Combining clinical-radiomics features with machine learning methods for building models to predict postoperative recurrence in patients with chronic subdural hematoma: Retrospective Cohort Study

Cheng Fang, Yifeng Pan, Xiao Ji, Sai Li, Guanchao Xie, Hongsheng Zhang, Jinghai Wan

Submitted to: Journal of Medical Internet Research
on: November 28, 2023

Table of Contents

Combining clinical-radiomics features with machine learning methods for building models to predict postoperative recurrence in patients with chronic subdural hematoma: Retrospective Cohort Study

Cheng Fang[1*] MD; Yifeng Pan[2*] MSE; Xiao Ji[3*] MM; Sai Li[4] MM; Guanchao Xie[4] MM; Hongsheng Zhang[4] MM; Jinghai Wan[5] MD

[1]Department of Neurosurgery National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital Chinese Academy of Medical Sciences and Peking Union Medical College Beijing CN
[2]The School of Big Data and Artificial Intelligence Anhui Xinhua University Hefei CN
[3]Department of Neurosurgery Anhui Provincial Institute of Translational Medicine The First Affiliated Hospital of Anhui Medical University Hefei CN
[4]Department of Neurosurgery The Second Affiliated Hospital of Anhui Medical University Anhui Medical University Hefei CN
[5]Department of Neurosurgery National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital Beijing CN
[*]these authors contributed equally

Corresponding Author:
Jinghai Wan MD
Department of Neurosurgery
National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital
Chinese Academy of Medical Sciences and Peking Union Medical College
Beijing
CN

Abstract

Background: Chronic subdural hematoma (CSDH) represents a prevalent medical condition, posing significant challenges in postoperative management due to risks of recurrence. Currently, prognosis determination largely depends on clinician expertise, revealing a dearth of precise prediction models in clinical settings.

Objective: This study sought to employ machine learning (ML) techniques for the construction of predictive models to assess the likelihood of CSDH recurrence post-surgery.

Methods: Data from 133 patients were amassed and partitioned into a training set (n=114) and a test set (n=19). Radiomics features were extracted from preoperative cranial computed tomography (CT) scans utilizing 3D Slicer software. These features, in conjunction with clinical data and composite clinical-radiomics features, served as input variables for model development. Four distinct ML algorithms were utilized to build predictive models, and their performance was rigorously evaluated via accuracy (ACC), area under the curve (AUC), and Recall metrics. The optimal model was identified, followed by recursive feature elimination (RFE) for feature selection, leading to enhanced predictive efficacy. External validation was conducted using datasets from additional healthcare facilities.

Results: Following rigorous experimental analysis, the Support Vector Machine (SVM) model, predicated on clinical-radiomics features, emerged as the most efficacious for predicting postoperative recurrence in CSDH patients. Subsequent to feature selection, key variables exerting significant impact on the model were incorporated as the input set, thereby augmenting its predictive accuracy. The model demonstrated robust performance, with metrics including accuracy (ACC) of 92.72%, area under the curve (AUC) of 91.34%, and Recall of 93.16%. External validation further substantiated its effectiveness, yielding an ACC of 90.32%, AUC of 91.32%, and Recall of 88.37%, affirming its clinical applicability.

Conclusions: Conclusion: The present study substantiates the feasibility and clinical relevance of a machine learning-based predictive model, utilizing clinical-radiomics features, for precise prognostication of postoperative recurrence in CSDH patients. This model holds considerable import for enhancing the quality and efficiency of clinical decision-making processes.

Conclusion: The present study substantiates the feasibility and clinical relevance of a machine learning-based predictive model, utilizing clinical-radiomics features, for precise prognostication of postoperative recurrence in CSDH patients. This model holds

considerable import for enhancing the quality and efficiency of clinical decision-making processes.

(JMIR Preprints 28/11/2023:54944)
DOI: https://doi.org/10.2196/preprints.54944

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Combining clinical-radiomics features with machine learning methods for building models to predict postoperative recurrence in patients with chronic subdural hematoma: Retrospective Cohort Study

Cheng Fang[1†], Yifeng Pan[2†], Xiao Ji[3,4†], Guanchao Xie[5], Hongsheng Zhang[5], Sai Li[5] and Jinghai Wan[1,5*].

[1]*Department of Neurosurgery, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China*

[2]*The School of Big Data and Artificial Intelligence, Anhui Xinhua University, Hefei, China*

[3]*Department of Neurosurgery, The First Affiliated Hospital of Anhui Medical University, Hefei, China*

[4] *Anhui Provincial Institute of Translational Medicine, Hefei, China*

[5]*Department of Neurosurgery, The Second Affiliated Hospital of Anhui Medical University, Anhui Medical University, Hefei, China*

[†]These authors have contributed equally to this work and share first authorship

*Corresponding Author:
Jinghai Wan, Department of Neurosurgery, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College.
Beijing, China
Phone: 86 13426261848
Email: wanjinghai@sina.com

Abstract

Background: Chronic subdural hematoma (CSDH) represents a prevalent medical condition, posing significant challenges in postoperative management due to risks of recurrence. Such recurrences not only cause physical suffering to the patient, but also add to the financial burden on the family and the healthcare system. Currently, prognosis determination largely depends on clinician expertise, revealing a dearth of precise prediction models in clinical settings. In the above clinical context, this study sought to employ machine learning (ML) techniques for the construction of predictive models to assess the likelihood of CSDH recurrence post-surgery, which leads to greater benefits for patients and the healthcare system.

Methods: Data from 133 patients were amassed and partitioned into a training set (n=93) and a test set (n=40). Radiomics features were extracted from preoperative cranial computed tomography (CT) scans utilizing 3D Slicer software. These features, in conjunction with clinical data and composite clinical-radiomics features, served as input variables for model development. Four distinct ML algorithms were utilized to build predictive models, and their performance was rigorously evaluated via accuracy (ACC), area under the curve (AUC), and Recall metrics. The optimal model was identified, followed by recursive feature elimination (RFE) for feature selection, leading to enhanced predictive efficacy. External validation was conducted using datasets from additional healthcare facilities.

Results: Following rigorous experimental analysis, the Support Vector Machine (SVM) model, predicated on clinical-radiomics features, emerged as the most efficacious for predicting postoperative recurrence in CSDH patients. Subsequent to feature selection, key variables exerting significant impact on the model were incorporated as the input set, thereby augmenting its predictive accuracy. The model demonstrated robust performance, with metrics including accuracy (ACC) of 92.72%, area under the curve (AUC) of 91.34%, and Recall of 93.16%. External validation further substantiated its effectiveness, yielding an ACC of 90.32%, AUC of 91.32%, and Recall of 88.37%, affirming its clinical applicability.

Conclusion: The present study substantiates the feasibility and clinical relevance of a machine learning-based predictive model, utilizing clinical-radiomics features, for relatively accurate prognostication of postoperative recurrence in CSDH patients. If the model is integrated into clinical practice, it will be of great significance in enhancing the quality and efficiency of clinical decision-making processes, which can improve the accuracy of diagnosis and treatment, reduce unnecessary tests and surgeries, and reduce the waste of medical resources.

KEYWORDS

chronic subdural hematoma; convolutional neural network; machine learning; neurosurgery; radiomics; support vector machine

Background

Chronic subdural hematoma (CSDH) is a prevalent neurosurgical pathology, disproportionately affecting middle-aged and elderly populations. Epidemiological data indicate incidence rates of 13.5/100,000, escalating to 58.1/100,000 in individuals 65 years or older [1, 2]. Manifestations commonly include headache, nausea, vomiting, and diplopia, indicative of elevated intracranial pressure. Diagnosis is generally confirmed through cranial computed tomography (CT) or magnetic resonance imaging (MRI). Established as a medical condition since 1857, surgical intervention remains a proven, efficacious treatment modality for CSDH. However, postoperative recurrence

serves as a critical metric for evaluating surgical success [3]. Such recurrence imposes not only physical suffering on patients but also accentuates financial burden on families and healthcare systems. For patients who are older, have a history of multiple surgeries, or have other complications, an accurate predictive tool can help physicians identify these high-risk patients in advance, allowing for a more precise treatment and follow-up plan. Each patient with CSDH has a different condition and clinical background, and thus requires a personalized treatment plan that can assess the risk of postoperative recurrence based on the patient's specific clinical information (age, gender, medical history, symptoms, signs, and imaging manifestations, etc.). And with the increasing strain on medical resources, how to optimize the allocation and utilization of medical resources and improve the operational efficiency of hospitals has become an important issue. Hence, in the above clinical context, Hence, the development of a predictive tool for postoperative recurrence risk is integral for informed clinical decision-making and optimized treatment outcomes, which can bring greater benefits to both patients and the healthcare system.

Recent advancements in computer technology have facilitated the construction of predictive models anchored on clinically pertinent data. Machine learning (ML) has emerged as a particularly robust paradigm, capable of delineating complex, nonlinear relationships between variables and outcomes. A plethora of studies substantiate the impressive levels of accuracy and reliability achieved through ML applications [4-7]. This current study employs four ML methodologies—convolutional neural networks (CNN), support vector machines (SVM), random forests (RF), and linear regression (LR)—each enjoying widespread academic acceptance and demonstrated applicability in predictive research across various domains, including healthcare and food sciences [6, 8, 9].

By surveying existing literature, it is evident that machine learning (ML) models integrated with radiomics are garnering increased scholarly attention [10, 11]. Radiomics constitutes a novel approach in medical image analysis, principally centered on quantitative feature extraction. This technique transforms medical imagery into high-dimensional structures, facilitating the comprehensive analysis of regions of interest in conjunction with relevant clinical, diagnostic, and prognostic data. A typical radiomics workflow encompasses stages of image acquisition, reconstruction, preprocessing and processing, feature extraction, selection, and eventually, classification or regression modeling [12]. Although previous research has melded radiomics and ML for diagnostic and prognostic applications in other medical specialties such as dermatology, oncology, and cardiology [13-15], studies targeting CSDH remain comparatively scant.

The objective of this investigation is to amalgamate ML algorithms with radiomics and clinical variables for the construction of a predictive model aimed at gauging the risk of CSDH recurrence post-surgery. The study will rigorously compare various methodologies and models to identify the most efficacious predictive framework for CSDH recurrence.

Materials and methods

Ethical review

The retrospective study was approved (approval: PJ-YX2024-021) by the Ethics Committee of the First Affiliated Hospital of Anhui Medical University (Anhui Public Health Clinical Center).

Participants

We compiled clinical and radiological data from patients diagnosed with CSDH who were treated at the neurosurgery department of the Second Affiliated Hospital of Anhui Medical University between May 2012 and May 2022. The inclusion criteria were as follows: (1) a confirmed clinical diagnosis of CSDH; (2) subjects must have undergone surgical intervention, either Burr hole craniostomy or craniotomy; and (3) comprehensive clinical records must be available, encompassing treatment histories, preoperative and postoperative imaging examinations, laboratory analyses, among other pertinent data. Exclusion criteria included: (1) patients who exhibited symptomatic improvement via pharmacological intervention, obviating the need for surgical treatment; (2) any prior neurosurgical procedures that could potentially induce CSDH; and (3) cases where follow-up data was incomplete or where recurrence post-surgery was undetermined. Following these criteria, 133 patients were incorporated into the study, there are no missing values for all case data in this study. The process of patient selection and enrollment is delineated in Figure 1. Furthermore, an external validation set was generated by screening data from 20 CSDH patients who underwent treatment at the First Affiliated Hospital of Anhui Medical University.

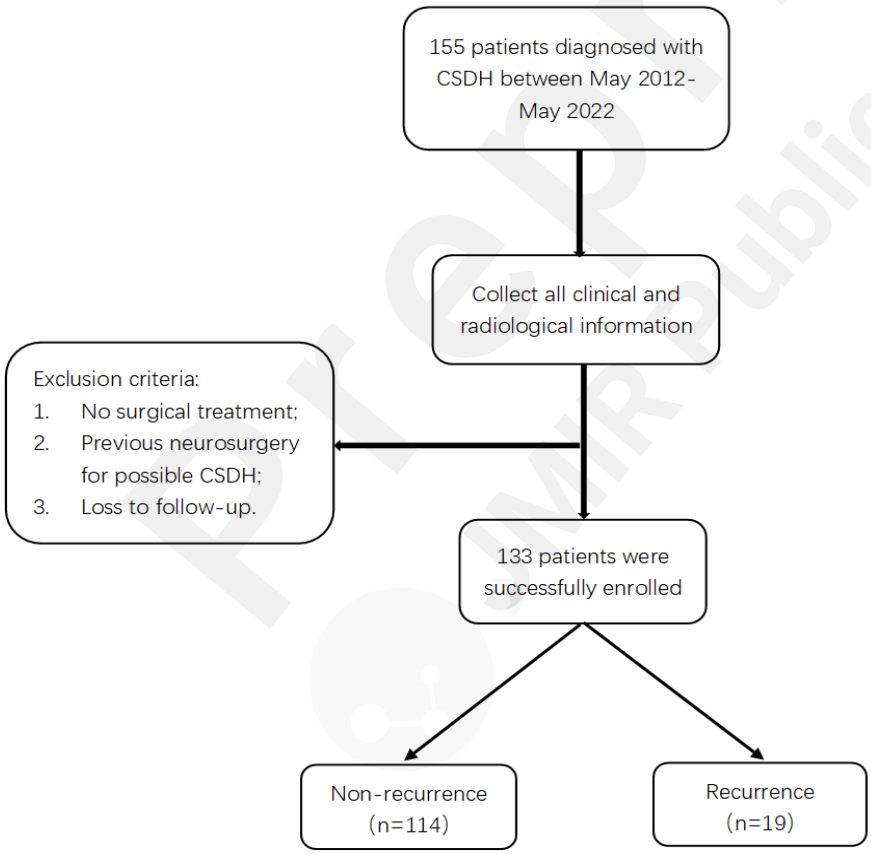


Figure1. Patient selection and enrollment process

Clinical data

Upon a rigorous review of existing literature and consultations with experienced neurosurgeons, specific data parameters were established for model construction. These parameters were extracted

from the electronic case files of the participants and included: patient demographics (age, gender), pertinent clinical history (smoking or drinking habits, prior medical history, history of head trauma, history of antiplatelet or anticoagulant therapy), preoperative clinical grading based on the Markwalder Grading Scale, and duration of hospitalization. Imaging data encompassed variables such as the location of the CSDH (unilateral or bilateral), hematoma classification, preoperative hematoma volume, preoperative midline shift, postoperative midline improvement, and cranial CT scans. Subsequent to the follow-up period, 19 patients exhibited postoperative recurrence. These patients were bifurcated into two cohorts: those with recurrence and those without. Clinical indices employed in model construction are elaborated in Table 1. Continuous variables were subjected to t-test analyses and are represented as mean ± standard deviation (SD). Categorical variables underwent $\chi 2$ testing and are conveyed as percentages. Statistical analysis was performed using IBM SPSS Statistics 27.0.

Table 1. Clinical variables used to construct the model

	Non-recurrence (n=114)	Recurrence (n=19)	*P*-value
Age (years)	70.1±9.4	73.7±8.3	0.126
Gender (%)			0.497
Male	91(79.8)	17(89.5)	
Female	23(20.2)	2(10.5)	
Smoking or drinking (%)			0.119
Yes	56(49.1)	13(68.4)	
No	58(50.9)	6(31.6)	
Hypertension (%)			0.097
Yes	43(37.7)	11(57.9)	
No	71(62.3)	8(42.1)	
Cerebral infarction (%)			0.023
Yes	11(9.6)	6(31.6)	
No	103(90.4)	13(68.4)	
History of head trauma (%)			0.342
Yes	73(64.0)	10(52.6)	
No	41(36.0)	9(47.4)	
History of antiplatelet or anticoagulation (%)			0.003
Yes	10(8.8)	7(36.8)	
No	104(91.2)	12(63.2)	
Markwalder Grading Scale (%)			0.004
0	1(0.9)	0(0)	
1	38(33.3)	0(0)	
2	56(49.1)	12(63.2)	
3	18(15.8)	7(36.8)	
4	1(0.9)	0(0)	
Length of stay in hospital (days)	14.0±4.3	14.5±3.2	0.639
CSDH location (%)			0.014
Unilateral	104(91.2)	13(68.4)	

Bilateral	10(8.8)	6(31.6)	
Classification of hematoma（%）			0.022
Homogenous	50(43.9)	2(10.5)	
Laminar	12(10.5)	1(5.3)	
Separated	35(30.7)	8(42.1)	
Trabecular type	17(14.9)	8(42.1)	
Preoperative hematoma volume（ml）	103.4±25.0	127.1±49.3	0.054
Preoperative midline shift（cm）	1.1±0.3	0.9±0.3	0.010
Postoperative midline improvement（cm）	0.6±0.2	0.4±0.2	0.011

In Table 2, we delineate the grading criteria and definitions associated with the Markwalder Grading Scale (MGS) [16]. Initially proposed in the 1980s, extensive research has validated the MGS as a robust metric for evaluating postoperative neurological recovery and prognosis in CSDH patients. Specifically, a grade of 0 indicates normal neurological function, grades 1-2 signify good neurological function, and grades 3-4 represent poor neurological function.

Table 2．Descriptions of the categories of the Markwalder Grading Scale [16]

Grading score	
Grade 0	Neurologically normal
Grade 1	Alert and orientated: absence of mild symptoms such as headache, or mild neurological deficit such as reflex asymmetry
Grade 2	Drowsy or disorientated, or variable neurological deficit such as hemiparesis
Grade 3	Stuporous, but responding appropriately to noxious stimuli, several focal signs such as hemiplegia
Grade 4	Comatose with absent motor responses to painful stimuli, decerebrate or decorticate posturing

Image data

Complementing the comprehensive clinical dataset, preoperative cranial CT scans were acquired for all enrolled participants. Hematoma images were systematically categorized into four distinct types: Homogenous, Laminar, Separated, and Trabecular. Representative CT scans for each category are furnished in Figure 2. This taxonomic approach to hematoma classification was initially delineated by Nakaguchi et al. [17]. Their work posited that hematomas with more irregular structures correlated with elevated recurrence rates. Subsequently, this classification schema has been integrated into a scoring system aimed at assessing recurrence risk [18]. All procured CT images were stored in DICOM format.

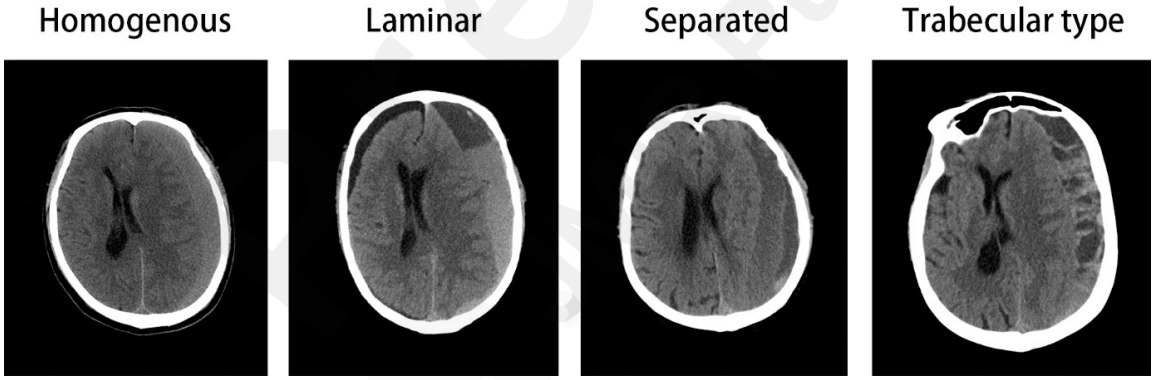


Figure2. Classification of hematoma

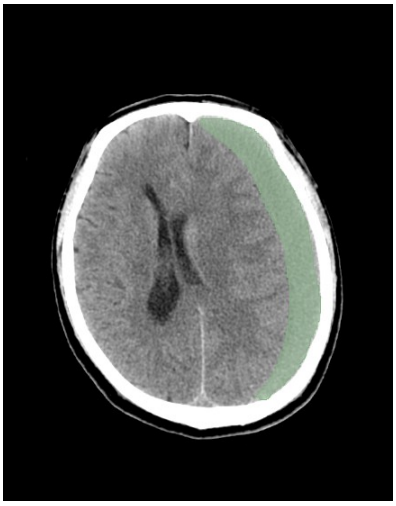


Figure 3. Hand-delineated regions of interest

Radiomics

Hematoma segmentation was executed through semi-automatic techniques utilizing the 3D Slicer open-source software platform (version 4.10, www.slicer.org). The region of interest (ROI), specifically the hematoma, was further segmented using the PyRadiomics package, an open-source plugin available on the 3D Slicer platform (Figure 3). The utilization of open-source software enabled the direct computation of three-dimensional (3D) features without the necessity for slice-wise combination or averaging. A total of 107 radiomic features were manually extracted from each patient's CT images. These features were allocated to seven distinct feature categories: 18 were first-order statistics, 14 were shape-based, 24 were derived from gray-level co-occurrence matrices, 16 from gray-level run-length matrices, 16 from gray-level size-zone matrices, five from adjacent gray-level difference matrices, and 14 from gray-level dependence matrices. This conversion from image-based to data-driven features optimizes the dataset for subsequent computational analyses and research endeavors.

Modeling

Neuroscientists participated comprehensively in this research, inclusive of data collection and categorical predictive model development. To safeguard patient confidentiality and adhere to the ethical guidelines stipulated by the Ethics Committee, patient data were anonymized through numerical coding. This approach also ensured the clinical applicability of the study's findings.

Four machine learning algorithms were employed for model development: Convolutional Neural Networks (CNN), Support Vector Machines (SVM), Random Forests (RF), and Logistic Regression (LR). Hyperparameter tuning was conducted via a grid search algorithm to optimize model performance. The main components of the CNN structure include an input layer, three convolution-pooling layers, a flat layer, two fully connected layers, and an output layer. Tables 3 and 4 detail the parameter configurations of the CNN and SVM. The dataset was stratified into a training set,

comprising 70% of the samples (N=93), and a test set accounting for the remaining 30% (N=40). Given the dataset's limited sample size, 5-fold cross-validation was executed on the training set to ensure robustness and validity. This cross-validation technique is a standard practice in machine learning for its ability to produce reliable performance metrics, mitigate the risks of overfitting and underfitting, and assess the model's generalization capability. The models were developed using the scikit-learn framework and implemented in the Python 3.9 programming environment.

Input Dataset Selection: Multiple input dataset configurations were assessed to optimize the predictive model. Initial models were constructed using three distinct types of data: clinical data, CT images, and radiomics from the enrolled patient cohort. These data types were then aggregated in two specific combinations to generate models based on mixed input data, specifically, clinical-CT and clinical-radiomics features. Furthermore, a composite model utilizing clinical data, CT images, and histological imaging was also developed. Comparative analysis of these dataset configurations was performed to identify the algorithms and input datasets most conducive for predictive modeling.

Feature selection: Our evaluation indicated that the Support Vector Machine (SVM) model, when configured with the clinical-radiomics dataset, demonstrated superior predictive efficacy. However, the complexity of the input variables compromised both the model's performance and computational efficiency, thereby not meeting the study's predefined objectives. To address this, a feature selection strategy was employed with the principal aim of isolating the most impactful variables. This was anticipated to enhance model performance, increase computational efficiency, and minimize algorithmic complexity. For this task, the Recursive Feature Elimination (RFE) method was selected, as supported by references [29-31]. Specific steps for the RFE implementation will be outlined in subsequent sections.

1. Initial Feature Subset Evaluation: The composite dataset of radiomic and clinical variables serves as the initial feature subset for the Support Vector Machine (SVM) model. Each feature's importance is quantified, and the classification accuracy of this initial feature set is assessed through cross-validation techniques.

2. Iterative Feature Removal and Recalculation: The least impactful feature is excised from the current feature subset, creating a modified feature set. This new set is subsequently input into the SVM model. Feature importance is recalculated and the modified feature subset's classification accuracy is evaluated using cross-validation methods.

3. Optimizing Feature Selection: The procedure delineated in step 2 is recursively applied until no features remain in the subset. Through this iterative process, a total of 15 distinct feature subsets are generated, each comprising a varying number of features. The feature subset yielding the highest classification accuracy is identified as the optimal feature combination for the predictive model.

The clinical settings for our predictive model include:1. Facility type: it is mainly applied to neurosurgery wards in general hospitals or neurosurgery specialty hospitals, which are capable of handling complex neurosurgical procedures.2. Size: it is more suitable to be implemented in large or medium-sized hospitals because these hospitals usually have more case data and experience, which is conducive to the training and validation of the model, and it can be generalized to smaller hospitals after it passes the clinical practice. The modeling background of the prediction model includes:1. Data duration: long-term clinical data, covering relevant case information over the past 10 years, are needed to ensure the stability and accuracy of the model. 2.Data characteristics: the data come from multiple sources, including medical records, imaging, and laboratory tests, reflecting the multiple and complex factors affecting recurrence.3. Modeling purpose: to improve the accuracy of

recurrence prediction through machine learning, help doctors develop more personalized treatment plans, and optimize the allocation of hospital resources. In summary, the clinical environment of the target prediction model is mainly set in the neurosurgical wards of large or medium-sized hospitals, and the modeling background involves long-term clinical data collection and analysis, aiming to improve prediction accuracy and optimize the allocation of medical resources.

Table 3. Setting hyperparameters for CNN.

Hyperparameter	Setting
Activation function	ReLu
Optimizer	Adam
Batch size	64
Dropout	0.5
loss	Binary crossentropy

Table 4. Setting hyperparameters for SVM.

Hyperparameter	Setting
Kernel	rbf
Gamma	0.01
C	50

Results

The initial objective of this study aimed to streamline the clinical application of the predictive model, specifically by enabling direct input of cranial CT images for generating predictive outcomes. Contrary to expectations, models utilizing CT images as the sole input data type demonstrated suboptimal performance across all evaluation metrics, irrespective of the machine learning (ML) algorithm employed. Moreover, models incorporating both clinical data and CT images as input data yielded prediction outcomes significantly inferior to those relying solely on clinical data. Upon utilizing 3D Slicer for radiological feature extraction from the CT images, the resultant model performance exhibited notable improvement. This underscores the inadequacy of current ML algorithms in directly utilizing CT images for clinical research; while predictive results can be generated, they remain unsatisfactory (refer to Table 5 and Figure 4). Consequently, the study abandoned the notion of utilizing image-based input data. On the other hand, radiomics data are already a distillation and abstraction of the CT image information, and using the original image as

input again may lead to information redundancy and even introduce noise, thus affecting the performance of the model. Second, from the perspective of computational efficiency, direct processing of raw CT images usually requires more computational resources and time, whereas radiomics data, as a more compact and higher-level feature representation, can significantly improve the training and prediction speed of the model. Therefore, in this study, we did not choose to try radiomics data + CT images as input data for model validation. Instead, an exploration into the viability of employing clinical data, radiomic features, or a combination thereof as input datasets was conducted to optimize predictive model performance.

Predictive model evaluation

This study employs Accuracy (ACC), Area Under the Curve (AUC), and Recall as evaluative metrics for the predictive models, with corresponding results delineated in Figure 4. ACC serves as a direct indicator of the model's consistency in aligning predictive and actual outcomes. Specifically, ACC represents the ratio of correctly classified samples to the overall sample pool, offering both intuitive understanding and straightforward implementation. It is principally utilized to assess the model's ability to accurately categorize target variables in the predictive outcomes. However, ACC possesses limitations as it exclusively considers the classification of positive samples, thereby omitting negative samples. This lack of comprehensiveness limits ACC's capacity to measure the overlap between predicted and true outcomes. To address this limitation, AUC is incorporated as it holistically considers both positive and negative samples, thereby providing a more nuanced evaluation of model performance. Recall, another metric utilized, is particularly pertinent given the study's objective to predict the recurrence of Chronic Subdural Hematoma (CSDH). Recall quantifies the model's proficiency in accurately identifying positive samples, focusing on True Positive (TP) cases. It is especially vital for this study, as it emphasizes the model's ability to correctly predict patient recurrence. Distinct from ACC and AUC, Recall remains unaffected by the selection of a decision threshold, rendering it more apt for comparing various models, particularly when the decision threshold is ambiguous or challenging to ascertain. In summary, ACC, AUC, and Recall are deployed as multifaceted evaluative metrics to assess the predictive models constructed in this study.

In a comprehensive evaluation across all designated metrics—Accuracy (ACC), Area Under the Curve (AUC), and Recall—the Support Vector Machine (SVM) model consistently outperformed the Convolutional Neural Network (CNN), Random Forest (RF), and Logistic Regression (LR) models. This was observed irrespective of the input dataset employed, be it clinical data, radiomics features, or a hybrid of both. For models utilizing radiomics features, SVM demonstrated improvements of 2.77%, 5.06%, and 13.87% in ACC; 3.41%, 7.30%, and 15.37% in AUC; and 4.18%, 6.93%, and 15.48% in Recall compared to CNN, RF, and LR models, respectively. Similarly, when clinical data served as the input, SVM enhanced ACC by 9.38%, 11.24%, and 19.50%; AUC by 8.84%, 10.13%, and 13.12%; and Recall by 2.98%, 4.79%, and 15.88%. Interestingly, a combination of clinical and radiomics features as input to the SVM model resulted in further performance gains: ACC improved by 5.52%, 7.94%, and 13.23%; AUC by 1.81%, 6.00%, and 9.22%; and Recall by 3.47%, 5.08%, and 13.94% in comparison to CNN, RF, and LR models. These outcomes substantiate the efficacy of the SVM model in predicting postoperative recurrence in Chronic Subdural Hematoma (CSDH) patients. Moreover, it was observed that the hybrid input set comprising both clinical and radiomics data enhanced the performance of the SVM model itself. Specifically, the ACC, AUC, and Recall were higher by 5.6%, 2.97%; 0.56%, 3.84%; and 1.88%, 3.54% respectively when compared to SVM models that utilized either radiomics features or clinical data as standalone inputs. In conclusion, the SVM model, when constructed based on a fusion of clinical and radiomics features, exhibited

superior predictive capabilities, making it the optimal choice for this study.

Table 5. Classification results of different datasets in four machine learning models

Model		Radiomics	Clinical	Clinical-radiomics	CT image	CT image-clinical
SVM	ACC（%）	82.19	84.82	87.79	54.27	57.86
	AUC（%）	85.57	82.29	86.13	56.91	62.75
	Recall（%）	80.82	79.16	82.70	60.54	65.32
CNN	ACC（%）	79.42	75.44	82.27	48.95	58.19
	AUC（%）	82.16	73.45	84.32	51.07	60.24
	Recall（%）	76.64	76.18	79.23	54.29	62.85
RF	ACC（%）	77.13	73.58	79.85	43.55	56.97
	AUC（%）	78.27	72.16	80.13	47.82	60.11
	Recall（%）	73.89	74.37	77.62	46.31	58.63
LR	ACC（%）	68.32	65.32	74.56	42.75	54.32
	AUC（%）	70.13	69.17	76.91	45.68	52.19
	Recall（%）	65.34	63.28	68.76	47.54	55.13

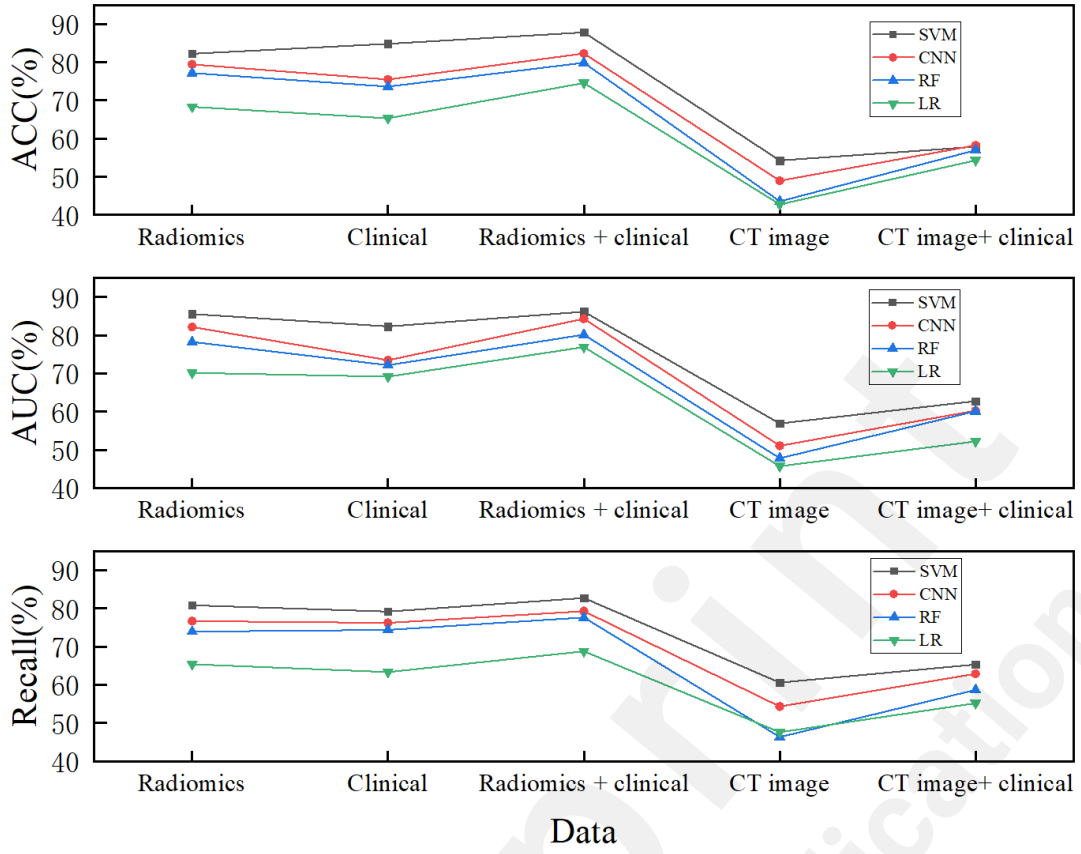


Figure 4. Classification results of different datasets in four machine learning models

While the Support Vector Machine (SVM) model utilizing clinical-radiomics features demonstrated superior performance, it did not meet our predefined target of exceeding 90% across key evaluation metrics. To address this, a feature selection process was implemented to refine the input variables for the predictive model. Our analysis identified the top five influential variables impacting postoperative recurrence in Chronic Subdural Hematoma (CSDH) patients as: history of head trauma, Markwalder Grading Scale, postoperative midline improvement, preoperative midline shift, and radiomics features (Figure 5). History of head trauma is one of the main factors leading to the formation of CSDH; Markwalder Grading Scale is used to assess the severity of CSDH, including hematoma volume, midline shift, state of consciousness and other factors; Postoperative midline improvement is an important index for assessing surgical results and patient's recovery, preoperative midline shift reflects the degree of compression of hematoma on brain tissue, if postoperative midline improvement is poor or larger midline shift usually indicates a poor prognosis and a higher risk of recurrence; Radiomics features can provide detailed information about CSDH, which can help doctors more accurately assess the disease and predict the risk of recurrence. These features are valuable in predicting the risk of recurrence after CSDH and should be emphasized by clinicians during diagnosis.

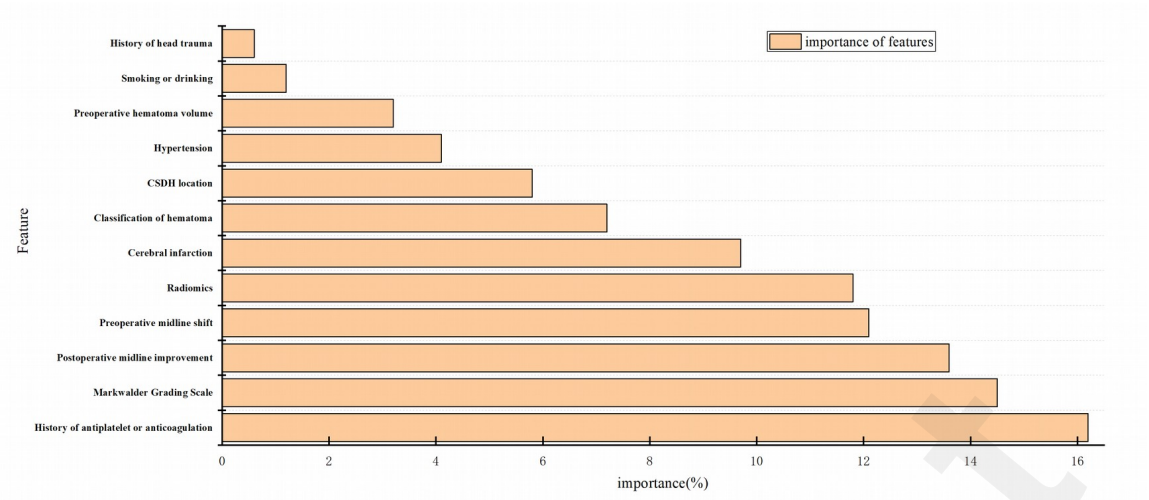


Figure 5. Ranking the importance of these 12 features screened by the RFE method in descending order of importance

Subsequent feature selection experiments, conducted using the Recursive Feature Elimination (RFE) method, indicated an optimal combination of 12 variables (Figure 6). When variables such as Length of Stay in hospital, Gender, and Age were excluded, the model yielded the highest performance across Accuracy (ACC), Area Under the Curve (AUC), and Recall, registering 92.72%, 91.34%, and 93.16%, respectively. Consequently, these 12 features were incorporated into the refined SVM model to yield an optimal predictive tool for assessing the likelihood of postoperative recurrence in CSDH patients.

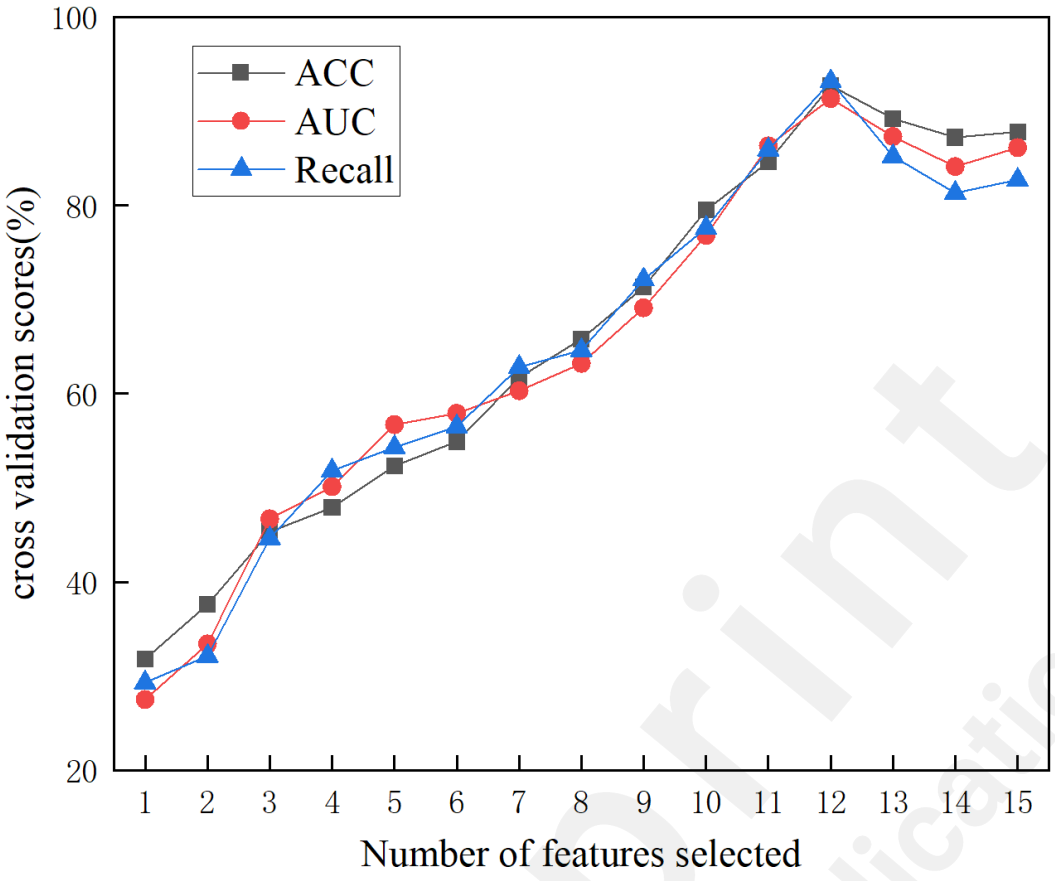


Figure 6. Results of ACC, AUC and Recall after selecting the number of features by cross-validation method

External dataset validation

To corroborate the reliability and generalizability of the developed SVM model, an external validation was performed using clinical-radiomics feature data from 20 Chronic Subdural Hematoma (CSDH) patients, sourced from the First Affiliated Hospital of Anhui Medical University. The inclusion criteria, exclusion criteria, outcome metrics, and predictors for the dataset used for external validation (including clinical data and preoperative head CT images) are identical to those for the modeling dataset (Table 6). There is no significant difference between the distribution of significant variables between the model validation dataset and the model development dataset when comparing Tables 1 and 6. The obtained data were fed into the four pre-established models, and the outcomes are depicted in Figure 7. The ACC metrics for the four models (SVM, Convolutional Neural Networks (CNN), Random Forest (RF), and Logistic Regression (LR)) registered at 90.32%, 84.67%, 81.30%, and 75.78%, respectively. Area Under the Curve (AUC) outcomes were 91.32%, 86.73%, 82.15%, and 72.16%, respectively. Recall rates were recorded at 88.37%, 87.12%, 80.01%, and 74.68%, respectively. Across all evaluation parameters, the SVM model consistently exhibited superior performance. With context, we find that the SVM model constructed based on the fusion of clinical and radiomics features has consistent results in both internal validation and external testing, and performs best among the four models. Consequently, these results reconfirm that the SVM model is the most effective predictive tool for assessing postoperative recurrence in CSDH patients, as further substantiated by this external dataset validation.

The data we used for modeling came from the Second Affiliated Hospital of Anhui Medical University, which mainly serves local patients, so its medical record data may reflect more of the disease characteristics and treatment experiences of local patients. The data used for the external validation of the model came from the First Affiliated Hospital of Anhui Medical University, which attracts patients from all over the province and even the neighboring regions due to the hospital's reputation and geographic location. These patients may have different cultural backgrounds, living habits and medical needs. The two hospitals mentioned above therefore provide a context for the differences between the internal and external datasets, ensuring the general applicability of the prediction model.

Table 6. Clinical variables used to validate the model.

	Non-recurrence（n=17）	Recurrence （n=3）	*P*-value
Age（years）	72.3±10.6	68.0±9.6	0.523
Gender（%）			0.531
Male	14(82.4)	3(100.0)	
Female	3(17.6)	0(0.0)	
Smoking or drinking（%）			0.891
Yes	12(70.6)	2(66.7)	
No	5(29.4)	1(33.3)	
Hypertension（%）			0.718
Yes	4(23.5)	1(33.3)	
No	13(76.5)	2(66.7)	
Cerebral infarction（%）			—
Yes	0(0.0)	0(0.0)	
No	17(100.0)	3(100.0)	
History of head trauma（%）			0.948
Yes	10(58.8)	2(66.7)	
No	7(41.2)	1(33.3)	
History of antiplatelet or anticoagulation（%）			0.335
Yes	2(11.8)	1(33.3)	
No	15(88.2)	2(66.7)	
Markwalder Grading Scale (%)			0.308
0	0(0.0)	0(0.0)	
1	5(29.4)	0(0.0)	
2	9(53.0)	3(100.0)	
3	3(17.6)	0(0.0)	
4	0(0.0)	0(0.0)	
Length of stay in hospital（days）	14.8±5.7	11.7±3.8	0.379
CSDH location（%）			0.579
Unilateral	13(76.5)	3(100.0)	
Bilateral	4(23.5)	0(0.0)	
Classification of hematoma（%）			0.113
Homogenous	4(23.5)	0(0.0)	
Laminar	3(17.7)	0(0.0)	
Separated	6(35.3)	1(33.3)	

Trabecular type	4(23.5)	2(66.7)	
Preoperative hematoma volume（ml）	105.7±27.3	126.2±14.6	0.226
Preoperative midline shift（cm）	1.1±0.4	1.3±0.2	0.385
Postoperative midline improvement（cm）	**0.5±0.3**	**0.6±0.3**	0.585

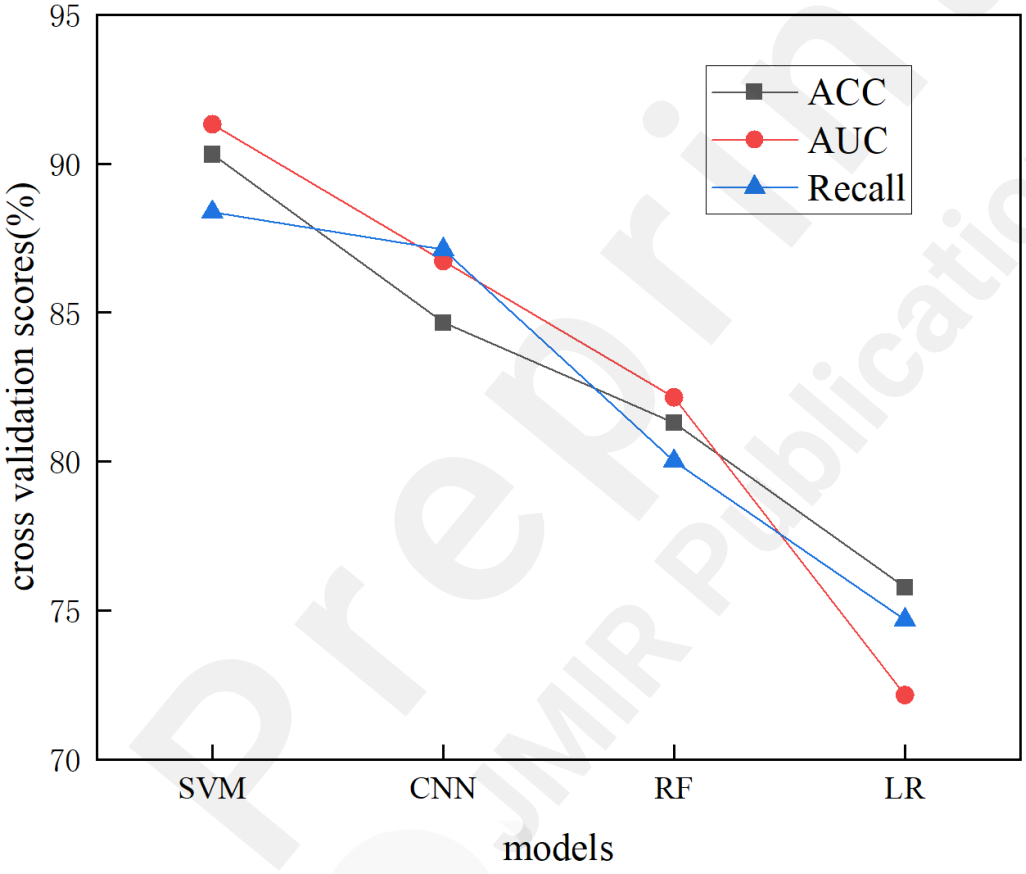


Figure 7. ACC, AUC and Recall for the four models in the external validation set

Discussion

In the current study, the observed postoperative recurrence rate for CSDH patients was 14.3%, a figure that aligns closely with previously published data, ranging from 5% to 26% [19, 20]. This congruence lends credibility to the representativeness of our dataset and suggests that the study's findings have broad applicability.

Within the neurosurgical landscape, CSDH has received limited investigative focus. This is likely due to its relatively high prevalence and standardized treatment approach, coupled with its lower mortality risk compared to other neurosurgical conditions. However, given the global demographic shift toward an older population, the incidence of CSDH—and consequently, its recurrence rate—is witnessing a steady uptick. This escalating trend underscores the need for ongoing research attention, a need that remains largely unmet.

Utilizing extensive data collection, processing, and iterative model optimization, we arrived at an SVM model predicated on clinical-radiomics features that exhibited optimal predictive performance. The final model not only met our predetermined efficacy criteria but also carries clinical utility. To our understanding, this constitutes the inaugural effort to leverage clinical-radiomics features in tandem with ML methodologies for the relatively accurate forecasting of postoperative recurrence in CSDH patients. Thus, the present study paves the way for future research, offering a novel paradigm for evaluating surgical outcomes in this patient cohort.

Our review of existing literature reveals that although there are studies incorporating machine learning (ML) with medical imaging (CT or MRI) for clinical applications, the prevailing approach does not typically leverage raw images for model construction. Instead, researchers employ various software tools or techniques to extract radiomics features from these images, upon which subsequent models are built. Consistent with these findings, our empirical results confirm that predictive models constructed directly from unprocessed images perform suboptimally. The limitations are primarily attributable to the inherent difficulty traditional ML algorithms face in extracting key lesion characteristics—such as location, size, and morphology—from unprocessed CT images. To address these limitations, we utilized the open-source software, 3D Slicer, which offers an array of algorithms for feature extraction, including edge and corner detection, as well as texture analysis. By segmenting the region of interest (ROI) and extracting features using 3D Slicer, we acquired meaningful mathematical attributes, such as gradient and curvature. This enhanced feature set enabled more effective computerized analysis of lesions and improved predictive outcomes. Consequently, we shifted our strategy from using raw CT images to combining extracted radiomics features with clinical data for CSDH patients as input variables in our model construction. The comparative evaluation of models developed through four ML algorithms (SVM, CNN, RF, and LR) revealed the superiority of the hybrid data approach over models built solely on clinical or radiomics data. Further, a performance matrix based on metrics such as ACC, AUC, and Recall indicated marked variations among the four methods, with the SVM model outperforming CNN, RF, and LR on all three metrics (ACC: 2.77%, 5.06%, 13.87%; AUC: 3.41%, 7.30%, 15.44%; Recall: 4.18%, 6.93%, 15.48%). Therefore, our analysis corroborates the high reliability of the SVM-based predictive model constructed using the amalgamated dataset.

Our analysis of the reasons for the poor performance of other models: 1. Mismatch between data characteristics and model complexity: Some highly complex models, such as CNN for deep learning, may encounter overfitting problems when confronted with datasets that are relatively small or not rich enough in feature dimensions. This means that the model performs well on training data, but performs poorly on new, unseen data.2. Feature extraction and exploitation capabilities: while RF and LR have some advantages in dealing with non-linear problems and feature selection, they may not be as good as specially designed algorithms such as SVM with appropriate kernel functions.3. Sensitivity to unbalanced data: models such as RF and LR may be biased towards the majority class if the positive and negative samples are unevenly distributed, which is a common problem in medical image analysis. This may lead to a decrease in the recall of the model and weak identification of the minority class (recurrent cases in this study). And the main reason why SVM performs well in

predicting postoperative recurrence in CSDH patients is that its principle of maximum interval (SVM improves the generalization ability of the model by maximizing the interval between the decision boundary and the support vectors) and the kernel trick (SVM can map the input space to a high-dimensional feature space through the kernel function, in which the nonlinear problem may become linearly differentiable) provide powerful support for dealing with small samples, high dimensionality, and unbalanced medical data provide powerful support.

Our model achieved 90.32% ACC, 91.32% AUC, and 88.37% Recall on an independent external validation set. These metrics indicate that the model has high accuracy in predicting postoperative recurrence in patients with CSDH. For socio-economic purposes, accurate prediction can reduce unnecessary examinations and follow-up visits, which can save a large amount of medical resources every year; accurate prediction can also help doctors to take interventions in advance to reduce the incidence of postoperative complications and readmission rate of patients. This not only reduces patient suffering, but also lowers hospital readmission costs. For patients, it reduces their burden by decreasing unnecessary examinations and follow-up visits, and timely interventions based on the predicted results can help reduce the occurrence of complications and improve their quality of life; in addition, through the model prediction, doctors can provide patients with more personalized treatment plans and care recommendations, which can improve patients' satisfaction and trust. In summary, our model has excellent performance and significant impact on clinical practice and economic benefits.

As early as 2009, Abouzari M et al. [21] explored the utilization of machine learning (ML) algorithms, specifically Artificial Neural Networks (ANN) and Logistic Regression (LR), for prognosis prediction in Chronic Subdural Hematoma (CSDH) patients. Given the technological limitations of that era, these models exhibited low accuracy and questionable evaluation metrics. However, their pioneering work served as a catalyst for us to implement contemporary ML techniques in this research domain.

With respect to postoperative recurrence in CSDH patients, extensive studies have been conducted on hematoma staging using CT imaging. Historically, hematomas were simplistically categorized into four density-based types: low-density, isodense, high-density, and mixed-density [22]. Tsutsumi et al., however, found no statistically significant difference in postoperative recurrence rates when using these classifications [23]. Subsequently, Nakaguchi H [17] introduced an alternative, more nuanced, classification—comprising Homogenous, Laminar, Separated, and Trabecular types—which garnered wide acceptance in the research community.

In the current investigation, we adhered to this latter classification scheme when analyzing CT images. Notably, our data analysis revealed that the "Separated" type constituted a greater fraction of the recurrence group, aligning with prior research findings. However, the proportion of cases classified as "Trabecular" diverged from existing literature. We hypothesize that this discrepancy may be attributable to selection bias arising from our limited dataset.

In the realm of Chronic Subdural Hematoma (CSDH) postoperative recurrence, the scholarly focus has predominantly been on surgical methodologies—recently emphasizing middle meningeal artery embolization—patient age, and the administration of antiplatelet or anticoagulant medications [24-26]. This narrow concentration likely stems from the ubiquity of CSDH and the established efficacy of existing surgical treatments, which generally yield favorable outcomes without posing immediate life-threatening risks to patients. Consequently, research has largely remained at the clinical echelon. However, as technological advancements continue to pervade medical practice, the incorporation of these innovations not only streamlines clinical operations but also enhances patient outcomes, thereby advancing the objective of precision medicine.

In the current investigation, we diverged from the conventional practice of utilizing either clinical data or imaging histology data exclusively. Rather, we integrated both data types and, through comparative analysis, substantiated the superior predictive performance of combined clinical-radiomics features. Furthermore, Support Vector Machine (SVM) was identified as an efficacious classification algorithm particularly suited for the unique characteristics of medical imaging data, which are high-dimensional and often limited in sample size. SVM achieves classification by constructing a hyperplane that aptly segregates distinct feature sets in medical imaging data, thereby facilitating more accurate identification and prediction of postoperative recurrence in CSDH patients. Additionally, SVM exhibits robustness in mitigating the influence of noise and outliers commonly present in radiomics features, thus bolstering the reliability of model predictions.

In the realm of medical research, feature selection predominantly employs filtering methods, including but not limited to correlation coefficients, chi-square tests, and mutual information, to identify variables that highly correlate with the target outcome. Filtering methods excel in computational efficiency, capable of swiftly processing large datasets and thereby reducing dimensionality. These methods also offer adaptability, accommodating user-defined criteria for application-specific scenarios.

Nevertheless, the present study employs Recursive Feature Elimination (RFE) in lieu of filtering methods, and for several substantive reasons:

1. Capability to Manage Highly Correlated Features: Filtering methods struggle with the presence of a multitude of highly correlated variables, a challenge more effectively navigated by RFE.
2. Distributional Assumption Sensitivity: Filtering methods often rest on certain statistical distribution assumptions (e.g., normal, t-distribution), which, if incorrect, compromise feature selection accuracy. Conversely, RFE operates independently of such assumptions.
3. Computational Efficiency: Contrary to common perception, filtering methods, while efficient with smaller datasets, may demand substantial computational resources and time when applied to larger datasets. RFE, on the other hand, demonstrates superior computational efficiency in such contexts.
4. Applicability to Nonlinear Problems: Filtering methods generally rely on linear models, limiting their efficacy for nonlinear challenges. RFE exhibits no such constraint.
5. Automated and Robust Feature Selection: Unique to RFE is its ability to automatically discern the most pertinent feature subset, obviating the need for manual selection. This automation further minimizes overfitting risks and enhances model interpretability by focusing on the most salient features [27,28].

Given these advantages, RFE was selected as the feature selection methodology for this study.

Upon implementing Recursive Feature Elimination (RFE) for feature selection, the predictive model demonstrated robust performance metrics, including an accuracy (ACC) of 92.72%, area under the curve (AUC) of 91.34%, and recall rate of 93.16%. Experimental outcomes identified the top five variables influencing postoperative recurrence in CSDH patients as follows: history of head trauma, (MGS), postoperative midline improvement, preoperative midline shift, and radiomics. Notably, the substantial impact of postoperative midline improvement and preoperative midline shift on the prognosis of CSDH has not been highlighted in extant literature. Consequently, we advocate for the inclusion of these novel factors in future CSDH studies, given their potential clinical relevance.

In the clinical setting, the specific steps we take to implement the model prediction function are as follows:1. Integrate the predictive model into an existing healthcare information system, such as an electronic medical record system or medical image processing software.2. Input data into the system,

which consists of the patient's medical images and relevant clinical information.3. Predictive results are presented to the physician in an easy-to-understand manner, such as a percentage of probability representation.4. Regularly update and maintain the prediction model. This can be done by collecting new clinical data, optimizing algorithm parameters, or adjusting feature selection strategies. Through the above steps, patients can be provided with personalized treatment plans, reducing unnecessary tests and surgeries and improving the efficient use of medical resources. For patients with a higher risk of recurrence, doctors can take treatment measures earlier, thus improving the patient's prognosis. In the event of a discrepancy between the model prediction and the physician's judgment, the physician should first adopt a conservative treatment strategy to ensure patient safety. For example, patients whose model predictions are at low risk of recurrence but whom physicians believe are at higher risk should continue to be closely monitored and followed up. The physician can then compare the model predictions with the patient's actual treatment results and feed this information back to the model developer. Through continuous data feedback and model optimization, the predictive accuracy and generalization ability of the model can be improved. For now, machine learning models are only supplementary tools, and physicians should always make the final decision in conjunction with their own expertise and experience.

In order to realize the effective application of predictive models in clinical practice, we need to establish a stable and reliable data pipeline. First, we need to collect clinical data from CSDH patients, including radiomic features, medical record information, and surgical records. Then, we need to preprocess and feature extract the data in order to feed it into a predictive model. Next, we need to train and validate the model using machine learning algorithms to ensure its predictive accuracy and reliability. Finally, we need to integrate the predictive model into an existing healthcare information system so that it can automatically receive and process patients' clinical data and generate predictions. During the establishment of the data pipeline, we need to consider the quality, integrity and security of the data. We need to ensure the accuracy and consistency of the data to avoid adverse effects on the predicted outcomes. At the same time, we need to ensure data security and privacy to protect patients' privacy rights.

While our findings hold considerable clinical utility and prospective applicability, it is imperative to acknowledge the following limitations of the study.

1. Assumed input and output data formats: The machine learning model used in this paper is based on a specific input data format (e.g., radiological features and clinical data extracted via 3D Slicer software) and assumes that these data are fully representative of the patient's health status and subdural hematoma characteristics. However, this assumption may omit other important biomarkers or unquantified clinical parameters [29,30], which may have an impact on the predictive power of the model. The output data format is assumed to be measured in terms of specific predictive accuracy metrics (e.g., ACC, AUC, and Recall), which may not adequately reflect the utility and sensitivity of the model in different clinical settings.
2. Potential pitfalls in interpreting the model: Although the support vector machine (SVM) model showed good performance in this study, SVMs and other machine learning models are often considered to be "black-box" models, in which the model's decision-making process may not be transparent. may not be transparent, and this lack of interpretability may produce a lack of trust in settings where the model is used to guide clinical decision-making.
3. Potential bias of the data used in modeling: the study was conducted based on a retrospective dataset from a specific healthcare organization, which may be subject to selection bias (e.g., only patients who received surgical treatments were included) and informational bias (data records

may not be completely accurate). In addition, due to the relatively small sample size (133 patients), the complex relationship between CSDH recurrence and multiple underlying factors may not have been adequately captured, which may have affected the model's ability to generalize and predict accuracy.

4. Generalizability of the data: Although the study was externally validated, the validation set consisted of only 20 patients from another healthcare facility, which may not be sufficient to comprehensively assess the ability of the model to generalize across populations and geographic regions. Patient populations in different regions may have different clinical characteristics, such as different treatment modalities and different healthcare resources, all of which may affect the generalizability and accuracy of the model.

Our subsequent research will work to address the above issues.

Conclusion

In this study, we constructed four models to predict postoperative recurrence in patients with CSDH, utilizing ML algorithms and an amalgamated dataset comprising both radiomics attributes and clinical variables. Comparative evaluation revealed that the SVM model, employing this integrated dataset, demonstrated superior predictive accuracy. The model not only outperforms previously established methods but also provides a more specific and comprehensive framework for predicting outcomes. These predictive findings enable healthcare teams to refine clinical decision-making and offer individualized treatment plans. Moreover, patients can engage in proactive follow-up and informed participation in their treatment protocols based on these results. The developed method offers the advantage of real-time updates and holds considerable clinical implications.

Acknowledgments

This research was funded by the Natural Science Foundation of the Anhui Xinhua university (No.2023zr013), and the Anhui Provincial University Students Innovation and Entrepreneurship Program (No.S202312216119, No.S202212216022, and the CAMS Innovation Fund for Medical Sciences (CIFMS) (2022-I2M-C&T-B-063). We did not use generative AI in any processes of the manuscript writing.

Authors' Contributions

JW and XJ designed the study and revised the manuscript and was responsible for the whole study. SL, GX, and HZ collected the data.CF screened and checked the data. YP built the models. CF and YP drafted the manuscript. All authors made substantial contributions to the study and provided the approval of the submitted version.

Conflicts of interest

The authors declare no conflicts of interest.

Abbreviations

ACC: accuracy
AUC: area under the curve
CNN: convolutional neural networks
CSDH: chronic subdural hematoma
LR: linear regression
MGS: Markwalder Grading Scale
ML: machine learning
RF: random forests
RFE: recursive feature elimination
ROI: regions of interest
SVM: support vector machines

References

References

1. Kudo, H.; Kuwamura, K.; Izawa, I.; Sawa, H.; Tamaki, N., Chronic subdural hematoma in elderly people: present status on Awaji Island and epidemiological prospect. *Neurol Med Chir (Tokyo)* **1992,** *32* (4), 207-9. [PubMed]
2. Rust, T.; Kiemer, N.; Erasmus, A., Chronic subdural haematomas and anticoagulation or anti-thrombotic therapy. *J Clin Neurosci* **2006,** *13* (8), 823-7. [PubMed]
3. Feghali, J.; Yang, W.; Huang, J., Updates in Chronic Subdural Hematoma: Epidemiology, Etiology, Pathogenesis, Treatment, and Outcome. *World Neurosurg* **2020,** *141*, 339-345. [PubMed]
4. Deo, R. C., Machine Learning in Medicine. *Circulation* **2015,** *132* (20), 1920-30. [PubMed]
5. Li, C.; Liu, M.; Zhang, Y.; Wang, Y.; Li, J.; Sun, S.; Liu, X.; Wu, H.; Feng, C.; Yao, P.; Jia, Y.; Zhang, Y.; Wei, X.; Wu, F.; Du, C.; Zhao, X.; Zhang, S.; Qu, J., Novel models by machine learning to predict prognosis of breast cancer brain metastases. *J Transl Med* **2023,** *21* (1), 404. [PubMed]
6. Wang, X.; Zhong, J.; Lei, T.; Chen, D.; Wang, H.; Zhu, L.; Chu, S.; Liu, L., An Artificial Neural Network Prediction Model for Posttraumatic Epilepsy: Retrospective Cohort Study. *J Med Internet Res* **2021,** *23* (8), e25090. [PubMed]
7. Yao, Q.; Jia, W.; Chen, S.; Wang, Q.; Liu, Z.; Liu, D.; Ji, X., Machine learning was used to predict risk factors for distant metastasis of pancreatic cancer and prognosis analysis. *J Cancer Res Clin Oncol* **2023,** *149* (12), 10279-10291. [PubMed]
8. Monteiro, M.; Newcombe, V. F. J.; Mathieu, F.; Adatia, K.; Kamnitsas, K.; Ferrante, E.; Das, T.; Whitehouse, D.; Rueckert, D.; Menon, D. K.; Glocker, B., Multiclass semantic segmentation and quantification of traumatic brain injury lesions on head CT using deep learning: an algorithm development and multicentre validation study. *Lancet Digit Health* **2020,** *2* (6), e314-e322. [PubMed]
9. Fang, C.; Pan, Y.; Zhao, L.; Niu, Z.; Guo, Q.; Zhao, B., A Machine Learning-Based Approach to Predict Prognosis and Length of Hospital Stay in Adults and Children With Traumatic Brain Injury: Retrospective Cohort Study. *J Med Internet Res* **2022,** *24* (12), e41819. [PubMed]
10. Gillies, R. J.; Kinahan, P. E.; Hricak, H., Radiomics: Images Are More than Pictures, They Are Data. *Radiology* **2016,** *278* (2), 563-77. [PubMed]

11. Lambin, P.; Leijenaar, R. T. H.; Deist, T. M.; Peerlings, J.; de Jong, E. E. C.; van Timmeren, J.; Sanduleanu, S.; Larue, R.; Even, A. J. G.; Jochems, A.; van Wijk, Y.; Woodruff, H.; van Soest, J.; Lustberg, T.; Roelofs, E.; van Elmpt, W.; Dekker, A.; Mottaghy, F. M.; Wildberger, J. E.; Walsh, S., Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol* **2017,** *14* (12), 749-762. [PubMed]
12. Vallières, M.; Zwanenburg, A.; Badic, B.; Cheze Le Rest, C.; Visvikis, D.; Hatt, M., Responsible Radiomics Research for Faster Clinical Translation. *J Nucl Med* **2018,** *59* (2), 189-193. [PubMed]
13. Esteva, A.; Kuprel, B.; Novoa, R. A.; Ko, J.; Swetter, S. M.; Blau, H. M.; Thrun, S., Dermatologist-level classification of skin cancer with deep neural networks. *Nature* **2017,** *542* (7639), 115-118. [PubMed]
14. Schwier, M.; van Griethuysen, J.; Vangel, M. G.; Pieper, S.; Peled, S.; Tempany, C.; Aerts, H.; Kikinis, R.; Fennessy, F. M.; Fedorov, A., Repeatability of Multiparametric Prostate MRI Radiomics Features. *Sci Rep* **2019,** *9* (1), 9441. [PubMed]
15. Zhao, B.; Tan, Y.; Tsai, W. Y.; Qi, J.; Xie, C.; Lu, L.; Schwartz, L. H., Reproducibility of radiomics for deciphering tumor phenotype with imaging. *Sci Rep* **2016,** *6*, 23428. [PubMed]
16. Markwalder, T. M., Chronic subdural hematomas: a review. *J Neurosurg* **1981,** *54* (5), 637-45. [PubMed]
17. Nakaguchi, H.; Tanishima, T.; Yoshimasu, N., Factors in the natural history of chronic subdural hematomas that influence their postoperative recurrence. *J Neurosurg* **2001,** *95* (2), 256-62. [PubMed]
18. Stanišic, M.; Pripp, A. H., A Reliable Grading System for Prediction of Chronic Subdural Hematoma Recurrence Requiring Reoperation After Initial Burr-Hole Surgery. *Neurosurgery* **2017,** *81* (5), 752-760. [PubMed]
19. Martinez-Perez, R.; Tsimpas, A.; Rayo, N.; Cepeda, S.; Lagares, A., Role of the patient comorbidity in the recurrence of chronic subdural hematomas. *Neurosurg Rev* **2021,** *44* (2), 971-976. [PubMed]
20. Weigel, R.; Schmiedek, P.; Krauss, J. K., Outcome of contemporary surgery for chronic subdural haematoma: evidence based review. *J Neurol Neurosurg Psychiatry* **2003,** *74* (7), 937-43. [PubMed]
21. Abouzari, M.; Rashidi, A.; Zandi-Toghani, M.; Behzadi, M.; Asadollahi, M., Chronic subdural hematoma outcome prediction using logistic regression and an artificial neural network. *Neurosurg Rev* **2009,** *32* (4), 479-84. [PubMed]
22. Nomura, S.; Kashiwagi, S.; Fujisawa, H.; Ito, H.; Nakamura, K., Characterization of local hyperfibrinolysis in chronic subdural hematomas by SDS-PAGE and immunoblot. *J Neurosurg* **1994,** *81* (6), 910-3. [PubMed]
23. Tsutsumi, K.; Maeda, K.; Iijima, A.; Usui, M.; Okada, Y.; Kirino, T., The relationship of preoperative magnetic resonance imaging findings and closed system drainage in the recurrence of chronic subdural hematoma. *J Neurosurg* **1997,** *87* (6), 870-5. [PubMed]
24. Cofano, F.; Pesce, A.; Vercelli, G.; Mammi, M.; Massara, A.; Minardi, M.; Palmieri, M.; D'Andrea, G.; Fronda, C.; Lanotte, M. M.; Tartara, F.; Zenga, F.; Frati, A.; Garbossa, D., Risk of Recurrence of Chronic Subdural Hematomas After Surgery: A Multicenter Observational Cohort Study. *Front Neurol* **2020,** *11*, 560269. [PubMed]
25. Shotar, E.; Meyblum, L.; Premat, K.; Lenck, S.; Degos, V.; Grand, T.; Cortese, J.; Pouvelle, A.; Pouliquen, G.; Mouyal, S.; Boch, A. L.; Carpentier, A.; Sourour, N. A.; Mathon, B.; Clarençon, F., Middle meningeal artery embolization reduces the post-operative recurrence rate of at-risk chronic subdural hematoma. *J Neurointerv Surg* **2020,** *12* (12), 1209-1213. [PubMed]
26. Andersen-Ranberg, N. C.; Poulsen, F. R.; Bergholt, B.; Hundsholt, T.; Fugleholm, K., Bilateral chronic subdural hematoma: unilateral or bilateral drainage? *J Neurosurg* **2017,** *126* (6), 1905-1911. [PubMed]

27. Yan, K., & Zhang, D. Feature selection and analysis on correlated gas sensor data with recursive feature elimination. *Sensors and Actuators B: Chemical,* **2015,** *25(2), 212,* 353–363. [Free Full Text]
28. Chen, X., & Jeong, J. C. Enhanced recursive feature elimination. *Sixth International Conference on Machine Learning and Applications (ICMLA 2007).* [doi:10.1109/icmla.2007.35]
29. Hirai, S.; Yagi, K.; Hara, K.; Kanda, E.; Matsubara, S.; Uno, M., Postoperative recurrence of chronic subdural hematoma is more frequent in patients with blood type A. *J Neurosurg* **2021,** *135* (4), 1203-1207. [PubMed]
30. Matsubara, M.; Yagi, K.; Minami, Y.; Kanda, E.; Sunada, Y.; Tao, Y.; Takai, H.; Shikata, E.; Hirai, S.; Matsubara, S.; Uno, M., Preoperative elevated eosinophils in peripheral blood for prediction of postoperative recurrence of chronic subdural hematoma. *J Neurosurg* **2023,** *139* (3), 708-713. [PubMed]
31. Pan Y, Fang C, Zhu X, Wan J. Construction of a predictive model based on MIV-SVR for prognosis and length of stay in patients with traumatic brain injury: Retrospective cohort study. *Digit Health*. **2023**;9:20552076231217814. [PubMed]

Appendix

Items of reporting predictive models in biomedical research.

Item number	Section	Topic	Checklist item	Is this entry included
1	Title	Nature of study	Identify the report as introducing a predictive model.	Yes
2	Abstract	Structured summary	Background.	Yes
			Objectives.	Yes
			Data sources.	Yes
			Performance metrics of the predictive model or models.	Yes
			Conclusion including the practical value of the developed predictive model or models.	Yes
3	Introduction	Rationale	Identify the clinical goal.	Yes
			Review the current practice and prediction accuracy of any existing models.	Yes
4		Objectives	State the nature of study being predictive modeling, defining the target of prediction.	Yes
			Identify how the prediction problem may benefit the clinical goal.	Yes
5	Methods	Describe the setting	Identify the clinical setting for the target predictive model.	Yes
			Identify the modeling context in terms of facility type, size, volume, and duration of available data.	Yes
6		Define the prediction	Define a measurement for the prediction goal.	Yes

		problem	Determine that the study is retrospective or prospective.	Yes
			Identify the problem to be prognostic or diagnostic.	Yes
			Determine the form of the prediction model: (1) classification if the target variable is categorical, (2) regression if the target variable is continuous, (3) survival prediction if the target variable is the time to an event.	Yes
			Translate survival prediction into a regression problem, with the target measured over a temporal window following the time of prediction.	No
			Explain practical costs of prediction errors.	
			Defining quality metrics for prediction models.	Yes
			Define the success criteria for prediction.	Yes
7		Prepare data for model building	Identify relevant data sources and quote the ethics approval number for data access.	Yes
			State the inclusion and exclusion criteria for data.	Yes
			Describe the time span of data and the sample or cohort size.	Yes
			Define the observational units on which the response variable and predictor variables are defined.	Yes
			Define the predictor variables. Extra caution is needed to prevent information leakage from the response variable to predictor variables.	Yes
			Describe the data preprocessing performed, including data cleaning and transformation.	Yes
			Remove outliers with impossible or extreme responses; state any criteria used for outlier removal.	Yes
			State how missing values were handled.	Yes
			Describe the basic statistics of the dataset, particularly of the response variable. These include	Yes

			the ratio of positive to negative classes for a classification problem and the distribution of the response variable for regression problem.	
			Define the model validation strategies.	Yes
			Internal validation is the minimum requirement; external validation should also be performed whenever possible.	Yes
			Specify the internal validation strategy.	Yes
			Common methods include random split, time-based split, and patient-based split.	Yes
			Define the validation metrics. For regression problems, the normalized root-mean-square error should be used. For classification problems, the metrics should include sensitivity, specificity, positive predictive value, negative predictive value, area under the ROC curve, and calibration plot	Yes
			For retrospective studies, split the data into a derivation set and a validation set. For prospective studies, define the starting time for validation data collection.	Yes
8		Build the predictive model	Identify independent variables that predominantly take a single value.	Yes
			Identify and remove redundant independent variables.	Yes
			Identify the independent variables that may suffer from the perfect separation problem.	Yes
			Report the number of independent variables, the number of positive examples, and the number of negative examples.	Yes
			Assess whether sufficient data are available for a good fit of the model. In particular, for classification, there should be a sufficient number of observations in both positive and negative	Yes

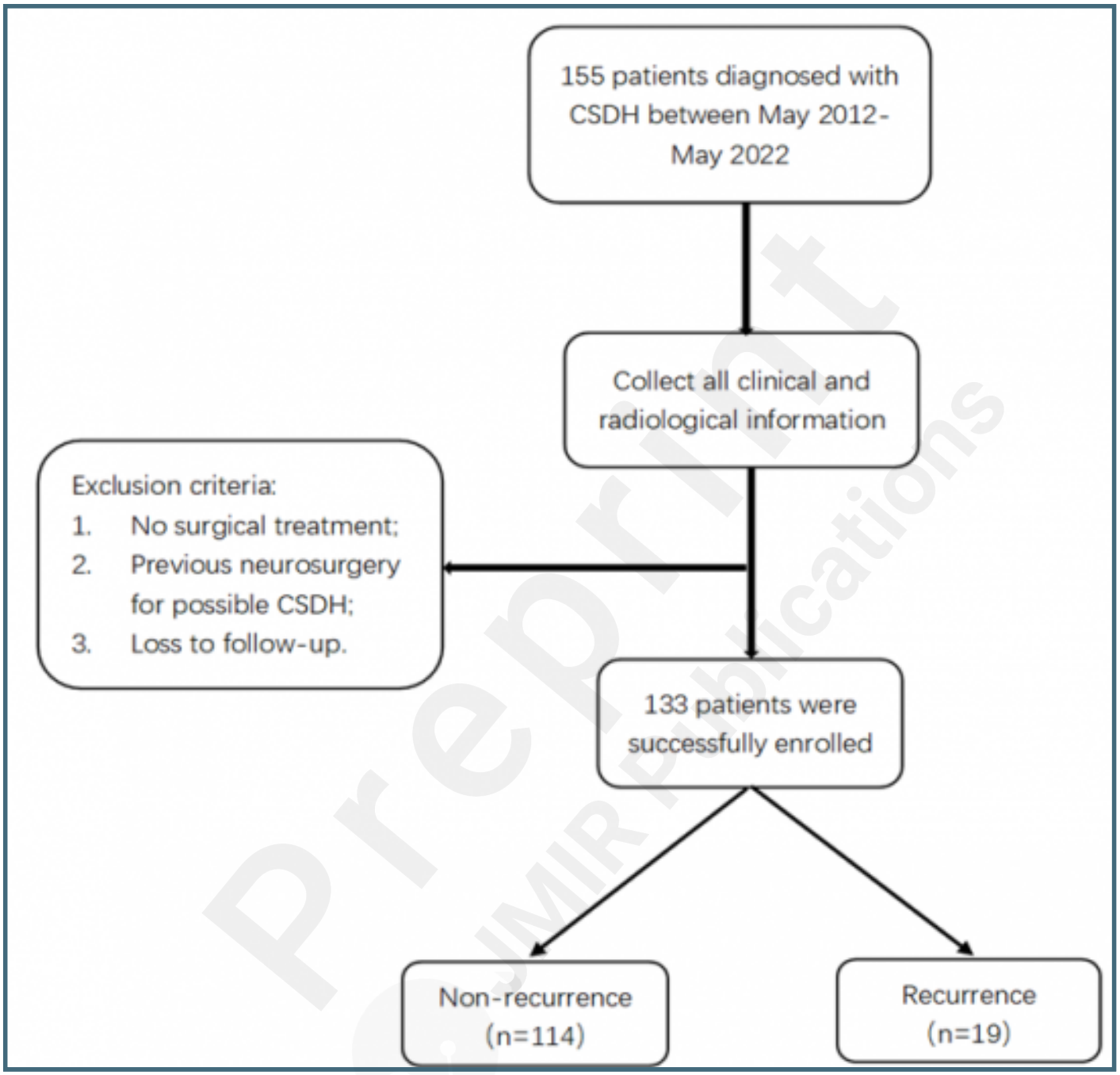
			classes.	
			Determine a set of candidate modeling techniques (eg, logistic regression, random forest, or deep learning). If only one type of model was used, justify the decision for using that model.	Yes
			Define the performance metrics to select the best model.	Yes
			Specify the model selection strategy.	Yes
			Common methods include K-fold validation or bootstrap to estimate the lost function on a grid of candidate parameter values. For K-fold validation, proper stratification by the response variable is needed.	Yes
			For model selection, include discussion on (1) balance between model accuracy and model simplicity or interpretability, and (2) the familiarity with the modeling techniques of the end user.	Yes
9	Results	Report the final model and performance	Report the predictive performance of the final model in terms of the validation metrics specified in the methods section.	Yes
			If possible, report the parameter estimates in the model and their confidence intervals. When the direct calculation of confidence intervals is not possible, report nonparametric estimates from bootstrap samples.	No
			Comparison with other models in the literature should be based on confidence intervals.	No
			Interpretation of the final model. If possible, report what variables were shown to be predictive of the response variable. State which subpopulation has the best prediction and which subpopulation is most difficult to predict.	Yes
10	Discussion	Clinical implications	Report the clinical implications derived from the obtained	Yes

			predictive performance.	
11		Limitations of the model	Discuss the following potential limitations: • Assumed input and output data format • Potential pitfalls in interpreting the model[a] • Potential bias of the data used in modeling • Generalizability of the data	Yes
12		Unexpected results during the experiments	Report unexpected signs of coefficients, indicating collinearity or complex interaction between predictor variables.	No

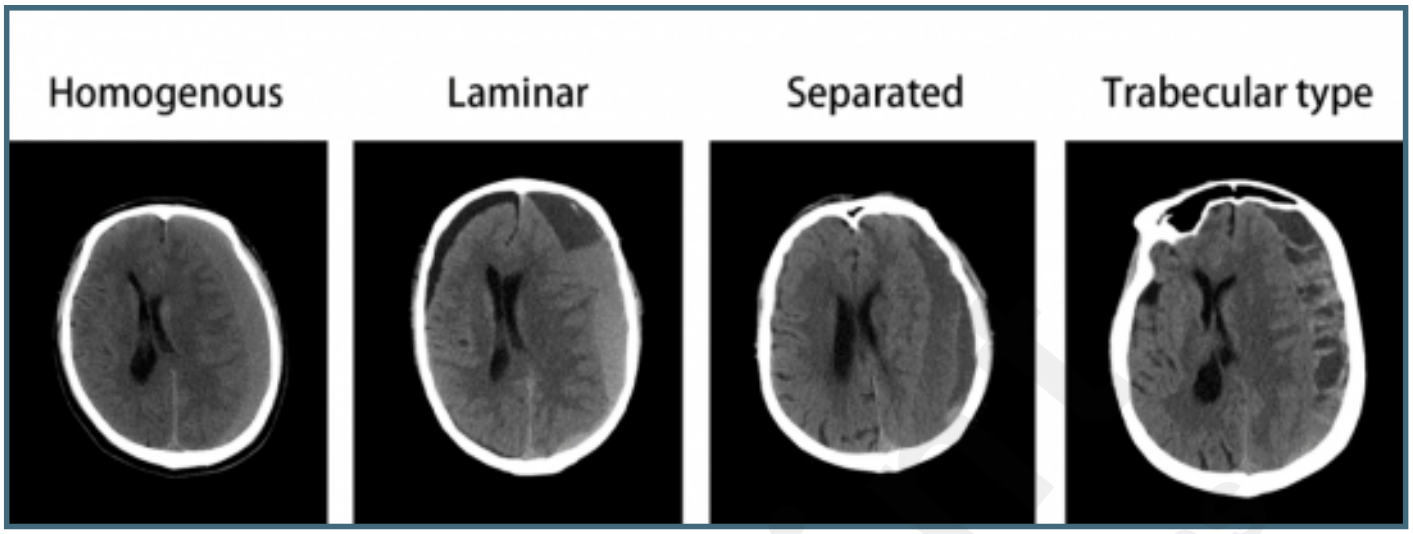
Supplementary Files

Figures

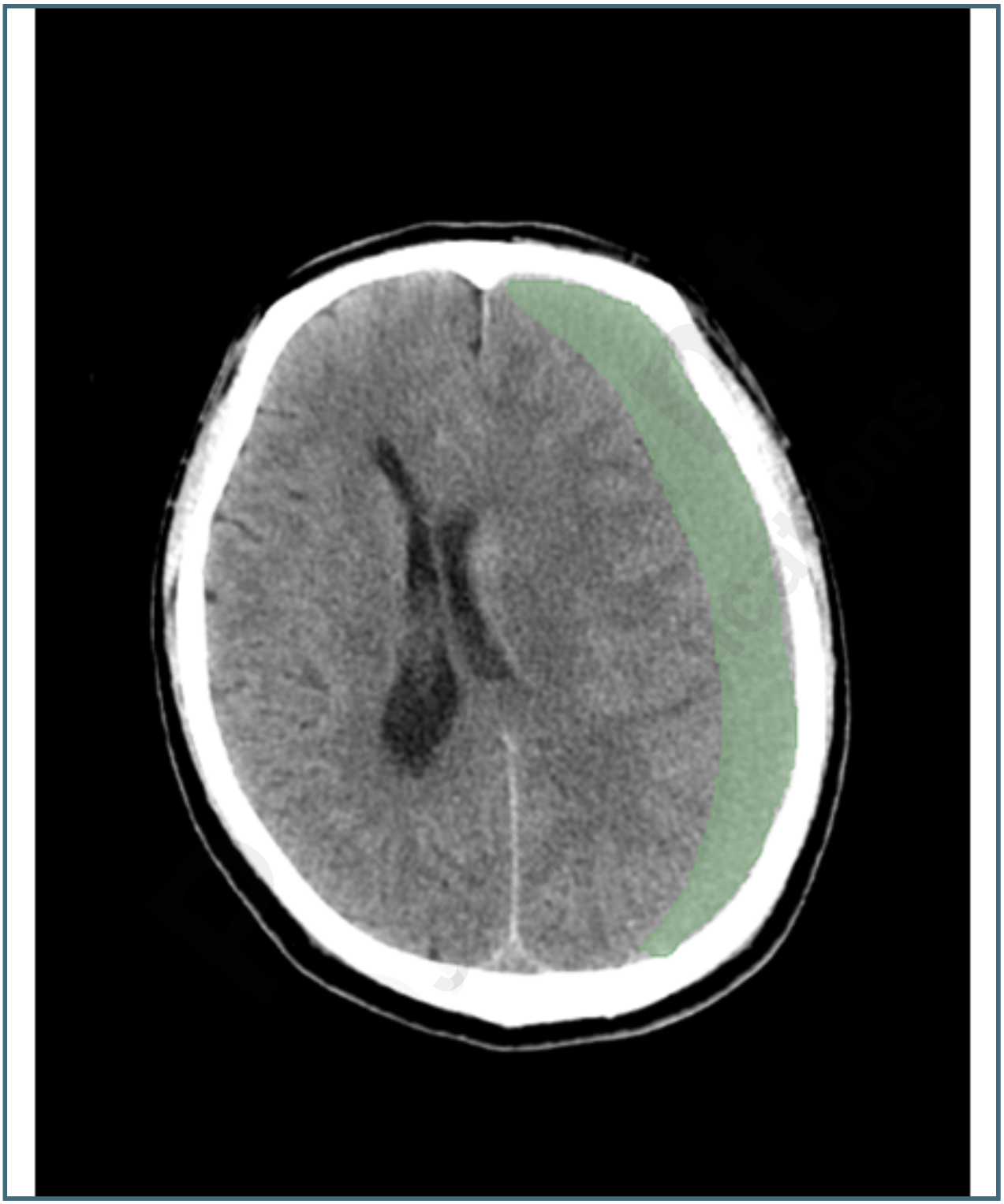
Patient selection and enrollment process.



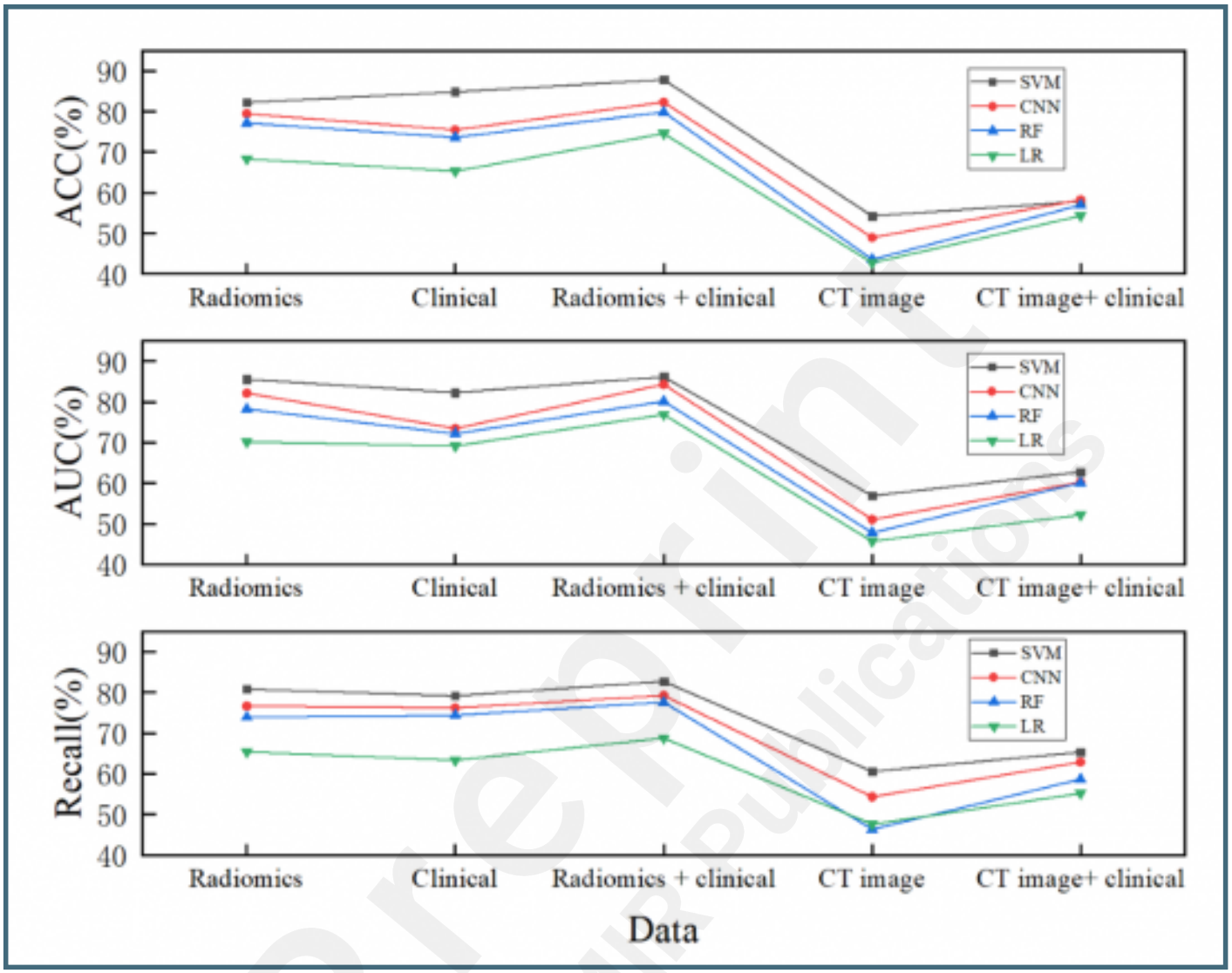
Classification of hematoma.



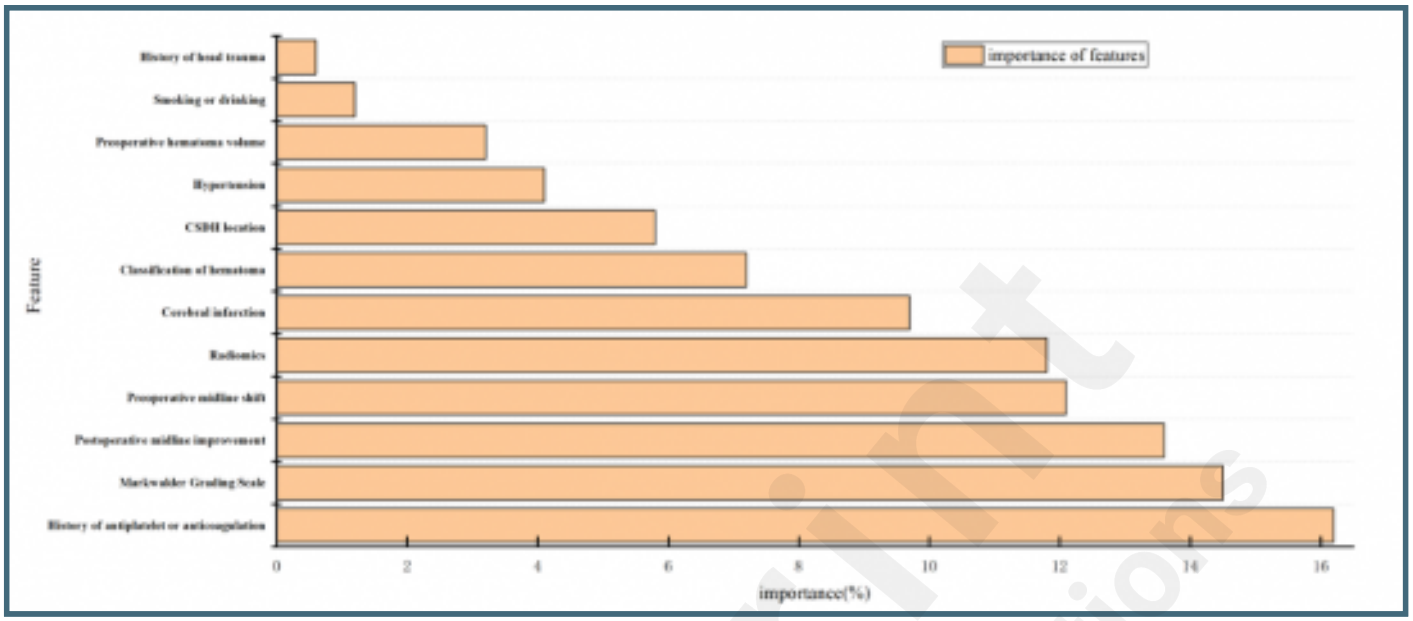
Hand-delineated regions of interest.



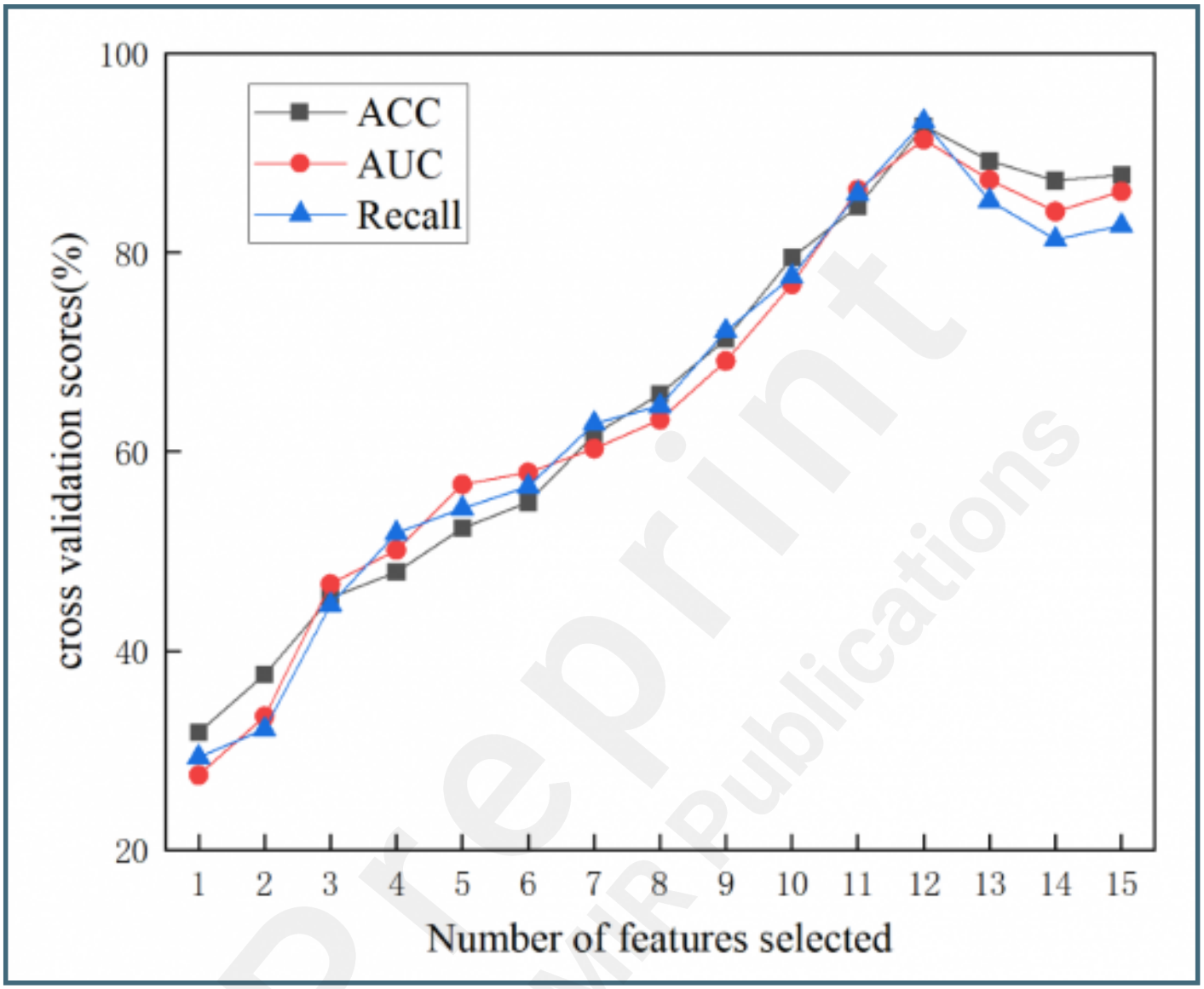
Classification results of different datasets in four machine learning models.



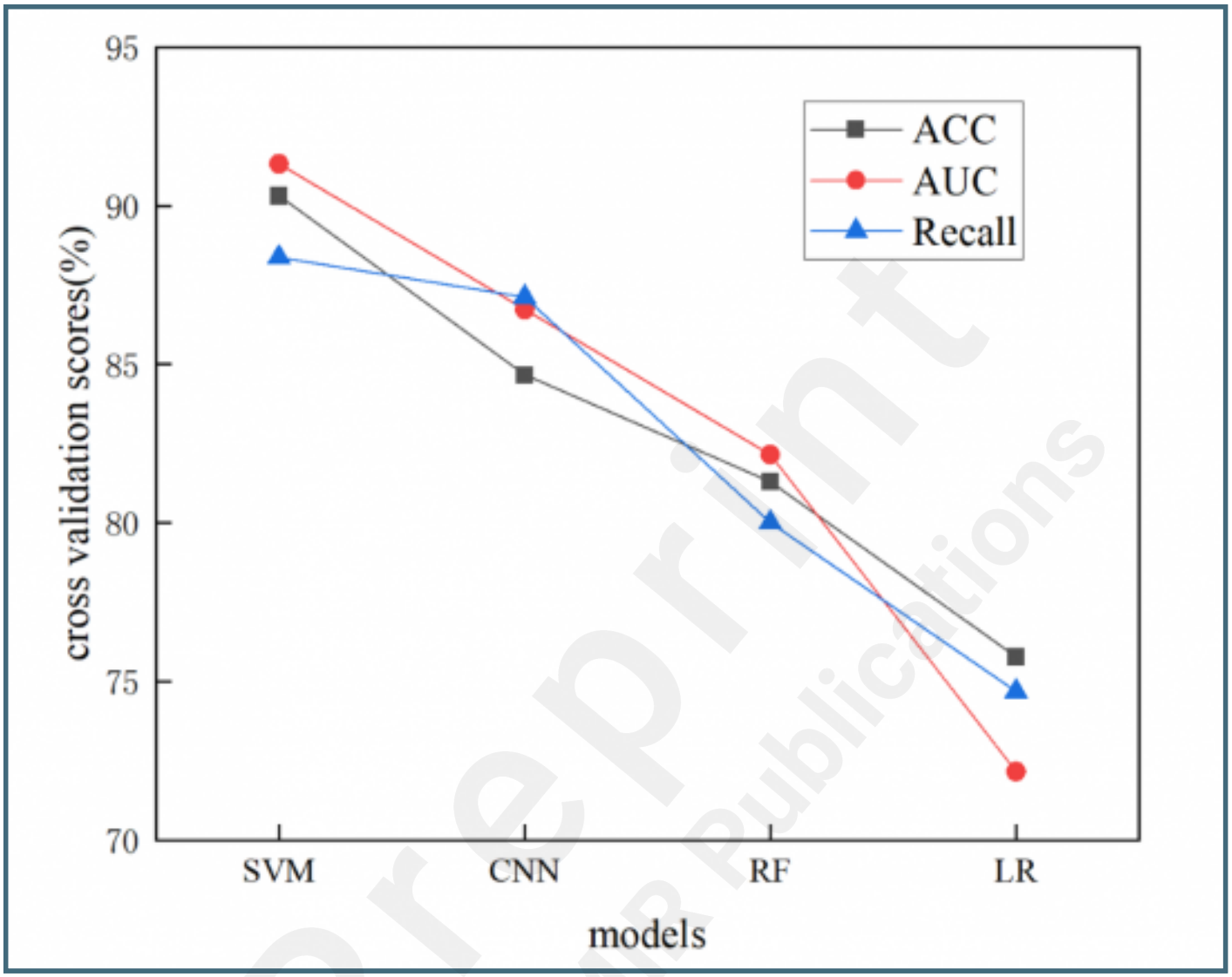
Ranking the importance of these 12 features screened by the RFE method in descending order of importance.



Results of ACC, AUC and Recall after selecting the number of features by cross-validation method.



ACC, AUC and Recall for the four models in the external validation set.



Multimedia Appendixes

Items of reporting predictive models in biomedical research.
URL: http://asset.jmir.pub/assets/2c45c42ce61eb1b1d58a1ea06ab4a820.docx