

## **Prediction of etiology and prognosis based on hematoma location of spontaneous intracerebral hemorrhage**

Jingjing Liang, MS<sup>1,2†</sup>, Weixiong Tan, MS<sup>3†</sup>, Shijia Xie, MS<sup>4</sup>, Lijuan Zheng, MD<sup>5</sup>, Chuyan Li, BS<sup>1</sup>, Zhong Yi, MS<sup>1</sup>, Jianrui Li, MS<sup>1</sup>, Changsheng Zhou, MS<sup>1</sup>, Zhiqiang, Zhang, MD<sup>1</sup>, Zhen Zhou, MS<sup>3</sup>, Ping Gong, MS<sup>3</sup>, Xingzhi Chen, MS<sup>3</sup>, Longjiang Zhang, MD<sup>1</sup>, Xiaoqing Chen, MD<sup>1\*</sup>, Qirui Zhang, MD<sup>16\*</sup>, Guangming Lu, MD<sup>1\*</sup>

### **Author affiliations:**

<sup>1</sup> Department of Radiology, Jinling Hospital, Medical School of Nanjing University, Nanjing 210002, China

<sup>2</sup> Department of Radiology, Lu'an Hospital of Anhui Medical University, Lu'an 237005, China

<sup>3</sup> Deepwise AI Lab, Beijing Deepwise & League of PhD Technology Co.Ltd, Beijing 100080, China

<sup>4</sup> Departments of Radiology, Yijishan Hospital of Wannan Medical College, Wuhu 241001, China.

<sup>5</sup> Jiangsu Key Laboratory of Molecular and Functional Imaging, Department of Radiology, Zhongda Hospital, Medical School of Southeast University, Nanjing 210009, China.

<sup>6</sup> Farber Institute for Neuroscience, Department of Neurology, Thomas Jefferson University, Philadelphia, Pennsylvania, USA

<sup>†</sup>, These authors contribute equally to this work.

### **Corresponding Authors:**

Guangming Lu, E-mail: [cjr.luguangming@vip.163.com](mailto:cjr.luguangming@vip.163.com), Qirui Zhang, E-mail: [Qirui.Zhang@jefferon.edu](mailto:Qirui.Zhang@jefferon.edu), Xiaoqing Chen, E-mail: [rabbitkiller80@126.com](mailto:rabbitkiller80@126.com)

**Running head:** Etiology and prognosis of sICH by location

**Number of characters in the title and running head :** 136

**Total article word count:** 3881

**Figures:** 5

**Table:** 1

## **Abstract**

### **Background**

The characteristics of the hemorrhagic location of spontaneous intracerebral hemorrhage (sICH) is pivotal for both identifying its etiology and prognosis. While empirical conclusions have been obtained in clinical practice, a comprehensive and quantitative modeling approach has yet to be thoroughly explored.

### **Methods**

We employed lesion-symptom mapping to extract the location features of sICH. We registered patients' non-contrast computed tomography image and hematoma masks with standard human brain templates to identify specific affected brain regions. Then, we generated hemorrhage probabilistic maps of different etiologies and prognoses. By integrating radiomics and clinical features into multiple logistic regression models, we developed and validated optimal etiological and prognostic models across three centers, comprising 1162 sICH patients.

### **Results**

Hematomas of different etiology have unique spatial distributions. Location features robustly categorized the etiology sICH (mean area under the curve (AUC) = 0.825) across different datasets), yielding clear add on value in models (fusion model mean AUC = 0.915) compared to clinical features (mean AUC = 0.828). In the prognostic analysis, patients with more extensive hematoma had a worse prognosis, the location (mean AUC = 0.762) and radiomic features (mean AUC = 0.837) also providing reliable add on value effects (fusion model mean AUC = 0.873) compared to clinical features alone (mean AUC = 0.771).

### **Interpretation**

Our results show that location features were more intrinsically robust, generalizable relative, strong interpretability to the complex modeling of radiomics, our approach demonstrated a novel interpretable, streamlined, comprehensive etiologic classification and prognostic prediction framework for sICH.

## Introduction

Spontaneous intracerebral hemorrhage (sICH) is characterized by the extravasation of blood within the intracranial space due to the rupture of cerebral vessels, encompassing brain parenchymal hemorrhage, intraventricular hemorrhage, and subarachnoid hemorrhage. The heightened mortality rate and unfavorable prognosis associated with sICH poses a significant challenge for clinicians and patients. Non-contrast computed tomography (NCCT) imaging is the preferred method of diagnosis<sup>1</sup>. Within the context of sICH (nontraumatic), the identification of disparate etiologies gives rise to distinct and varied treatment indications. For instance, aneurysms may require interventions such as embolism or craniotomy, while hypertensive intracerebral hemorrhage is typically addressed through conservative or decompressive measures. Upon hospital admission the primary imperative for sICH patients is prompt determination of the underlying cause<sup>2</sup>. Survivors of sICH frequently encounter severe complications, enduring neurological impairments, even debilitating disabilities, if not treated promptly and correctly<sup>3,4</sup>.

In their quest to identify the underlying etiology of sICH, clinicians routinely incorporate brain spatial location into their assessments of hemorrhagic impact. For instance, aneurysmal hemorrhage tends to manifest in the subarachnoid space<sup>5</sup>, hypertensive hemorrhage predominantly presents in the deep nucleus region<sup>6</sup>, and vascular malformations are commonly observed in cortex<sup>7</sup>. And in terms of prognostic prediction, hyperglycemia, hypertension, advanced age, and hematoma location characteristics, all have been empirically identified as crucial factors that influence outcomes<sup>3,8</sup>. Accordingly, the localization of hematoma is of paramount importance for both etiologic and prognostic purposes.

Current analyses of sICH location often focus on specific cerebral regions, lacking a comprehensive investigation across the entire anatomical space of the brain. Lesion-symptom

mapping, a population-based analysis method, provides standardized location features of lesions by calculating normalized lesion locations. In recent years, this method has been used in studies of ischemic stroke and glioma, yielding significant advances in the precision and full context of localization<sup>9-12</sup>. Utilizing this approach, probabilistic maps can be drawn, and the impact of clinical variables can be statistically quantified, enabling a quantitative analysis of location features based on the visualized distribution of hematoma effects throughout the entire brain<sup>13, 14</sup>.

Based on these insights, we posit that the location features of hematoma bear a close association to etiology and possess predictive value for determining prognosis. In this project, we employed lesion-symptom mapping to extract the location features of sICH and generated probability maps of etiology and prognosis for each and every brain region. We then combined these radiomic metrics with clinical features (e.g. hypertension, sex, and chronological age) to train and multiple models of etiology and prognosis, utilizing model comparison statistics to determine final, optimal models. Our study encompassed a comprehensive cohort of 1162 sICH patients across three medical centers, ensuring a robust, diverse, and generalizable dataset for our analyses. This study's fundamental aim is to expand our understand how lesion location features affect the etiology and prognosis of sICH, and to provide a quantitative methodology for efficiently making these determinations.

## **Materials and Methods**

### **Study Population and clinical Information**

This retrospective study encompassed the collection of sICH patients from three distinct medical centers, Jinling Hospital (dataset 1, January 2015 to December 2019), Yi Jishan Hospital (dataset 2, January 2018 to December 2020), and Nanjing First Hospital (dataset 3,

January 2019 to December 2020). The confirmation of intracranial hemorrhage was achieved through NCCT. Patients were excluded if they presented with secondary hemorrhage resulting from head trauma, hemorrhagic transformation of ischemic infarction, brain tumors, or exhibited abnormalities in blood coagulation, liver, kidney function, or were subject to drug-induced cerebral hemorrhage. Supplementary Figure 1 illustrates the retrospective inclusion and exclusion flowchart in detail.

The following baseline clinical characteristics of patients were extracted from medical records: chronological age, sex, hypertension (defined as systolic blood pressure  $\geq 140$  mm Hg and/or diastolic blood pressure  $\geq 90$  mmHg, or self-reported use of antihypertensive medication within the past 2 weeks), diabetes mellitus (fasting plasma glucose concentration  $\geq 7.0$  mmol/L or exceeding 11.1 mmol/L 2 hours post-meal), hyperlipidemia (abnormal blood lipids were detected in serum), history of previous stroke, smoking, alcohol consumption, Glasgow Coma Scale (GCS) score at admission, and the known time interval from onset to baseline CT if witnessed or self-reported by the patient. All demographic and clinical information are shown in Table 1.

The etiologies of sICH were categorized into four classifications: hypertension, aneurysm rupture, vascular malformation, and clinically unclear diagnosis by discharge diagnosis. Modified Rankin Scale (mRS)<sup>15</sup> scores were determined based on clinical examinations or telephone follow-ups at twelve months post sICH event. Consistent with prior literature, we defined good outcome as "mRS = 0-2", poor outcome as "mRS = 3-6"<sup>16, 17</sup>.

Ethics review and institutional review board approval in each center were both obtained. Informed consents from participating patients were exempted because of the retrospective nature of this study.

## **Image Acquisitions and Preprocess**

NCCT covering the entire brain was acquired with a slice thickness of less than 10.0 mm in all patients. All images underwent meticulous quality checks to eliminate those displaying significant motion artifacts. The hematoma areas were delineated from the NCCT images utilizing artificial intelligence brain hemorrhage auto-segmentation software (Dr. Wise TM software available at <http://label.deepwise.com/>). A physician's review of the segmentation results revealed no obvious segmentation errors.

## **Lesion-symptom mapping and location features**

The NCCT scans and hematoma areas were subjected to spatial normalization and alignment with the standard Montreal Neurological Institute (MNI) template. The normalization process employed the CT normalization function offered by Standard Space Lesion Symptom Mapping (SPLSM , [https://github.com/JLhos-fmri/SPLSM\\_version1.0](https://github.com/JLhos-fmri/SPLSM_version1.0)) toolkit, which was based on the Clinical Toolbox of SPM12 (<https://github.com/spm/spm12>), utilizing a cost function masking method. Based on normalized hemorrhagic areas, the frequency occurrence of hematoma at each voxel was computed, creating a voxel-based probabilistic map for each patient (Figure 1A).

This yielded a location feature set that quantified hemorrhage involvement of intracerebral regions in the parenchyma (cortical, white matter, and sub-cortical), subarachnoid space, and ventricles. A hemorrhagic event was considered present if more than 20% of the voxels in a given region showed a positive indication of a hemorrhage.

The brain parenchyma analyses utilized the Johns Hopkins University (JHU) template, comprised of 178 annotated brain regions, encompassing 68 white matter and 110 gray matter regions (Figure 1B). The subarachnoid space (10 regions) and ventricles (4 regions) were



delineated using an in-house manually crafted template. List of intracranial regions see Supplementary Table 1.

In the etiological prediction analysis, 192 regions were sorted into 5 intracranial structures: ventricle, subcortical, white matter, cortical, and subarachnoid space. In each intracranial structure, a paired one-way ANOVA was used to calculate whether there was a significant difference in the rate of involvement in the region of the hemorrhage for the three etiologies. In the prognostic analysis, the rate of hemorrhage at 192 regions were compared by Chi-square test.

### **Radiomics features**

Radiomics features were extracted by the python package PyRadiomics (Version 3.0.1; <https://pypi.org/project/pyradiomics>). 1454 radiomic features were extracted for each hematoma region. It included first-order features, shape-based features, gray cooccurrence matrix, gray dependence matrix, gray running length matrix, gray size area matrix, and a neighborhood gray-tone difference matrix (Figure 1B). Given the possibility of multiple regions of hematoma of a portion of patients, it was essential to assign weights to the radiomic features of each region, creating a patient-level feature set. To achieve this, we used a deep learning multi-locus attention model algorithm to generate adaptive weights of focus-level features. The model underwent training and iteration within a training cohort using 5-fold cross-validation. The adjustment of model parameters occurred based on prognostic and error considerations (loss function). Subsequently, the radiomic features from multiple regions were amalgamated into patient-level radiological group features, as detailed in the Supplementary Method and Supplementary Figure 2. After adaptive weighting, 1454 radiomics features based on patient level were obtained.

### **Model development for etiology and prognosis**

The models developed in this study served dual purposes: classifying sICH etiology and predicting prognosis. In the actual clinical determination of etiology, doctors often judge by clinical characteristics and the location of hemorrhage. Therefore, we only included cases with known clinical and hemorrhage location features in our etiology classification models (n.b., by this criterion 53 patients with an unknown etiology at discharge were excluded). A multiclassification strategy was also used to predict whether the etiology is hypertension, ruptured aneurysm, or vascular malformation. Three models were constructed: clinical features only, location features only, and a fused location-clinical model (referred to as the loc-cli model).

In the prognostic task, clinical and hemorrhage location features, along with the above 1454 radiomic features, were utilized. Based upon these feature sets used to predict good (mRS = 0-2) or bad (mRS = 3-6) outcome<sup>16,17</sup>. Four models were tested: clinical, location, radiomics, and a combined or fused clinical-location-radiomics model (referred to as the loc-cli-rad model) (see Figure 1C). Dataset 1 was utilized for model training, while datasets 2 and 3 were reserved for independent testing and validation of the models.

Logistic regression in scikit-learn (<https://scikit-learn.org>) toolkit was used to construct the model. In training set, Spearman correlations were used to eliminate feature redundancy. Correlation coefficients greater than 0.8 between two or more features are retained for only one of them. L1 feature selection methods were adopted for reducing dimensionality in our models. Fivefold cross-validation and a grid search with F1 score as the optimization goal were implemented on the training set, thereby tuning the model hyperparameters and predictive performances. The fusion model utilized “model fusion” in the training set, a technique that employed an average information strategy to combine the prediction scores generated after running each individual model during cross validation.

In modeling etiology, we plotted confusion matrices and receiver operating characteristic (ROC) curves to evaluate the models. To obtain ROC curves from these tri-categorical models, we reduced the multiclassification problem to three two-level classification problems, and averaged the corresponding true positive rate at the same false positive rate on the ROC curve and plotted the macro average ROC curve using the average of false positive rate and true positive rate.

In the prognosis prediction model, ROC curves were used for model evaluation. To compare the accuracy between models, we calculated the difference in the area under the curve (AUC) values between the fusion model and the other models. Subsequently, we verified the significance of these differences using the Delong test (Figure 1D).

## **Results**

This study involved a total of 1162 patients (655 males, 507 females), with a mean age of  $57.65 \pm 13.19$  years. We categorized sICH etiology as hypertension, aneurysm, vascular malformation, or unknown, with respective counts of 563, 435, 111, and 53 patients. The clinical functional prognosis was stratified by outcome, with 721 and 441 patients for good and poor outcomes, respectively. Table 1 provides a comprehensive summary of sample demographic and clinical characteristics. Detailed demographics and clinical characteristics for each dataset are provided in Supplementary Results, Supplementary Tables 2 and 3.

### **Association between hematoma location and etiology**

The distribution of hematoma by etiology (hypertension, aneurysm, vascular malformation) is depicted in Figure 2A. Hypertensive hematoma manifested primarily in deep brain, particularly the basal ganglia area. Conversely, hematoma resulting from aneurysm rupture were concentrated primarily in the subarachnoid space, encompassing regions such as the lateral fissure cistern and the anterior longitudinal fissure cistern in subarachnoid area.

Notably, vascular malformation-related hematoma demonstrated a more even and widespread distribution, with a higher likelihood in the ventricles. The proportion of individuals affected in each intracranial region based on intracranial structure (ventricle, subcortical, white matter, cortical, and subarachnoid space) is summarized in Figure 2B. Significant differences were observed in the distribution of cerebral hemorrhage for the different etiologies. A noteworthy result is that vascular malformations and ruptured aneurysms appeared likely to involve the ventricles. Hypertension and vascular malformations tended to involve the brain parenchyma, but hypertension tended to be sub-cortical, whereas vascular malformations tended to be cortical. In addition to ruptured aneurysms, which are most likely to involve the subarachnoid space, vascular malformations also involved the subarachnoid space. The pattern of hematoma distribution was consistent across datasets 1, 2, and 3 (Supplementary Figure 3, 4, 5).

### **Association between hematoma location and prognosis**

Probability plots show that, in general, patients with a poorer prognosis have a wider distribution of hematoma in subarachnoid space, parenchyma, and ventricles (Figure 3A). When the hematoma was located in the radiation crown, marginal gyrus, cerebellum, bilateral lateral ventricle and fourth ventricle, the prognosis was generally poor. In contrast, patients with hematoma in the subarachnoid area had a relative good prognosis (Figure 3B). Detailed results for the separate datasets (1, 2, and 3) are provided in Supplementary Figure 6, 7, 8.

### **Etiology classification**

In the etiological classification task, the clinical model demonstrated an average AUC of 0.828 (range: 0.806 ~ 0.848 cross different dataset), while the location model achieved an AUC of 0.825 (range: 0.807 ~ 0.842). Notably, the fusion model exhibited a higher AUC of

0.915 (range: 0.909~0.923). See Supplementary Table 4 for details of model performances in each dataset. The confusion matrix in Figure 4A was employed to assess the mean accuracy of each model cross datasets. The location models achieved overall accuracies exceeding 86% for hypertension and aneurysm, with suboptimal performance (12.6%) for the prediction of vascular malformation. The fusion model, however, improved the diagnostic ability of all three etiological types (hypertension = 94.7%, aneurysm = 91.5%), especially for vascular malformations (42.3%). Detailed results of each dataset are provided in Supplementary Figure 9. Delong tests confirmed that the fusion model consistently outperformed the clinical models, showing significantly higher AUC in all datasets (all  $p < 0.001$ ) (Supplementary Table 10), indicating that location features provided significant predictive value (Supplementary Table 5).

In the clinical model, hypertension, male sex, and history of stroke displayed the strongest positive contribution to predicting hypertensive sICH, whereas these features showed the strongest negative contribution in association with aneurysms, indicating that the presence of these features made it less likely that an aneurysm was present. In vascular malformation, male sex had a positive contribution, whereas hypertension and age had a negative contribution (Supplementary Figure 11). In the location models, hypertension was strongly associated with deep regions such as the lentiform nucleus, striae terminalis, and hippocampus. In contrast, hemorrhage of the superior frontal gyrus and anterior longitudinal fissure contributed negatively to the model, reducing the likelihood that hypertensive sICH was present. Aneurysms ruptured mainly in the subarachnoid space, especially in the lateral fissure cistern and anterior longitudinal fissure cistern. Hemorrhage in the fornix and cerebellum, however, contributed negatively to the model, indicating that aneurysmal cerebral hemorrhage were less likely to reside in these areas. Vascular malformations were

primarily distributed in the posterior cingulate gyrus and superior frontal gyrus, whereas they were rare in the lenticular fasciculus and dorsal anterior cingulate regions (Figure 4B).

### **Prognosis prediction**

In the prognostic prediction task, the clinical model demonstrated an average AUC of 0.771(range: 0.762 ~ 0.799), while the radiomics model achieved an average AUC of 0.837 (range: 0.831 ~ 0.840). In contrast, the location model exhibited an average AUC of 0.760 (range: 0.785 ~ 0.725). The fusion model outperformed all other models with an average AUC of 0.873 (range: 0.895 ~ 0.850). ROC of each dataset is shown in Figure 5 Detailed model performances are shown in Supplementary Table 4. Delong tests confirmed that the fusion model consistently outperformed the clinical models, displaying significantly higher AUC in all datasets (all  $p < 0.01$ ), indicating that location and radiomics features consistently provided predictive value above and beyond standard clinical predictors (see Supplementary Table 5).

The most important features in the clinical model were lower GCS, age and presence of hypertension. In contrast, in the radiomics model, higher shape mesh volume original, lower GLDM, and lower NGTDM were the most important (see Supplemental Figure 12 for a complete list of feature weights). In the location model (see Figure 5D), the most important features making a positive contribution to poor prognosis were right inferior frontal gyrus orbitalis, right pontine, left posterior insula, and right cingulate gyrus (n.b., the suprasellar cisterna negatively contributed to poor prognosis).

### **Discussion**

This study provides a novel, interpretable, streamlined, comprehensive framework for etiologic classification and prognostic prediction of sICH patients after NCCT examination. In particular, we found that the hematoma location features in our models accurately

described the relationship between spatial hematoma distribution and etiology or prognosis, with these relationships consistent with a priori clinical knowledge. This consistency suggested that imaging biomarkers constructed using hematoma location features are highly interpretable. As our imaging markers were not affected by different centers or different NCCT scanners, they appeared robust, exhibiting a high degree of generalizability. Lastly, our data showed our imaging biomarkers provided predictive above and beyond standard clinical characteristics. Accordingly, our data not only enhances our understanding of the condition, but showed the potential to improve clinical decision-making that would be broadly applicable.

Clinical features have always been important information for clinicians to judge the cause of intracerebral hemorrhage and clinical treatment<sup>18</sup>. In our clinical model, as expected, hypertension was crucial for hypertensive sICH, as hemorrhage often occurs in the setting of poor blood pressure control<sup>19</sup>. For aneurysm patients, female showed higher risk than male, it is consistent with previous studies, noting that additional hemorrhage from a ruptured aneurysm is more common in females<sup>20, 21</sup>. Our results also showed age of vascular malformations patients is younger than hypertension and aneurysm patients. This is due to the fact that brain hemorrhages caused by congenital dysplasia usually precede those caused by structural degeneration of vessels<sup>22</sup>.

On the basis of our clinical model, we utilized lesion-symptom mapping for quantitative and precise specification of hematoma location, confirming that location features have a significant additive effect relative to standard clinical information for the prediction and classification of etiology. Our data showed that hypertensive hemorrhage manifested primarily in deep brain regions, specifically the lentiform nucleus, striae terminalis, and

hippocampus. This is consistent with the previous clinical knowledge, which contends that hypertensive hemorrhage is primarily located in the basal ganglia region<sup>23</sup>. Due to hemodynamic sensitivity, the cerebral arterial circle or deep penetrating arteries may exhibit heightened responsiveness to systemic blood pressure changes compared to cortical arteries<sup>3, 24</sup>, thereby increasing the likelihood of bleeding<sup>25, 26</sup>. Subarachnoid hemorrhage often results from rupture with sICH occurring in approximately 20-30% of aneurysm patients<sup>27</sup>. The cistern in which the hemorrhage is likely to be located is an important clue to finding the responsible aneurysm<sup>28, 29</sup>. In this light, it is important to note that our etiologic classification model found evidence that involvement of the lateral fissure cistern and the anterior longitudinal fissure cistern were key in predicting aneurysm rupture. Our evidence also suggesting that the corpus callosum and lenticular fasciculus are also important in terms of location for regions outside the subarachnoid space. Vascular malformations are congenital vascular abnormalities<sup>30</sup>. Hemorrhages associated with vascular malformations can be distributed across the white matter, cortex, subcortex, subarachnoid space, and other regions, indicating that a broad range of potential bleeding sites. Our data demonstrate the unique exist in the posterior cingulate gyrus and superior frontal gyrus.

Numerous predictors associated with an elevated risk of adverse prognosis following sICH have been identified, encompassing clinical factors such as age, elevated blood pressure and blood glucose levels, as well as hematoma location<sup>31, 32</sup>. Our clinical model broadly align with these established associations. Compared with the young, older patients have vascular distortion, accompanied with atherosclerosis<sup>33</sup>, and this can co-exist with other diseases, all factors that can indirectly increase disability and mortality. Previous studies have found that uncontrolled blood pressure will increase the risk of stroke recurrence, and controlling the stability of blood pressure is crucial to preventing such events. Our data showed that GCS



score is the strongest prognostic predictor<sup>34</sup> of sICH, with a low GCS score indicating severe disturbance of consciousness and a poor prognosis<sup>35</sup>.

When including location and radiomic features in our prognostic models, we found they provided predictive value significantly above what was provided by these more standard clinical features. In our prognostic models, brain regions located in areas of the cortex contributed more to the prognostic model, whereas subcortical regions contributed less to the prognostic prediction, which is consistent with previous studies that have concluded that lobar hemorrhages have a poor prognosis relative to nonlobar hemorrhages<sup>32,36</sup>. The regions most associated with poor outcomes predominantly involve the inferior frontal gyrus, pontine, and insula. Damage to the inferior frontal gyrus has been shown to lead to language deficits<sup>37</sup>, may further hinder the recovery of physical function<sup>38</sup>. Pontine hemorrhage is particularly grave in terms of prognosis. After pontine hemorrhage, various oculomotor abnormalities (such as pupil apex, medial longitudinal tract syndrome, ocular deviation and nystagmus) may be observed<sup>39</sup>. The insula is a key node within the supratentorial swallowing network. Previous studies have shown a strong association between lesions of the right insular cortex and a severe symptom of dysphagia<sup>40,41</sup>. At the same time, injuries involving the insula have also been found to increase cardiogenic mortality<sup>42</sup>.

Among the radiomic features, the shape mesh volume of the hemorrhage is positively correlated with prognosis, and this metric allows for a more precise characterization of the volume of irregular hemorrhage<sup>43</sup>. This is consistent with the common knowledge that the larger the hematoma, the worse the prognosis<sup>44, 45</sup>, and is also in line with our location-based analysis. The NGTDM features represent the contrast between voxels and neighboring voxels<sup>46</sup>, is used to quantify the intensity difference between neighboring voxels. The contrast which negatively contributes to prognosis in our data. This suggests that if the

density is uneven due to the large volume of the hematoma, slow absorption or secondary bleeding, the contrast will be relatively low, which indicating a poor prognosis<sup>47</sup>.

At the level of feature construction, this study incorporates both location features and radiomics features. Since location features are determined to describe the entire hemorrhagic region, they are not affected by NCCT scanning machines and parameters and are naturally compatible with the analysis of multiple hemorrhagic foci. In contrast, radiomics affected by the NCCT scanning machine and parameters<sup>48</sup> and are difficult to deal with the combined analysis of multiple hemorrhagic regions. To address this issue, our approach incorporates innovative techniques for radiomics feature extraction. Through multiple instance learning, we assign weights to the cross-focus of each patient, ensuring the preservation of both global information pertaining to the patient's lesions and the local information associated with each hemorrhagic location<sup>49</sup>. Even so, we believe that location features are intrinsically robust and generalizable relative to the complex modeling of radiomics, especially the high degree of interpretability that coincides with a priori knowledge.

Our study has several limitations. First, the study is a retrospective study, and the possibility of sample bias cannot be ruled out, e.g., the etiologic and prognostic distributions of the included samples may differ from the real situation. Second, we categorize the etiology of hemorrhage in general terms, which may reduce the applicability of our results in some cases affects the integrity of the etiological classification results. For example, sICH caused by amyloidosis is difficult to diagnose clinically, and there was no clear etiologic classification in our study.

In conclusion, our results show that the location features of hematoma are stable and reliable indicators in sICH etiology and prognosis prediction, and the combination of the radiomics features can effectively improve the prediction efficiency and overcome the defects of the

instability of the radiomics features. Provide an interpretable, comprehensive etiologic classification and prognostic prediction framework for sICH in a streamlined manner.

### **Acknowledgements**

We thank Prof. Joseph I. Tracy from Thomas Jefferson University for language editing and helpful discussions. This study received funding from National Natural Scientific Foundation of China (Grant Nos. 82271983 to X.C., 82127806 to G.L., 81790653 to G.L., and 81701680 to Q.Z.), National Key R&D Program of China (Grant Nos. 2020AAA0109505 to G.L., 2018YFA0701703 to Zhiqiang Zhang).

### **Author Contributions**

G.L., Q.Z., and X.C. contributed to the conception and design of the study; Q.Z. J.L., W.T. contributed to acquisition and analysis of data, drafting the text or preparing the figures; S.X., L.Z., C.L., Y.Z., J.L., C.Z., Zhiqiang Zhang., Zhen Zhou, P.G., X.C., L.Z. contributed to the acquisition and analysis of data

### **Conflicts of Interest**

Conflict of interest Weixiong Tan, Zhen Zhou, Ping Gong and Xingzhi Chen are employees of Deepwise Inc. The remaining authors have no conflicts of interest

### **Data Availability**

The data that support the findings of this study are available from the corresponding author on reasonable request.

### **References**

1. Greenberg SM, Ziai WC, Cordonnier C, et al. 2022 Guideline for the Management of Patients With Spontaneous Intracerebral Hemorrhage: A Guideline From the American Heart

- Association/American Stroke Association. *Stroke* 2022;53(7):e282-e361. doi:10.1161/STR.0000000000000407
2. Winstein CJ, Stein J, Arena R, et al. Guidelines for Adult Stroke Rehabilitation and Recovery: A Guideline for Healthcare Professionals From the American Heart Association/American Stroke Association. *Stroke* 2016;47:e98-e169.
  3. Schrag M, Kirshner H. Management of Intracerebral Hemorrhage: JACC Focus Seminar. *J Am Coll Cardiol* 2020;75:1819-1831.
  4. AA. R. Prognosis after ICH: A necessity despite uncertainties. *Neurology* 2016;86(20):1854-1855.
  5. Vergouwen MD, Vermeulen M, van Gijn J, et al. Definition of delayed cerebral ischemia after aneurysmal subarachnoid hemorrhage as an outcome event in clinical trials and observational studies: proposal of a multidisciplinary research group. *Stroke* 2010;41:2391-2395.
  6. An SJ, Kim TJ, Yoon BW. Epidemiology, Risk Factors, and Clinical Features of Intracerebral Hemorrhage: An Update. *J Stroke* 2017;19:3-10.
  7. El-Abtah ME, Kashkoush A, Achey R, Patterson T, Moore NZ, Bain MD. Diagnostic yield of cerebral angiography for intracranial hemorrhage in young patients: A single-center retrospective analysis. *Interv Neuroradiol* 2023;15910199231152505.
  8. Johansson K, Jansson JH, Johansson L, Wiklund PG, Nilsson TK, Lind M. D-Dimer Is Associated With First-Ever Intracerebral Hemorrhage. *Stroke* 2018;49:2034-2039.
  9. Weaver NA KH, Aben HP, et al. Strategic infarct locations for post-stroke cognitive impairment: a pooled analysis of individual patient data from 12 acute ischaemic stroke cohorts. *Lancet Neurol* 2021;20(6):448-459.
  10. Liu X, Zhang Q, Li J, et al. Coordinatized lesion location analysis empowering ROI-based radiomics diagnosis on brain gliomas. *Eur Radiol* 2023.
  11. Kim NY, Taylor JJ, Kim YW, et al. Network Effects of Brain Lesions Causing Central Poststroke Pain. *Annals of Neurology* 2022;92:834-845.
  12. Ding L, Liu H, Jing J, et al. Lesion Network Mapping for Neurological Deficit in Acute Ischemic Stroke. *Annals of Neurology* 2023;94:572-584.
  13. Karnath HO, Fruhmann Berger M, Kuker W, Rorden C. The anatomy of spatial neglect based on voxelwise statistical analysis: a study of 140 patients. *Cereb Cortex* 2004;14:1164-1172.
  14. Wang J, Gu M, Xiao L, et al. Association of Lesion Location and Fatigue Symptoms After Ischemic Stroke: A VLSM Study. *Front Aging Neurosci* 2022;14:902604.
  15. Quinn TJ, Dawson J, Walters MR, Lees KR. Reliability of the modified Rankin Scale: a systematic review. *Stroke* 2009;40:3393-3395.
  16. Sharma M HR, Connolly SJ, et al. Strategic infarct locations for post-stroke cognitive impairment: a pooled analysis of individual patient data from 12 acute ischaemic stroke cohorts. *Circulation* 2019;139(9):1134-1145.
  17. Ouyang M, Boaden E, Arima H, et al. Dysphagia screening and risks of pneumonia and adverse outcomes after acute stroke: An international multicenter study. *Int J Stroke* 2020;15:206-215.

18. Montaña A, Hanley DF, Hemphill JC. Hemorrhagic stroke. *Interventional Neuroradiology* 2021; 229-248.
19. Ng Y, Qi W, King NKK, et al. Initial antihypertensive agent effects on acute blood pressure after intracerebral haemorrhage. *Stroke and Vascular Neurology* 2022;7:367-374.
20. Algra AM KC, Helmerhorst FM, Algra A, Rinkel GJ. <female risk factors for subarachnoid hemorrhag.pdf>. 2012;79(12):1230-1236.
21. Fuentes AM, Stone McGuire L, Amin-Hanjani S. Sex Differences in Cerebral Aneurysms and Subarachnoid Hemorrhage. *Stroke* 2022;53:624-633.
22. Chen B, Saban D, Rauscher S, et al. Modifiable Cardiovascular Risk Factors in Patients With Sporadic Cerebral Cavernous Malformations. *Stroke* 2021;52:1259-1264.
23. An SJ, Kim TJ, Yoon B-W. Epidemiology, Risk Factors, and Clinical Features of Intracerebral Hemorrhage: An Update. *Journal of Stroke* 2017;19:3-10.
24. Matsukawa H, Shinoda M, Fujii M, et al. Factors associated with lobar vs. non-lobar intracerebral hemorrhage. *Acta Neurol Scand* 2012;126:116-121.
25. Dye JA, Rees G, Yang I, Vespa PM, Martin NA, Vinters HV. Neuropathologic analysis of hematomas evacuated from patients with spontaneous intracerebral hemorrhage. *Neuropathology* 2014;34:253-260.
26. Kanazawa T, Takahashi S, Minami Y, Jinzaki M, Toda M, Yoshida K. Early prediction of clinical outcomes in patients with aneurysmal subarachnoid hemorrhage using computed tomography texture analysis. *J Clin Neurosci* 2020;71:144-149.
27. Macdonald RL, Schweizer TA. Spontaneous subarachnoid haemorrhage. *The Lancet* 2017;389:655-666.
28. Yao A, Jia L, Li J, et al. Fusion of subarachnoid hemorrhage data and computed tomography angiography data is helpful to identify the rupture source in patients with multiple intracranial aneurysms. *Neurosurg Rev* 2021;44:1411-1416.
29. 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;44:3018-3026.
30. Shaligram SS, Winkler E, Cooke D, Su H. Risk factors for hemorrhage of brain arteriovenous malformation. *CNS Neuroscience & Therapeutics* 2019;25:1085-1095.
31. Sheth KN, Ropper AH. Spontaneous Intracerebral Hemorrhage. *New England Journal of Medicine* 2022;387:1589-1596.
32. Boe NJ, Hald SM, Jensen MM, et al. Major Cardiovascular Events After Spontaneous Intracerebral Hemorrhage by Hematoma Location. *JAMA Network Open* 2023;6.
33. Yue Q, Liu Y, Leng B, et al. A Prognostic Model for Early Post-Treatment Outcome of Elderly Patients With Aneurysmal Subarachnoid Hemorrhage. *World Neurosurgery* 2016;95:253-261.
34. Hemphill JC 3rd BD, Besmertis L, Manley GT, Johnston SC. <the ich score a simple reliable grading.pdf>. 2001;32(4):891-897.
35. Rutten-Jacobs LCA, Maaijwee NAM, Arntz RM, et al. Clinical characteristics and outcome of intracerebral hemorrhage in young adults. *Journal of Neurology* 2014;261:2143-2149.

36. Tharmaseelan H, Rotkopf LT, Ayx I, et al. Evaluation of radiomics feature stability in abdominal monoenergetic photon counting CT reconstructions. *Scientific Reports* 2022;12.
37. Adezati E, Thye M, Edmondson-Stait AJ, Szaflarski JP, Mirman D. Lesion correlates of auditory sentence comprehension deficits in post-stroke aphasia. *Neuroimage: Reports* 2022;2.
38. Nadeau SE. Neural mechanisms of emotions, alexithymia, and depression. *Disorders of Emotion in Neurologic Disease* 2021: 299-313.
39. Meguro T, Kuwahara K, Tomita Y, et al. Primary Pontine Hemorrhage in the Acute Stage: Clinical Features and a Proposed New Simple Scoring System. *Journal of Stroke and Cerebrovascular Diseases* 2015;24:860-865.
40. Hess F, Foerch C, Keil F, Seiler A, Lapa S. Association of Lesion Pattern and Dysphagia in Acute Intracerebral Hemorrhage. *Stroke* 2021;52:2921-2929.
41. Galovic M, Leisi N, Müller M, et al. Lesion Location Predicts Transient and Extended Risk of Aspiration After Supratentorial Ischemic Stroke. *Stroke* 2013;44:2760-2767.
42. Rincon F, Dhmoon M, Moon Y, et al. Stroke Location and Association With Fatal Cardiac Outcomes. *Stroke* 2008;39:2425-2431.
43. Joyce T, Buoso S, Stoeck CT, Kozerke S. Rapid inference of personalised left-ventricular meshes by deformation-based differentiable mesh voxelization. *Medical Image Analysis* 2022;79.
44. Morotti A, Arba F, Boulouis G, Charidimou A. Noncontrast CT markers of intracerebral hemorrhage expansion and poor outcome. *Neurology* 2020;95:632-643.
45. Delcourt C, Zhang S, Arima H, et al. Significance of Hematoma Shape and Density in Intracerebral Hemorrhage. *Stroke* 2016;47:1227-1232.
46. Doumou G, Siddique M, Tsoumpas C, Goh V, Cook GJ. The precision of textural analysis in 18F-FDG-PET scans of oesophageal cancer. *European Radiology* 2015;25:2805-2812.
47. Serrano E, López-Rueda A, Moreno J, et al. The new Hematoma Maturity Score is highly associated with poor clinical outcome in spontaneous intracerebral hemorrhage. *European Radiology* 2021;32:290-299.
48. Liu X, Zhang Q, Li J, et al. Coordinatized lesion location analysis empowering ROI-based radiomics diagnosis on brain gliomas. *European Radiology* 2023;33:8776-8787.
49. Jiang B, Li N, Shi X, et al. Deep Learning Reconstruction Shows Better Lung Nodule Detection for Ultra-Low-Dose Chest CT. *Radiology* 2022;303:202-212.

## Figures and Table Legends

**Figure 1. The workflow of constructing clinical-, location-, radiomics- and fusion models for prediction diagnosis of sICH.** (A) Lesion-symptom mapping and location features. (B) Clinical, imaging and location feature extraction. (C) Construction and fusion of etiologic model and prognostic model. (D) Model validation and evaluation.

ROC= Receiver operating characteristic.

**Figure 2. Association between hematoma location and etiology (n = 1109).** (A) Probability maps depicting the distribution of hematoma location in patients. The location patterns associated with hypertension, aneurysm, and vascular malformation are illustrated, with the frequency of hemorrhage location graded according to color. The gradient scale ranged from 5% to 40%. (B) Paired sample ANOVA was used to compare the involvement rate of each region in each intracranial structure. \*:  $p < 0.05$ , \*\*:  $p < 0.01$ , \*\*\*:  $p < 0.001$ , ns:  $p > 0.05$ .

**Figure 3. Association between hematoma location and prognosis (n = 1162).** (A) Probability maps depicting the distribution of hematoma location in different prognosis. The location patterns associated with subarachnoid space, parenchyma and ventricle are illustrated, with the frequency of hematoma location graded according to color. The gradient scale ranged from 5% to 40%. (B) Frequency plots of the top five cisterns, ten parenchyma and three ventricles showing a group difference are ranked according to chi-square values. \*:  $p < 0.05$ , \*\*:  $p < 0.01$ , \*\*\*:  $p < 0.001$ , ns:  $p > 0.05$ . Abbreviations of intracranial region see Supplementary Table 1.

**Figure 4. Etiology classification.** (A) The confusion matrix demonstrated the classification capability of the location model for predicting etiology in all datasets. Columns represent predicted classes, with rows representing ground truth. (B) The top ten location features associated with hypertension, aneurysm, and vascular malformation prediction models are shown. The horizontal axis represents the relative feature weights, with red coding the three features making the strongest contributions to the model. Positive values indicate a positive contribution, negative values a negative contribution. Abbreviations of intracranial region see Supplementary Table 1.

**Figure 5. Prognosis prediction.** (A-C) Receiver Operating Characteristic (ROC) curves of sICH predicted by the clinical, location, radiomics and multidimensional fusion models of dataset 1, 2 and 3. (D) displays the top ten location features of in the location model of prognosis prediction task. The horizontal axis represents the model weight value, red codes the three features making the strongest contribution. Abbreviations of intracranial region see Supplementary Table 1.

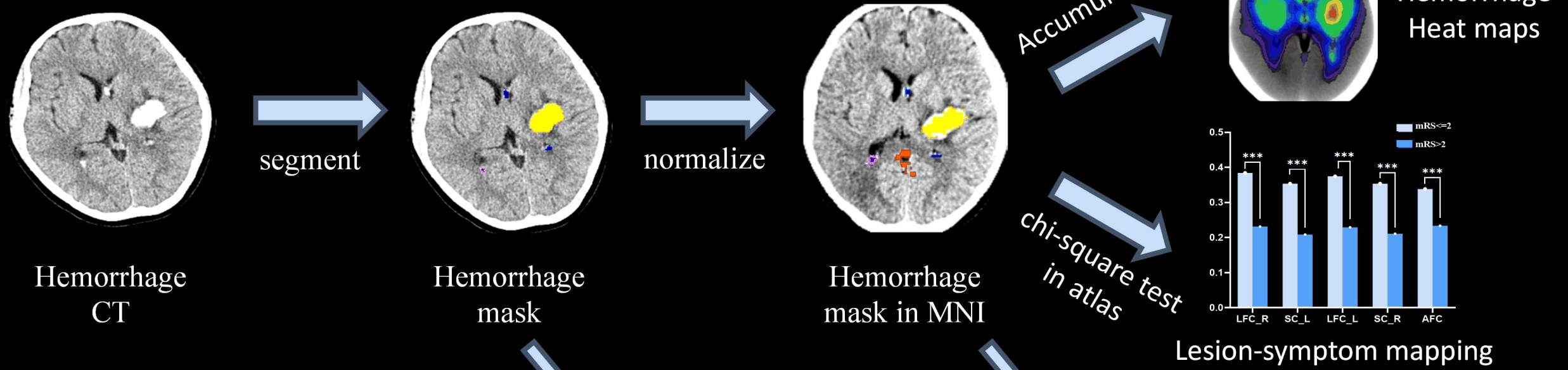


**Table 1. Demographic and clinical characteristics in sICH patients.**

<b>Variables</b>	<b>Subjects (n = 1162)</b>
Age, years, mean $\pm$	57.65 $\pm$ 13.19
GCS, median (IQR)	12 (10-14)
Sex, n (%)	
Male	655 (56.4%)
Female	507 (43.6%)
Hypertension, n (%)	791 (68.1%)
Diabetes mellitus, n	110 (9.5%)
Hyperlipidemia, n (%)	27 (2.3%)
Smoking, n (%)	220 (18.9%)
Drinking, n (%)	219 (18.9%)
Onset-to-CT time, n	
$\leq$ 24h	583 (50.2%)
24h-3d	428 (36.8%)
3-14d	128 (11.0%)
$>$ 14d	23 (2.0%)
History of stroke, n	93 (8.0%)
Etiology, n (%)	
Hypertension	563 (48.5%)
Aneurysm	435 (37.4%)
Vascular malformation	111 (10.0%)
Unknown	53 (4.6%)
Prognosis, n (%)	
mRS $\leq$ 2	721 (62.1%)
mRS $>$ 2	441 (37.9%)

GCS= Glasgow Coma Scale; mRS= the modified Rankin Scale

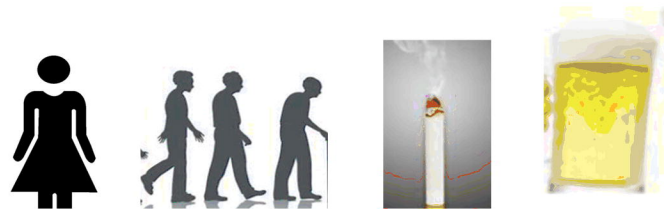
## A. Lesion-symptom mapping



## B. Feature extraction

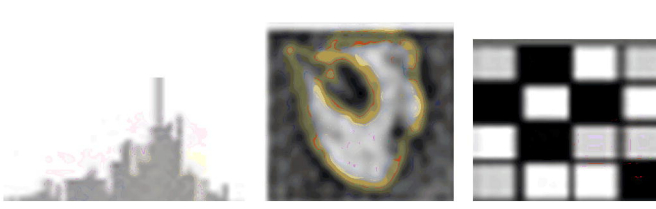
medRxiv preprint doi: <https://doi.org/10.1101/2024.05.22.24307743>; this version posted May 23, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

### Clinical features



Sex, Age, Smoking, Drinking...

### Radiomics features

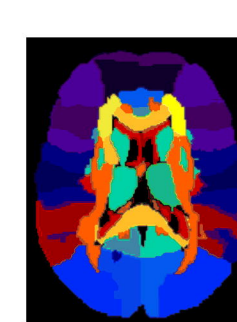


First-order features

Shape features

Texture features

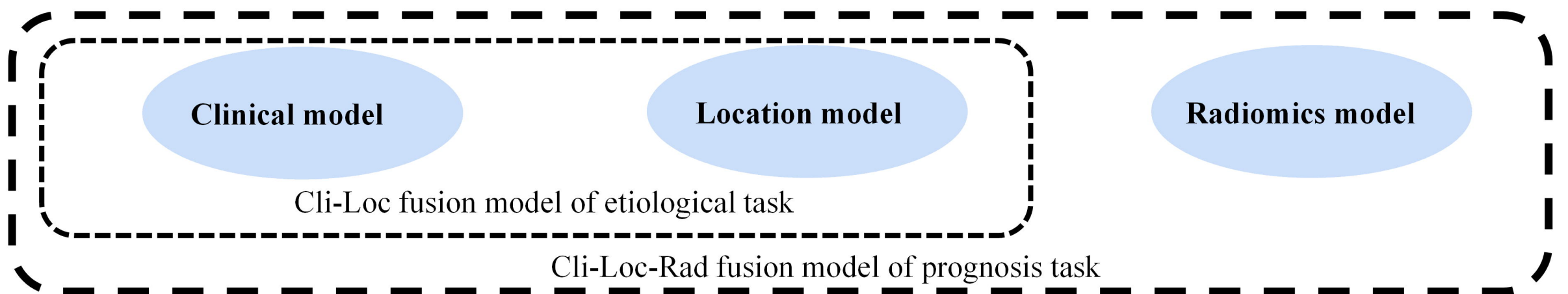
### Location features



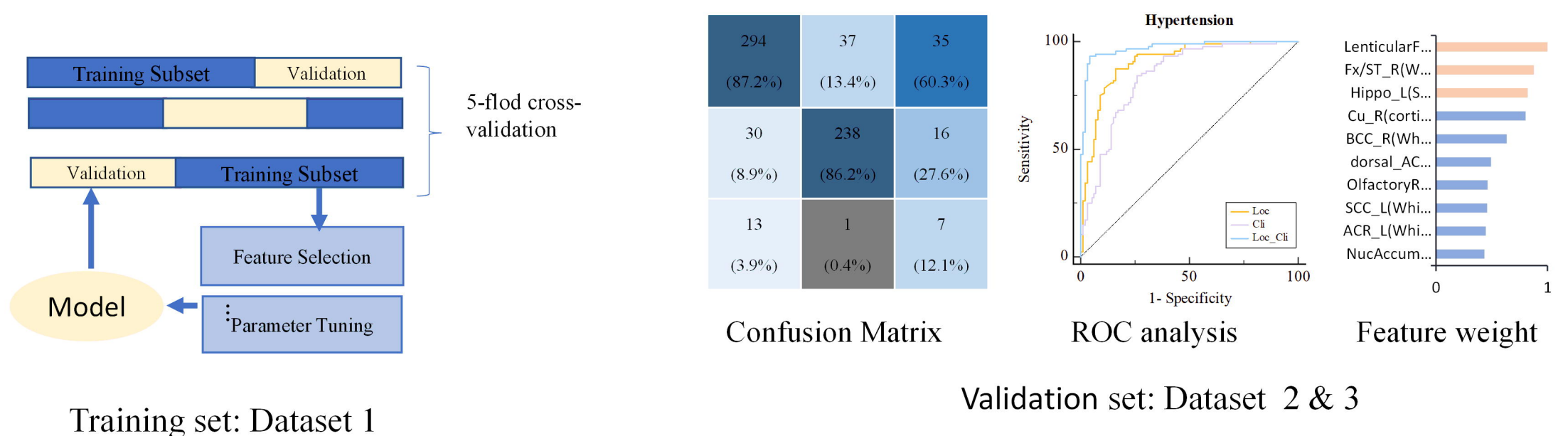
Hemorrhagic involvement of brain regions

- Subarachnoid atlas
- Parenchyma (JHU)
- Ventricle atlas

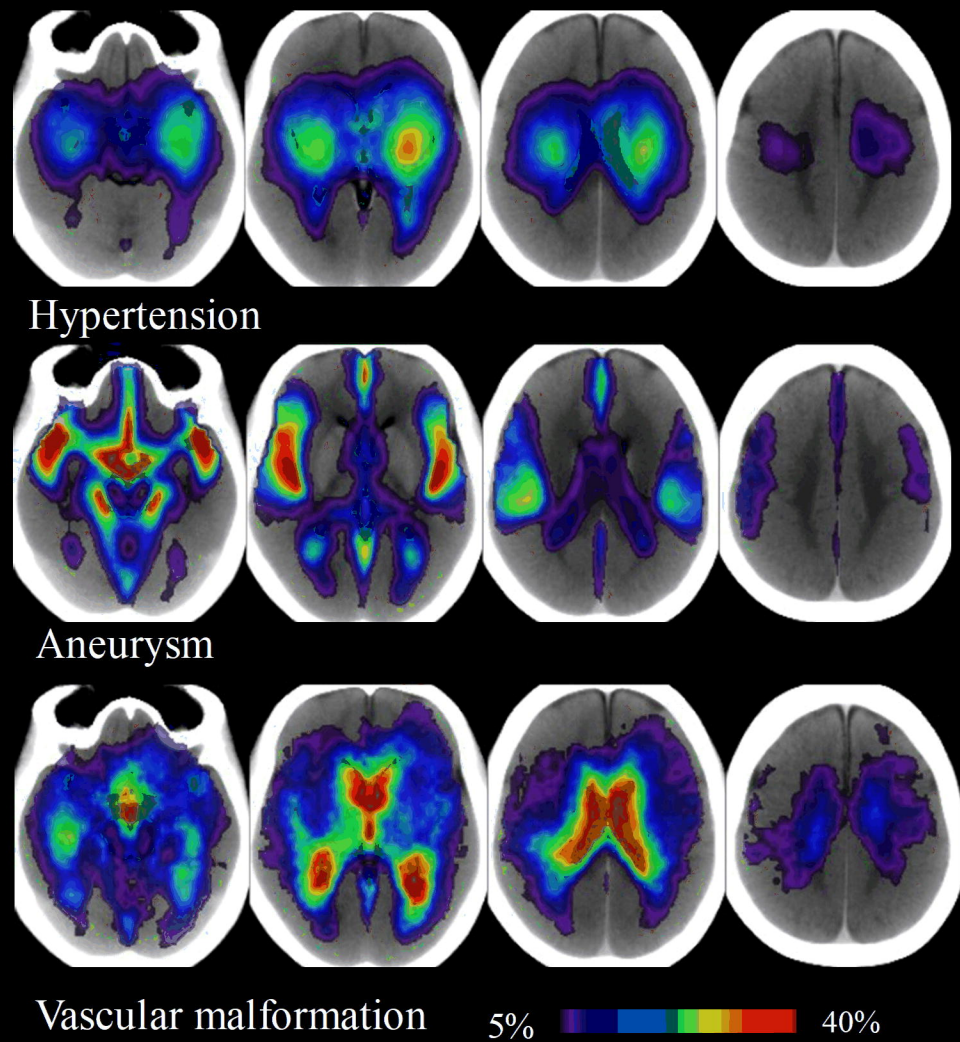
## C. Model construction and fusion



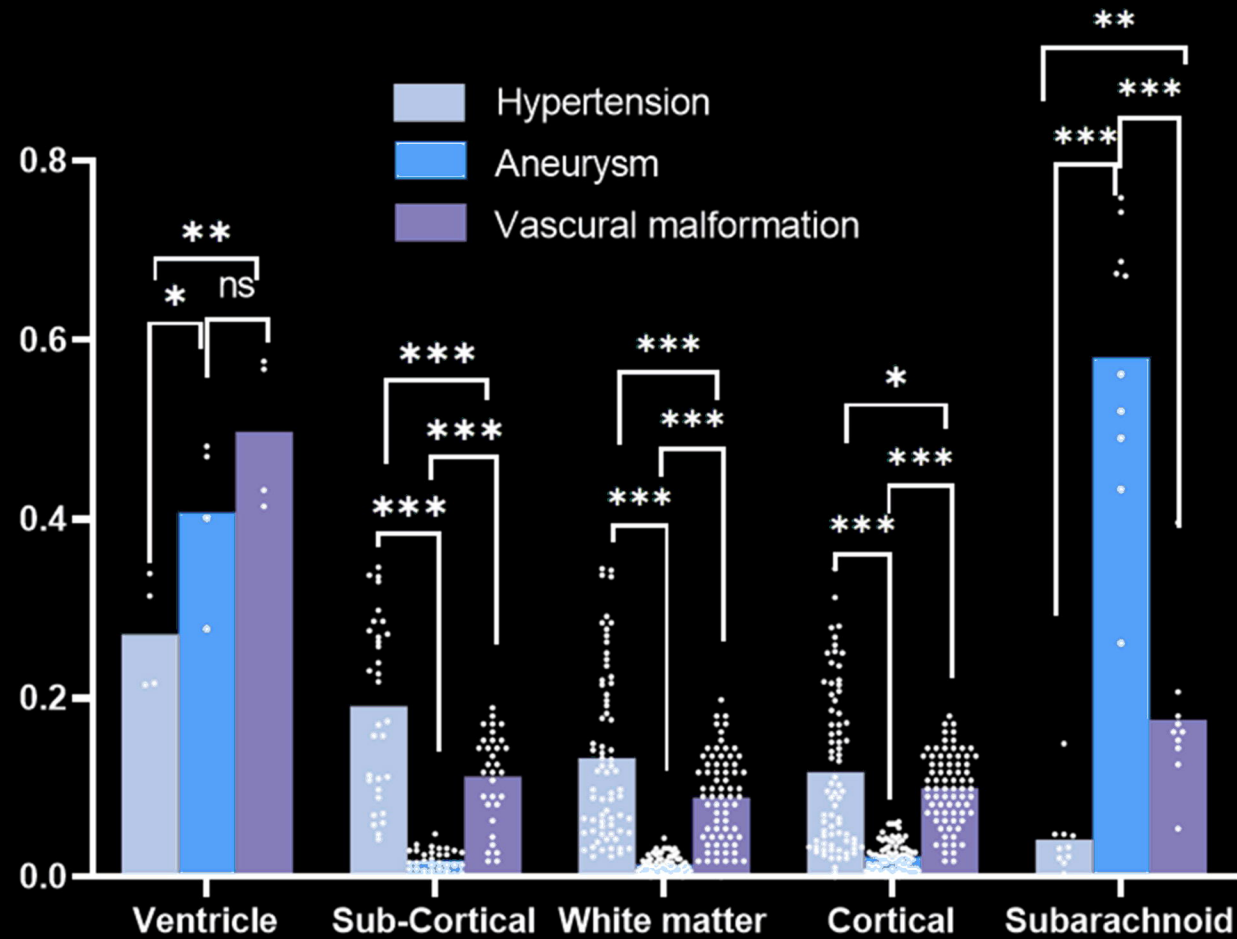
## D. Model validation and evaluation



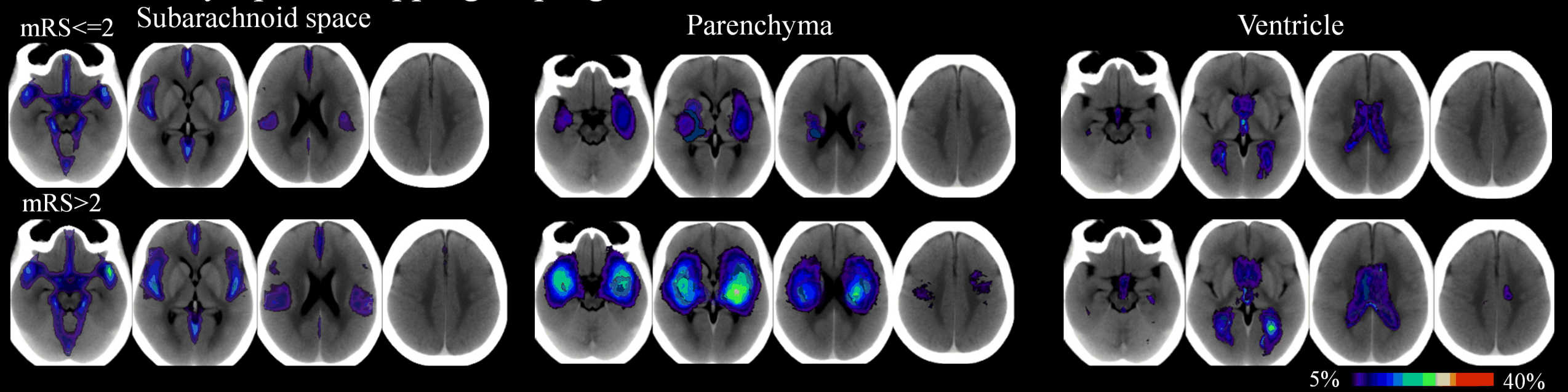
## A. Lesion-symptom mapping of etiologies



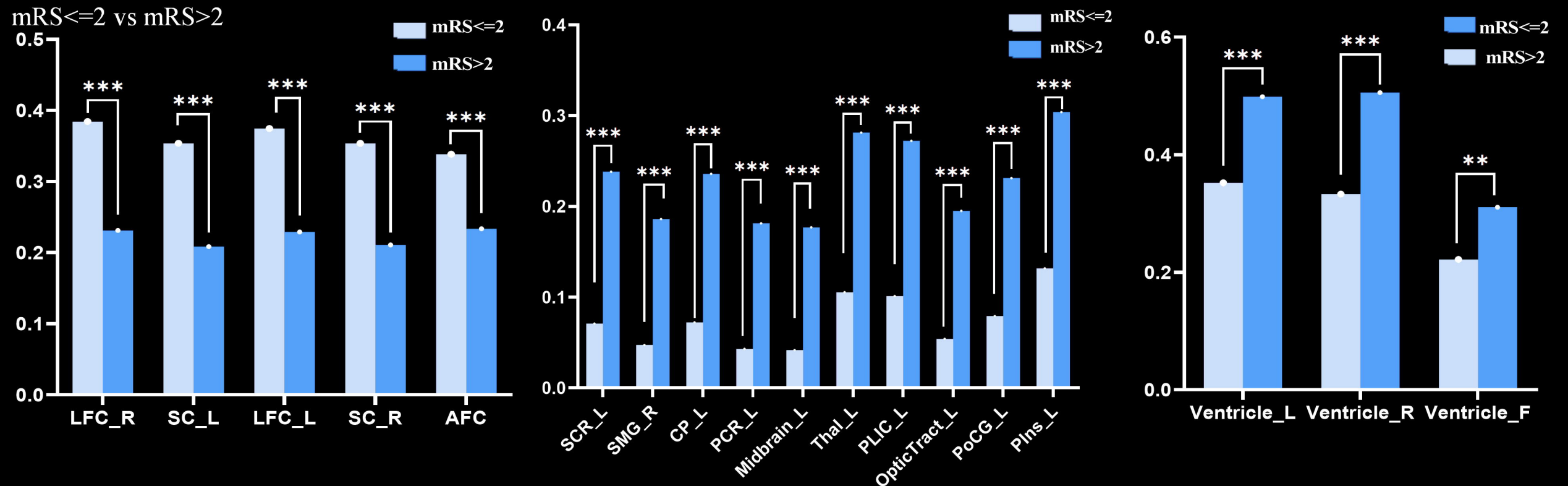
## B. Differences in location by etiology



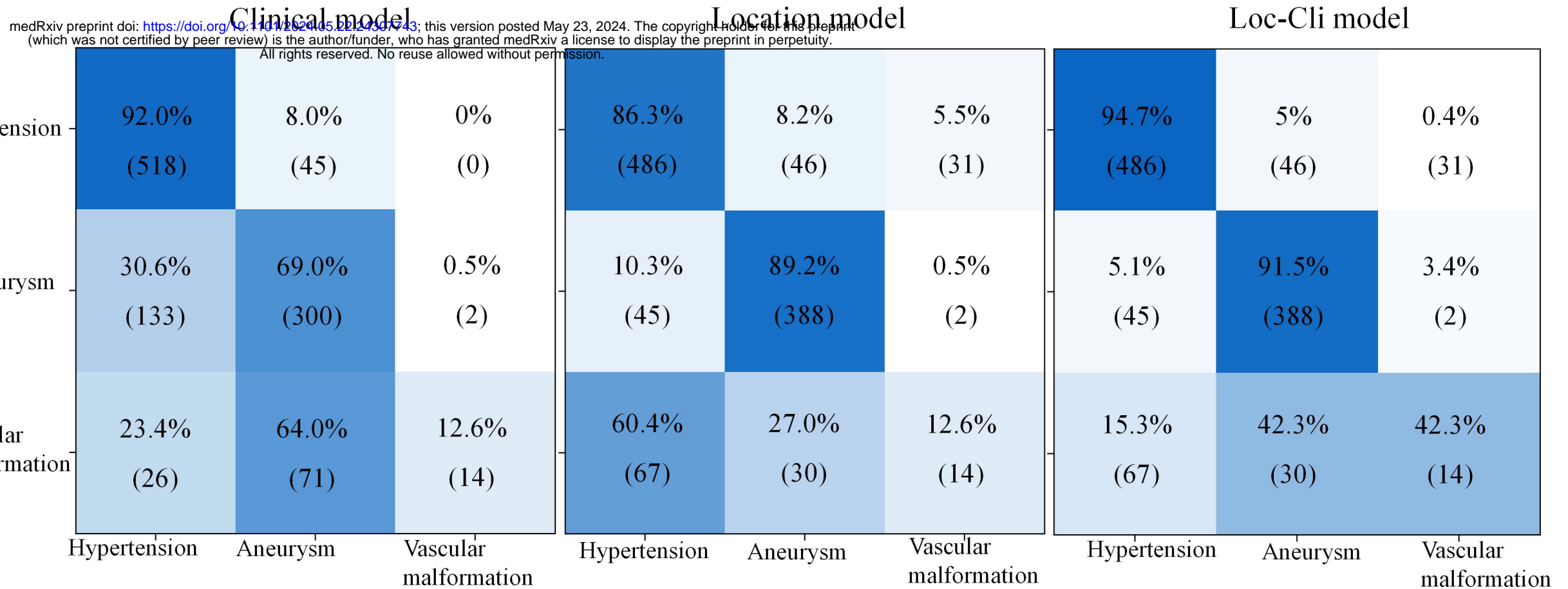
## A. Lesion-symptom mapping of prognosis



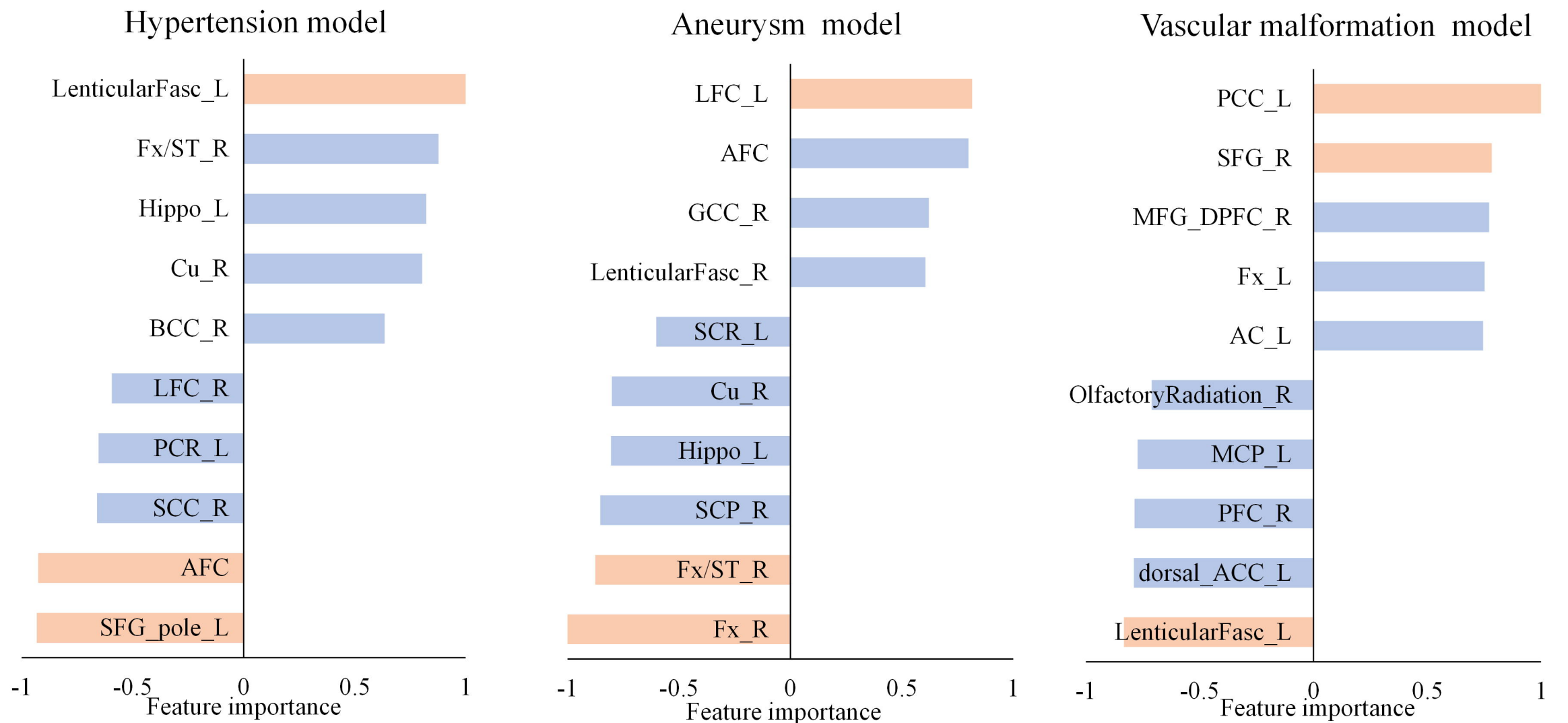
## B. Differences in location by prognosis



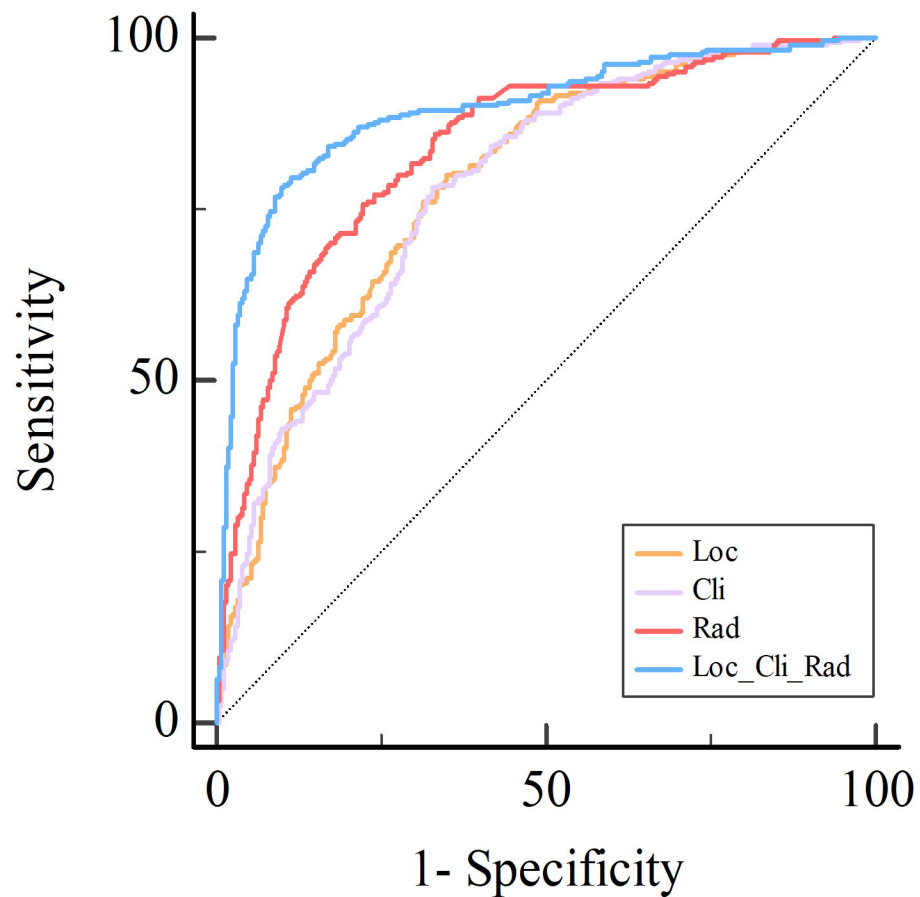
# A. Confusion matrix of etiology models



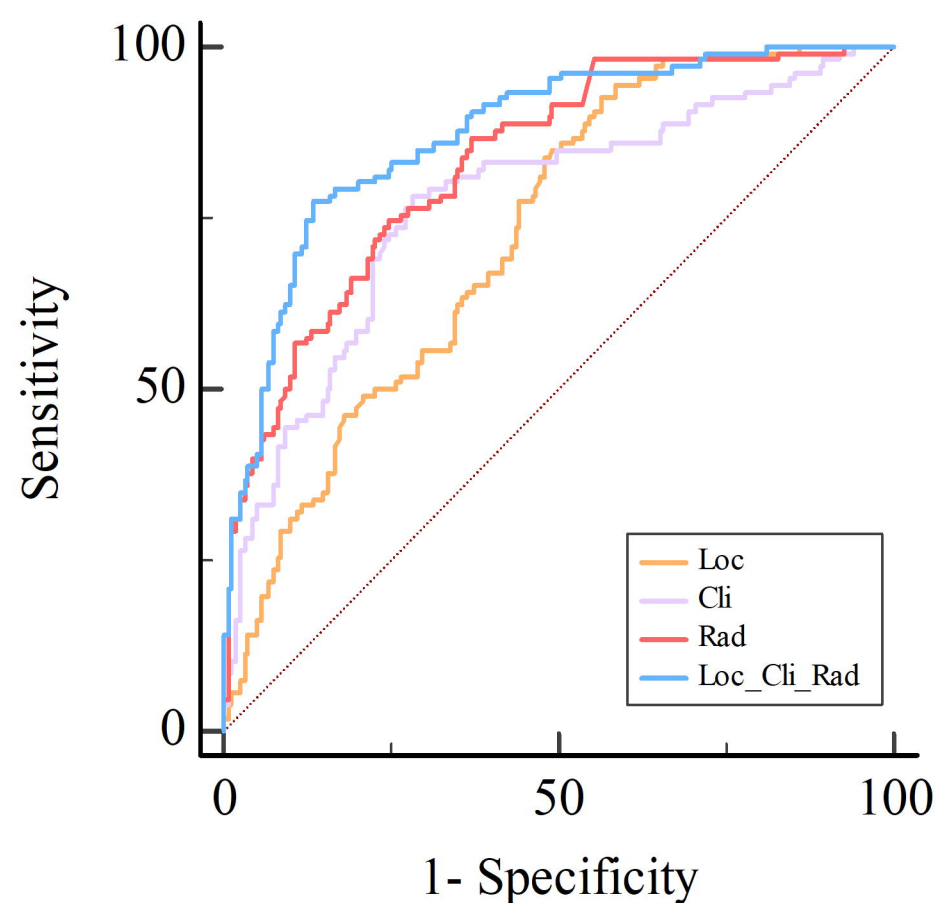
# B. Feature importance of etiology location models



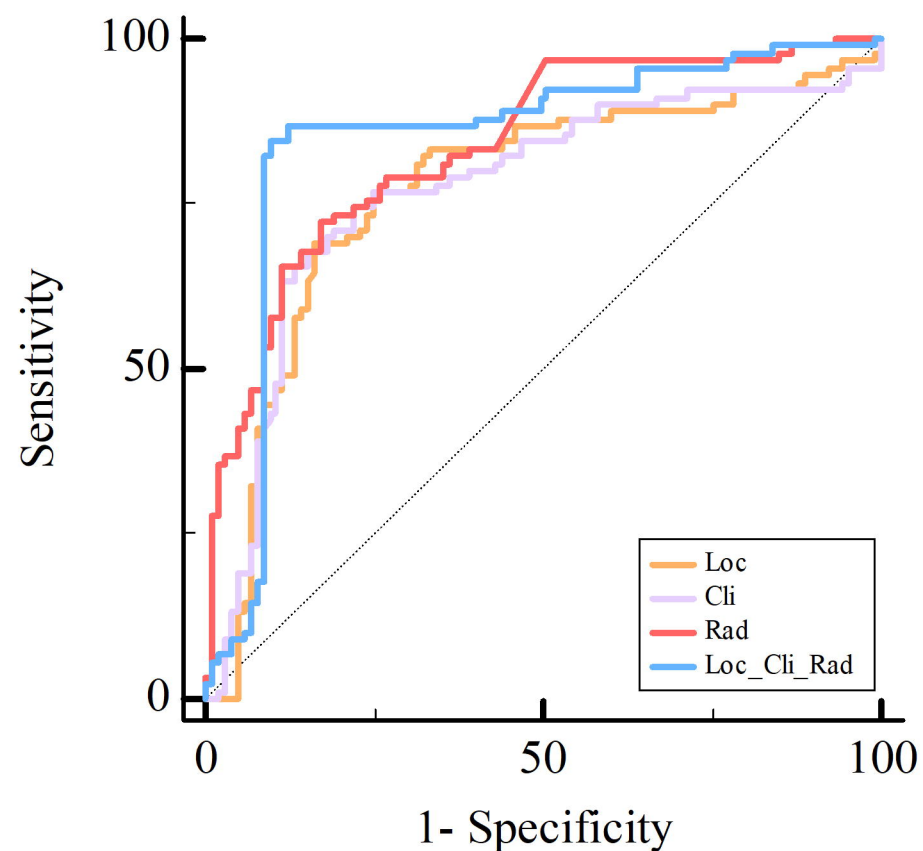
A. ROC of prognosis model(Dataset 1)



B. ROC of prognosis model(Dataset 2)



C. ROC of prognosis model(Dataset 3)



D. Feature importance of prognosis location model

