Pearson correlation coefficient (used, inter alia, for associations between metabolomic biomarkers and imaging features, and between perfusion/structural parameters and hemorrhage risk).

## 8. Regression Analysis

Multivariable regression will use the generalized estimating equation (GEE) model or logistic regression, as appropriate, to evaluate the independent associations of candidate imaging and metabolomic markers with hemorrhage events. The AUROC and the Hosmer-Lemeshow goodness-of-fit test will be used to evaluate discrimination and calibration of imaging-based assessments.

## 9. Model Development

1. **Candidate algorithms**: logistic regression, naïve Bayes, support vector machine, neural network, k-nearest neighbor, and decision tree will first be evaluated. If performance is unsatisfactory, ensemble learning algorithms will be tested.
2. **Hyperparameter tuning**: grid search with exhaustive enumeration to identify the optimal parameter combinations; the best AUROC (from internal validation) is the tuning endpoint.
3. **Feature inputs**: clinical baseline data, 5T MRI features (collateral vessel morphology and counts, hemodynamic parameters, vessel wall features, radiomic features), and metabolomic biomarkers.
4. **Deep learning components**: a VT-UNet 3D segmentation network extracts the vascular skeleton from TOF-MRA, and a Swin Transformer architecture extracts vessel features and generates predictions.

## 10. Model Validation and Performance Assessment

- **Internal validation**: 5 repetitions of 5-fold cross-validation (5 × 5-fold) on the training cohort.
- **External validation**: evaluation of the final model in the independent validation cohort.
- **Primary performance metric**: AUROC.
- **Auxiliary metrics**: accuracy, recall, precision, and F1 score.
- **Calibration**: Hosmer-Lemeshow goodness-of-fit test.
- **Stratification**: the validated model and core biomarkers will be used to construct a hemorrhage risk stratification system; risk discrimination across strata will be evaluated.

## 11. Imaging Analysis Quality Control

- **Vessel wall annotation**: manual annotation on high-resolution vessel wall imaging (Slicer 4.11) by two independent readers in a blinded manner; annotations are accepted when the intraclass correlation coefficient (ICC) exceeds 0.8; otherwise, annotations are corrected. The principal investigator's annotations serve as the study data after two independent experts (each with more than 20 years of cerebrovascular and imaging diagnosis experience) complete quality checks and corrections.

- **Radiomic feature extraction**: performed with the PyRadiomics package; approximately 1,790 features are expected (356 first-order; 14 shape; 432 GLCM; 312 GLSZM; 312 GLRLM; 90 NGTDM; 274 GLDM; including 744 wavelet features).
- **Segmentation quality**: min–max normalized input; model training monitored with the cross-entropy-based loss on the training set and Dice coefficient on the test set; convergence epochs recorded.

## 12. Metabolomic Data Analysis

- Preoperative blood (plasma) and urine samples undergo high-throughput metabolomic measurement of metabolic indices, mainly including amino acids.
- Differential metabolic profiling between hemorrhagic-type and ischemic-type MMD, and association of metabolic features with imaging features (collateral vessel morphology, hemodynamics, vessel wall characteristics), will be analyzed.
- Candidate biomarkers will be cross-validated in superficial temporal artery specimens by immunohistochemistry.
- Candidate biomarkers (target: 2–3) will be evaluated for their value in hemorrhage risk assessment, risk stratification, and prognosis, and will be incorporated into the prediction model.

## 13. Statistical Software

Statistical analyses will be performed using SAS 9.1 (SAS Institute, Cary, NC), SPSS 17.0 (SPSS Inc., Chicago, IL), and MedCalc 12.3 (MedCalc Software); imaging processing, radiomics, and deep learning/machine learning pipelines will be implemented in Python.

## 14. General Conventions

- Hypothesis tests are two-sided; a significance level of $\alpha = 0.05$ will be used, and 95% confidence intervals will be reported where appropriate.
- Patterns of missing data will be summarized descriptively; primary analyses will use available data.
- No formal interim analyses for efficacy are planned for this observational study; data and imaging quality are reviewed monthly by the research team.
- Analysis code and datasets will be versioned and archived together with the results to ensure reproducibility.