CT-based Intra-thrombus and Peri-thrombus Radiomics for Prediction of Prognosis After Endovascular Thrombectomy: A Retrospective Study Across Two Centers

Minda Li, Jingxuan Jiang, Gu Hongmei, Hu Su, Wang Jingli, and Chunhong Hu

ABSTRACT

BACKGROUND AND PURPOSE: Complications from endovascular thrombectomy (EVT) can negatively affect clinical outcomes, making the development of a more precise and objective prediction model essential. This research aimed to assess the effectiveness of radiomic features derived from pre-surgical CT scans in predicting the prognosis post- EVT in acute ischemic stroke patients.

MATERIALS AND METHODS: This investigation included 336 acute ischemic stroke patients from two medical centers, spanning from March 2018 to March 2024. The participants were split into a training cohort of 161 patients and a validation cohort of 175 patients. Patient outcomes were rated with the mRS: 0-2 for good, 3-6 for poor. A total of 428 radiomic features were derived from intra-thrombus and peri-thrombus regions in non-contrast CT and CT angiography images. Feature selection was conducted using a least absolute shrinkage and selection operator regression model. The efficacy of eight different supervised learning models was assessed using the area under the curve (AUC) of the receiver operating characteristic curve.

RESULTS: Among all models tested in the validation cohort, the logistic regression algorithm for combined model achieved the highest AUC (0.87, with a 95% confidence interval of 0.81 to 0.92), outperforming other algorithms. The combined use of radiomic features from both the intra-thrombus and peri-thrombus regions significantly enhanced diagnostic accuracy over models using features from a single region（0.81 vs 0.70, 0.77）, highlighting the benefit of integrating data from both regions for improved prediction.

CONCLUSIONS: The findings suggest that a combined radiomics model based on CT imaging serves as a potent approach to assessing the prognosis following EVT. The logistic regression model, in particular, proved to be both effective and stable, offering critical insights for the management of stroke.

ABBREVIATIONS: AUC=area under the curve; EVT=endovascular thrombectomy; KNN=k-nearest neighbors; LASSO=least absolute shrinkage and selection operator; LightGBM=Light Gradient Boosting Machine; LR=logistic regression; MLP=multi-layer perceptron; RF=random forest; SVM=support vector machine; XGBoost=extreme gradient boosting.

Received month day, year; accepted after revision month day, year.

From the Department of Radiology (M.L, H.S, C.H), The First Affiliated Hospital of Soochow University, Suzhou, China. Department of Radiology (M.L, J.J, G.H), Affiliated Hospital of Nantong University, Nantong, China. Institute of Diagnostic and Interventional Radiology (J.J), Shanghai Sixth People's Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai, China. Institute of Medical Imaging (H.S, C.H), Soochow University, Suzhou, China. Stroke center (W.J), Affiliated Hospital of Nantong University, Nantong, China.

The authors declare no conflicts of interest related to the content of this article.

Please address correspondence to Chunhong Hu, MD, Department of Radiology, The First Affiliated Hospital of Soochow University, No. 188, Shizi Street, Suzhou 215006, China.sdhuchunhong@sina.com.

SUMMARY SECTION

PREVIOUS LITERATURE: Prior studies have examined the predictive value of CT-derived thrombus radiomics in stroke, focusing mainly on intra-thrombus features for predicting thrombectomy or thrombolysis outcomes. These investigations highlighted the role of CT signs and thrombus properties in prognostication but also revealed limitations due to reliance on subjective interpretation and a singular focus on intra-thrombus analysis.

KEY FINDINGS: Our study validates a CT-based combined radiomics model utilizing both intra-thrombus and peri-thrombus features, with LR demonstrating the highest predictive accuracy for post-EVT outcomes in stroke.

KNOWLEDGE ADVANCEMENT: The research advances understanding by integrating peri-thrombus features into predictive modeling, offering a more comprehensive and objective analysis that surpasses traditional evaluations, thereby enhancing stroke outcome predictions.

INTRODUCTION

Acute ischemic stroke significantly contributes to death, disability, and high morbidity globally, significantly impacting global mortality rates[1]. Endovascular thrombectomy (EVT) is the standard treatment recommended for patients experiencing acute anterior-circulation large vessel occlusion[2, 3]. However, endovascular thrombectomy (EVT) is associated with certain complications, such as intracranial hemorrhage and malignant brain edema[4-8]. The presence of these conditions notably compromises the probability of positive clinical outcomes and elevates the risk of death. Given the heterogeneity of functional outcomes even after a successful procedure, there is an urgent need to both identify patients who are suitable for EVT and predict early poor outcomes. These efforts can subsequently enhance patient prognosis.

In recent years, certain radiological signs such as the high-density middle cerebral artery sign, intracranial high-density areas, large ischemic cores, and mismatch of CT perfusion have been identified as predictors of the clinical outcomes[9-12]. Although these CT features provide valuable insights, their evaluation depends heavily on the subjective interpretation by radiologists and may not be sufficient for accurate prognosis prediction. Thus, developing a more precise and objective prediction model is essential.

Radiomics analysis offers a quantitative approach by analyzing the variations in gray levels between pixels, allowing for a detailed, high-throughput examination of imaging data that surpasses the conventional visual assessments performed by experts. This method holds promise for enhancing diagnostic accuracy. Machine learning models such as support vector machine (SVM), logistic regression (LR), and random forest (RF) have proven effective in delivering precise predictions, thus aiding healthcare professionals in refining stroke management and enhancing patient outcomes[13]. Recent progress in CT-derived radiomics, particularly in analyzing thrombus properties, has shown potential in forecasting various clinical outcomes. While previous research has validated the effectiveness of CT-derived thrombus radiomics in determining thrombus age, composition, and origin, as well as predicting outcomes after thrombectomy or thrombolysis treatments, and employing a CTA-based thrombus radiomics model to estimate the timing of stroke onset[14-18], these studies have primarily focused on intra-thrombus features. There remains a significant research gap in examining peri-thrombus areas and their role in predicting clinical outcomes after EVT.

Consequently, this research seeks to evaluate the predictive capacity of both intra-thrombus and peri-thrombus radiomic features extracted from CT scans for clinical outcomes post-EVT. We also aim to identify the most effective machine learning classifier for this purpose through rigorous statistical analysis.

MATERIALS AND METHODS

Patients

This research adhered to the Declaration of Helsinki's guidelines and received approval from the ethics committees of the participating hospitals, and was granted a waiver from the need for informed consent. We performed a retrospective analysis on acute stroke patients who were admitted to two medical centers between March 2018 and March 2024. The inclusion criteria included: (1) acute stroke resulting from anterior-circulation large vessel occlusion; (2) visible thrombus related signs on initial NCCT or CTA at admission; (3) mRS score prior to stroke less than 3; (4) subsequent immediate EVT; and (5) comprehensive demographic and clinical data availability. Criteria for exclusion were inadequate imaging clarity due to motion or metal artifacts and incomplete clinical records. Collected clinical data encompassed age, sex, medical history (including hypertension, diabetes, hyperlipidemia, atrial fibrillation, and coronary artery disease), and NIHSS score at admission. In this study, “prognosis” is defined as the clinical outcomes observed 90 days after EVT, as gauged by the mRS. The assessment specifically targets the restitution of motor function and the frequency of major complications. At 90 days, two specialized stroke neurologists (JLW, JXJ) conducted a systematic evaluation of the mRS scores. Patients were categorized into two groups according to their mRS scores: the good outcome group with mRS scores less than 3, and the poor outcome group with mRS scores ranging from 3 to 6. Patients from Center A were assigned to the training cohort, while those from Center B were designated to the validation cohort. The patient selection process and analytical pathway are depicted in Figure 1. For further details on the code and model files, please contact the corresponding author via email.

CT data acquisition and thrombus segmentation

The radiomics process encompassed outlining the ROI, extraction of radiomic features, feature selection, and construction of predictive models (Figure 2). NCCT and CTA imaging was executed using 64- to 256-slice CT scanners from two vendors (SOMATOM Force by Siemens, Germany; GE Revolution and GE Optime CT680 by GE, USA), set with a reconstruction slice thickness between 0.63 and 1.00 mm. Prior to thrombus segmentation, all CT scans were subject to intensity normalization, adjusting the intensity values to a 0-600 range. The images were also adjusted to a uniform resolution of 1 × 1 × 1 mm to standardize voxel dimensions. Thrombus-associated ROIs were outlined using ITK-SNAP software (Version 3.6.0; http://www.itksnap.org/pmwiki/pmwiki.php)，referencing DSA images with the method used in our previous study[15]. After segmenting the intra-thrombus regions, the peri-thrombus areas were automatically segmented by increasing the radius by 1 mm from the original ROIs using Python (Version 2.7.13). To assess the precision of segmentation, 30 thrombi selected at random were delineated two times from CTA scans via one radiologist (MDL) within a two-week period and independently verified by another radiologist (HMG). Both readers were unaware of the patients' clinical data during the segmentation process.

Feature Extraction and Selection

Following the delineation of intra-thrombus and peri-thrombus regions, radiomic features were obtained via the PyRadiomics library (https://pypi.org/project/pyradiomic/). From both regions on NCCT and CTA scans, 428 features were derived in total. To normalize these features and reduce variability across variances, Z-score normalization was applied, scaling the features to a 0–1 range in the training cohort. This normalization process was replicated in the validation datasets as well. Feature selection was carried out on the training cohort via the Mann–Whitney U test to screen out redundant radiomic features, keeping only those significant at a level of $p <$

0.05. To assess the inter-feature relationships, Spearman's rank correlation coefficients were calculated, and only features that demonstrated a correlation coefficient above 0.9 with at least two other features were kept. The refined dataset was then subjected to the least absolute shrinkage and selection operator (LASSO) regression model to develop a predictive radiomic signature. In this study, we conducted k-fold cross-validation as part of our regularization process, specifically tuning the Lambda (λ) parameter to select features optimally.

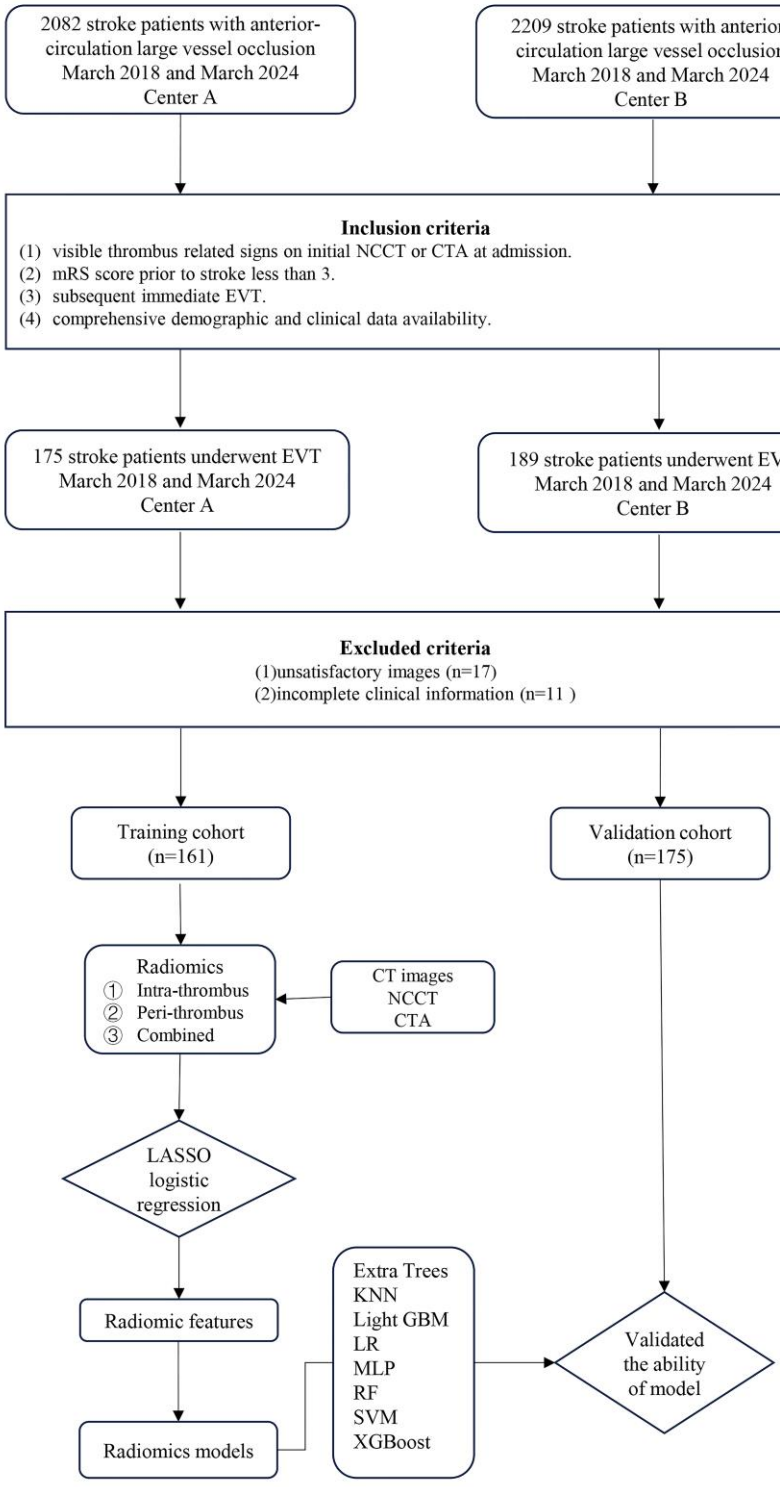


FIG 1. Flow chart of the patient-selection process.

Classifier Model Building and Evaluation

For the selection of radiomic features, the max-relevance and min-redundancy method followed by the LASSO technique were implemented sequentially. This method ranked radiomic features with an intraclass correlation coefficient above 0.90 based on their relevance-redundancy index. From this ranking, the top 10 features exhibiting the highest relevance were preserved. These chosen features were further refined through the LASSO classifier to pinpoint an optimized subset for model development. A radiomics signature was established via multiple logistic regression, utilizing the selected features, and a Rad-score was computed by summing

these features, each weighted by its respective coefficient.

To evaluate the clinical differences between patient groups with and without good outcome, both univariate and multivariate analyses were performed. Additionally, eight supervised machine learning algorithms—RF, LR, SVM, k-nearest neighbors (KNN), extra trees, Light Gradient Boosting Machine (LightGBM), multi-layer perceptron (MLP), and extreme gradient boosting (XGBoost)—were utilized as classifiers. After feature selection through the LASSO method, these features were incorporated into the models, and a 5-fold cross-validation strategy was adopted to confirm the final Rad signature. The DeLong test was used to statistically assess differences in predictive performance among the radiomics models (intra-thrombus, peri-thrombus, and combined models). The CheckList for EvaluAtion of Radiomics study served as the guideline for standardized reporting in this radiomics research. The optimal classification algorithm was identified to develop a clinical prediction model that incorporates selected clinical variables.

Statistical Analysis

Clinical characteristics were evaluated via the t-test, Mann–Whitney U test, or chi-squared test as appropriate. To analyze the correlations among features, Spearman's rank correlation coefficient was employed, retaining those features with a coefficient above 0.9. The consistency of the ROI delineation was verified using the intraclass correlation coefficient, with an intraclass correlation coefficient greater than 0.75 indicating strong reliability. The predictive models' efficacy post-EVT was assessed with receiver operating characteristic curve analysis and DeLong test for variations. A p-value below 0.05 was deemed statistically significant. To mitigate Type I errors from multiple comparisons, we utilized False Discovery Rate corrections in our analysis.

We hereby present this article following the STARD reporting checklist （Online Supplemental Data）.

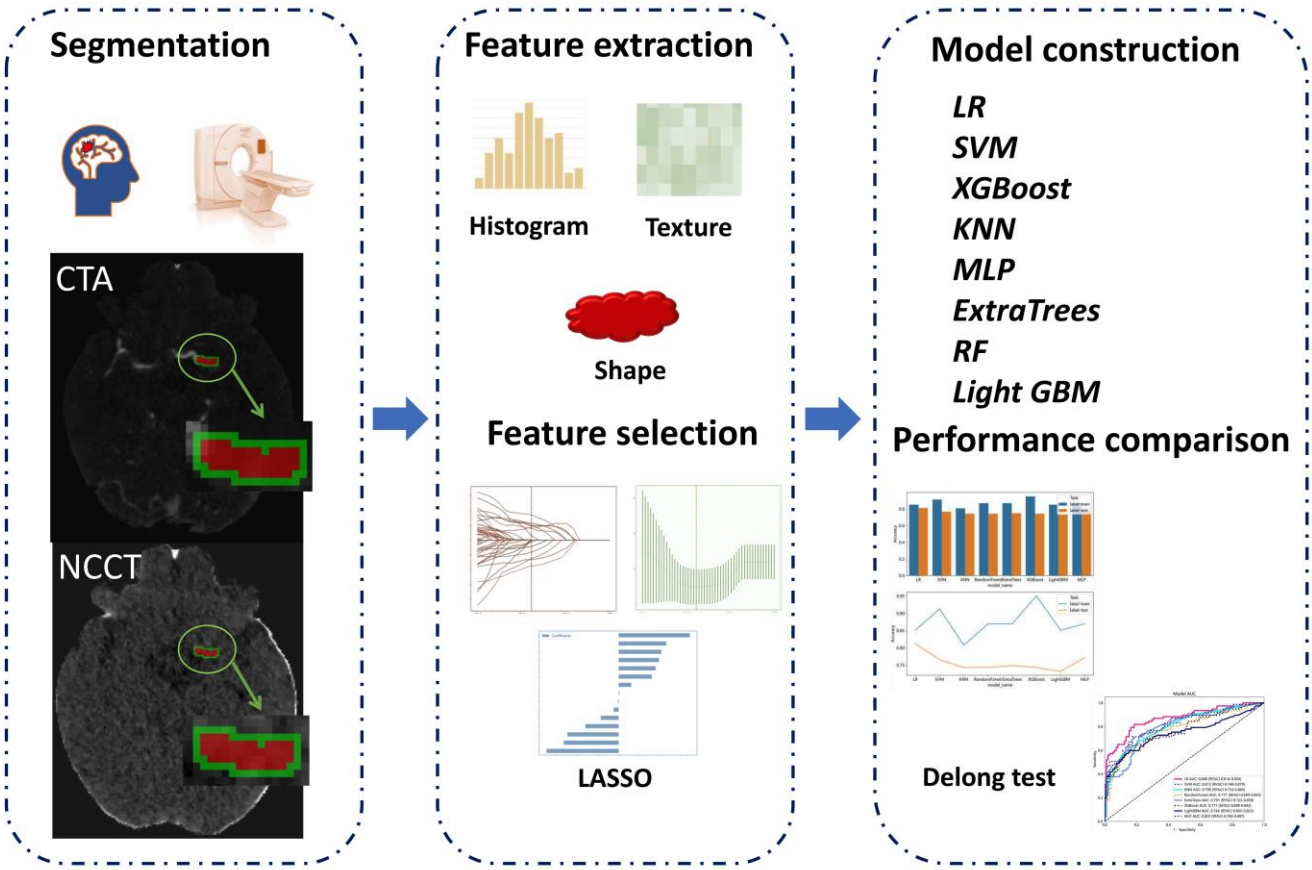


FIG 2. Workflow of the CT-based radiomics model.

RESULTS

Patient Characteristics

In our study, 336 stroke patients were carefully chosen based on defined criteria, of which 128 (38.1%) assessed as poor outcome following EVT. The study divided these patients into two cohorts: 161 from center A formed the training group, and 175 from center B comprised the validation group. Table 1 provides a detailed summary of the demographic and clinical characteristics of the patients, categorized based on the outcome after EVT, for both the training and validation groups. The analysis showed no significant statistical differences in gender, age, hypertension, hyperlipidemia, diabetes, smoking habits, coronary heart disease, or NIHSS scores between the good outcome and poor outcome groups across both cohorts. However, a notable statistical difference was found in the incidence of atrial fibrillation between the groups. No statistically significant variance was observed when comparing the two cohorts.

Feature Extraction and Selection

The intraclass correlation coefficient values demonstrated strong agreement (0.75-0.90) for the radiomic features. After confirming this consistency, all pertinent radiomic features were extracted and used to build predictive models. Ultimately, the Rad scores were formulated using 6, 12, and 15 features with non-zero coefficients for the intra-thrombus, peri-thrombus, and combined models, respectively, as shown in Figure 3. Detailed information on the chosen radiomic features can be found in the Online Supplementary Material. It was noted that all selected radiomic features originated from CTA images, with no features being chosen from NCCT

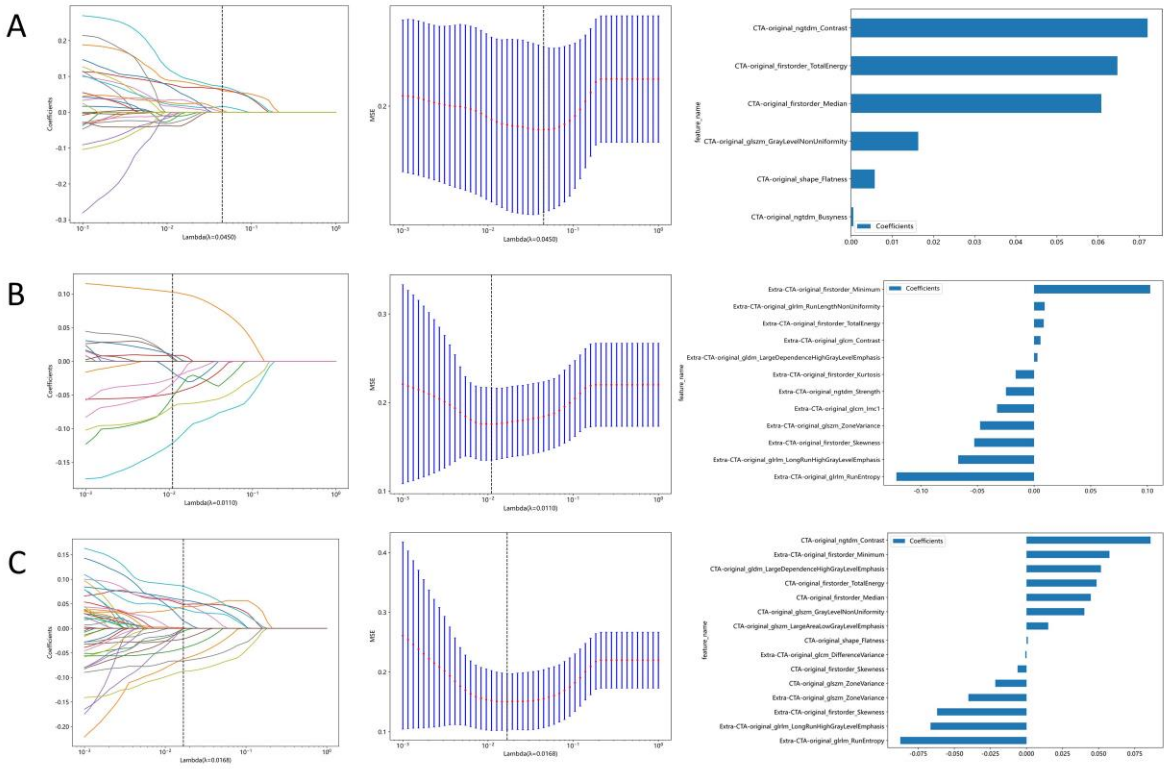


FIG 3. Radiomics feature selection based on the Least absolute shrinkage and selection operator (LASSO) algorithm and Rad score based on intra-thrombus(A), peri-thrombus(B) and combined regions (C).

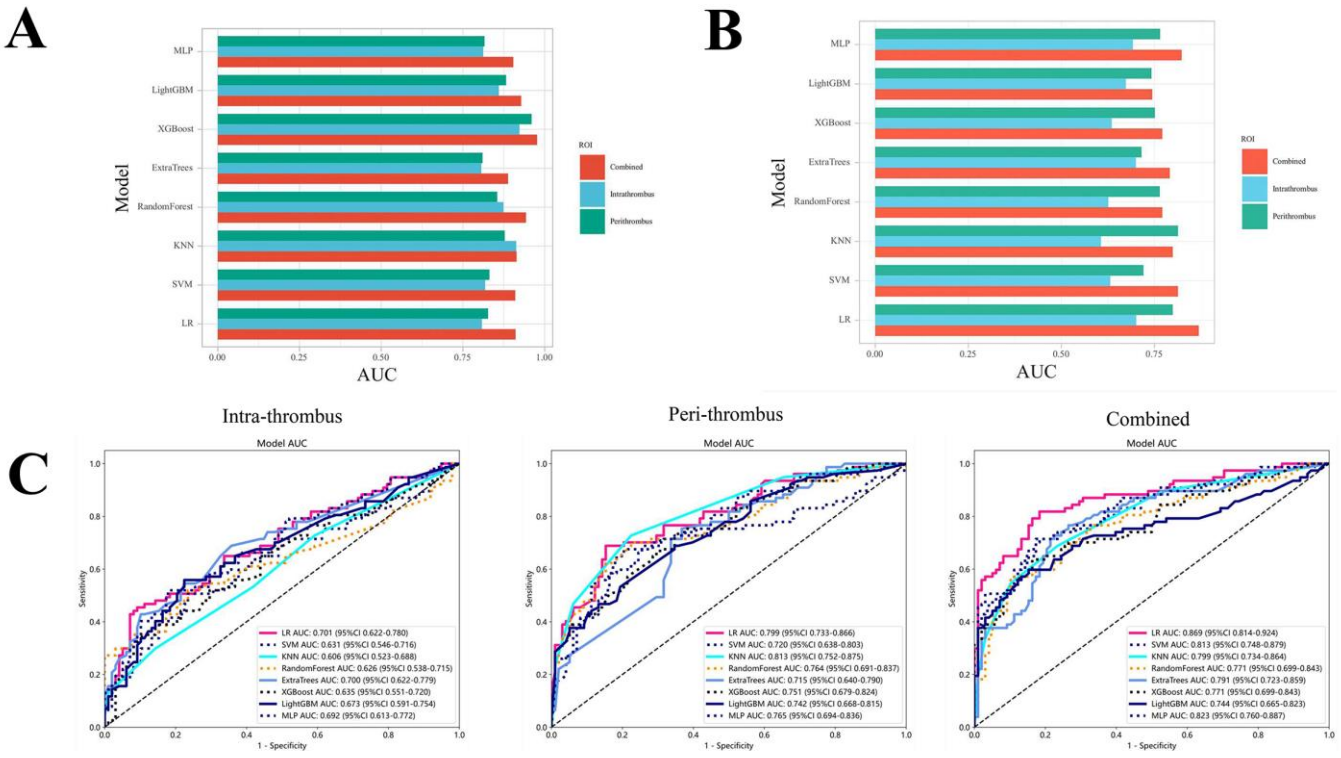


FIG 4. In the training(A) and validation(B) cohorts, the eight classifiers including the LR, SVM, KNN, RF, extra trees, XGBoost, Light GBM, MLP obtained AUCs for the models, respectively. In the validation cohort, the AUC range of each model based on intra-thrombus regions is 0.61-0.70, based on peri-thrombus regions is 0.74-0.87 and combined regions is 0.74-0.90 (C).

images.

Performance and Comparison of Models

The performance of eight classifiers—LR, SVM, KNN, RF, Extra Trees, XGBoost, Light GBM, and MLP—was assessed in both the training and validation cohorts. The results are detailed in Online Supplemental Data and Figure 4.

In the evaluation of models for the intra-thrombus, peri-thrombus, and combined regions, XGBoost consistently outperformed other algorithms in the training cohort. However, it experienced a noticeable decline in performance when evaluated on the validation cohort. KNN, LightGBM, and RF showed similar performance to XGBoost. LR performed slightly lower in the training cohort compared to these models and was relatively stable, but excelled in the validation cohort.

Within the intra-thrombus models of the validation cohort, the LR achieved the highest AUC of 0.70 (95% CI: 0.62-0.78), showing statistically significant differences compared to XGBoost (p=0.03), RF (p=0.03), and KNN (p=0.01), exhibiting no statistically significant disparities when compared to LightGBM (p=0.26), Extra Trees (p=0.97), MLP (p=0.36), and SVM (p=0.13). For the peri-thrombus models, the LR exhibited the greatest AUC of 0.80 (95% CI: 0.73-0.87), demonstrating statistically significant distinctions when compared to Extra Trees (p=0.002), LightGBM (p=0.04), MLP (p=0.03), and SVM (p=0.01), whereas it did not differ significantly from RF (p=0.18),XGBoost (p=0.05), and KNN (p=0.60). In the combined regions, the LR model reached an AUC of 0.87 (95% CI: 0.81 - 0.92), and was statistically different compared to Extra Trees (p=0.01), KNN (p=0.02), RF (p=0.01), XGBoost (p<0.001), LightGBM (p<0.001), MLP (p=0.01), and SVM (p=0.01).

190Furthermore, in the validation cohort, the diagnostic capability of the LR model using the combined regions significantly surpassed that of both the intra-thrombus (p<0.001) and peri-thrombus models (p=0.01). However, no significant difference in diagnostic performance was observed between the intra-thrombus and peri-thrombus models (p=0.05).

We integrated clinical parameters into a LR radiomics model but observed no significant predictive gains in the validation cohort (P > 0.05). An exclusively clinical LR model has also been formulated. The receiver operating characteristic curves for all three models can be examined in the Supplementary Data available online.

Table 1: Baseline demographic characteristics and clinical variables of enrolled patients

Variables	Training cohort (n=161)			Validation cohort (n=175)			P value
	Good outcome (n=110) mRS(0-2)	Poor outcome (n=51) mRS(3-6)	P value	Good outcome (n=98) mRS(0-2)	Poor outcome (n=77) mRS(3-6)	P value	
Age,years	65.7±11.65	63.76±13.51	0.45	66.76±12.45	66.96±13.28	0.83	0.20
Male, n (%)	69(62.73)	34(66.67)	0.76	68(69.39)	48(62.34)	0.41	0.66
History, n (%)							
Hypertension	90(81.82)	41(80.39)	1.00	71(72.45)	54(70.13)	0.87	0.05
Hyperlipidemia	36(32.73)	16(31.37)	1.00	23(23.47)	25(32.47)	0.25	0.33
Diabetes	34(30.91)	19(37.25)	0.54	25(25.51)	24(31.17)	0.51	0.33
Smoking	21(19.09)	15(29.41)	0.21	23(23.47)	16(20.78)	0.81	0.99
Atrial fibrillation	74(67.27)	24(47.06)	0.02	69(70.41)	42(54.55)	0.05	0.63
Coronary heart disease	22(20.00)	7(13.73)	0.46	16(16.33)	14(18.18)	0.90	0.83
NIHSS	14.26±8.45	14.51±8.24	0.92	14.32±7.57	15.27±6.90	0.19	0.64

DISCUSSION

In this retrospective analysis, our aim was to use radiomic features derived from intra-thrombus and peri-thrombus regions on CT scans to forecast prognosis following EVT in patients with acute ischemic stroke. A significant gap exists in the current research regarding the use of thrombus-related radiomics for predicting clinical outcomes following EVT, especially in peri-thrombotic areas, as some studies focus on recanalization following EVT[13, 17]. Our study developed and validated a radiomics model that utilizes features from both intra-thrombus and peri-thrombus regions to estimate prognosis after EVT. We employed eight different classifiers to determine which models offer robust diagnostic effectiveness and superior generalization capabilities. Among these, the LR model

using combined radiomics features proved to be the most precise in predicting outcomes.

Prior research has shown that the duration of thrombectomy and the frequency of interventions can influence the long-term outcomes in patients[19, 20]. There is also evidence suggesting a relationship between the thrombus's structural composition and the number and duration of EVT procedures[21, 22]. Variables like the use of stent retrievers, thrombotic makeup, and the count of thrombectomy sessions might lead to varying extents of vascular trauma in patients with acute ischemic stroke[23, 24], indicating that thrombus composition could be a critical factor in forecasting prognosis post-EVT. In this study, we constructed a combined model using 15 selected radiomic features derived solely from CTA scans, which included 9 features from intra-thrombus regions and 6 from peri-thrombus regions. Our analysis reveals that radiomic features from both intra- and peri-thrombus regions have potent predictive capabilities, particularly those from peri-thrombus areas. In the validation cohort, radiomics from the peri-thrombus areas notably enhanced the prediction of prognosis post-EVT over features from intra-thrombus regions alone. Based on prior research[25, 26] and actual measurements of vessel wall thickness in high-resolution MR images, we defined the peri-thrombus region as extending 1mm outward from the thrombus boundary. The peri-thrombotic region, which includes structures like the vessel wall and perivascular fat, may offer predictive insights into the disruption of the BBB associated with a heightened risk of complications, aligning with previous findings[27, 28]. Given that BBB disruption is commonly observed post-EVT and is linked to an increased risk of complications[27, 29], radiomic features from peri-thrombus regions could serve as crucial predictors of clinical outcomes post-EVT. Utilizing information from both regions significantly improves diagnostic performance, surpassing that of models using only intra- or peri-thrombus data, underscoring the value of integrating data from both regions for enhanced prediction accuracy.

The selection of classifiers is pivotal to the effectiveness of predictive models, yet there remains no universal standard guiding this choice, leading researchers to rely on personal preference and experience[30]. Consequently, this study assessed and compared eight different machine learning classifiers. The findings indicated that the LR model was consistently more effective than others in both training and validation cohorts, particularly when analyzing combined radiomic features, where it achieved the highest AUC across all classifiers. LR was favored over complex models for its statistical simplicity, interpretability, and robust performance for binary classification tasks associated with lesser risks of overfitting[15, 31, 32]. Given the scope and nature of our data, we believed that a parsimonious model like LR would be more appropriate. Complex models like SVM, MLP, RF, Extra Trees, XGBoost, KNN, and LightGBM, despite their high-dimensional data handling and robustness, were prone to overfitting without substantial data and careful tuning. The notable decline in XGBoost's validation performance highlighted the overfitting risk and the need for a balance between model complexity and generalizability. LR's consistent validation cohort performance affirmed its suitability and reliability for clinical diagnostics, justifying its choice due to effective generalization as demonstrated across datasets.

In this study, except for atrial fibrillation, no clinical variables exhibited statistically significant differences between the good outcome and poor outcome groups in either the training or validation cohorts. Notably, a significant difference in the occurrence of atrial fibrillation between the good outcome and poor outcome groups was identified, which stands in contrast to other studies that found no significant differences[31, 33, 34]. This inconsistency can be explained by variations in data arrangement and the sample sizes involved in our research. To further understand these discrepancies, additional research with larger sample sizes is recommended. The combined model did not demonstrably enhance predictive performance over the radiomics-only model, suggesting a greater reliance on superior-performing radiomics features rather than clinical variables.

This study has several notable limitations. Firstly, the patient sample size is relatively small, which may impact the stability of the machine learning model outcomes; applying these models to larger datasets could potentially provide more robust results. Secondly, the thrombus segmentation process was manually conducted, which could be time-consuming and might compromise the reliability of the results. Future research should focus on developing automated or semi-automated methods for more efficient and accurate thrombus segmentation. Thirdly, we sourced training and validation cohorts from separate centers, and despite image calibration and cross-validation efforts, potential bias may exist. Lastly, even with cross-validation and regularization techniques, overfitting is a challenge in high-dimensional data like ours. Future studies need to encompass more centers and larger samples to validate our findings with external data.

CONCLUSIONS

In summary, we developed and validated a CT imaging-based radiomics model to evaluate the prognosis following EVT in patients with acute ischemic stroke. This model could provide critical insights for clinical decision-making and outcome prediction. The analysis showed varied performance across different thrombus regions and classifiers, with models that combined features from multiple regions proving most effective. Specifically, the LR models exhibited high efficacy and stability in predicting clinical outcome.

ACKNOWLEDGMENTS

This study was supported by Jiangsu Province Capability Improvement Project through Science，Technology and Education （Jiangsu Provincial Medical Key Discipline Cultivation Unit，JSDW202242）and National Natural Science Foundation of China (Youth Program, Grant No. 82402364).

REFERENCES

1.Global, regional, and national burden of stroke, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet Neurology* 2019;18:439-58. https://doi.org/10.1016/s1474-4422(19)30034-1

2. Sarraj A, Hassan AE, Abraham MG, et al. Trial of Endovascular Thrombectomy for Large Ischemic Strokes. *The New England journal of medicine* 2023;388:1259-71. https://doi.org/10.1056/NEJMoa2214403

3.Yoshimura S, Sakai N, Yamagami H, et al. Endovascular Therapy for Acute Stroke with a Large Ischemic Region. *The New England journal of medicine* 2022;386:1303-13. https://doi.org/10.1056/NEJMoa2118191

4.Zhang X, Xie Y, Wang H, et al. Symptomatic Intracranial Hemorrhage After Mechanical Thrombectomy in Chinese Ischemic Stroke Patients: The ASIAN Score. *Stroke* 2020;51:2690-96. https://doi.org/10.1161/strokeaha.120.030173

5.Zi W, Qiu Z, Li F, et al. Effect of Endovascular Treatment Alone vs Intravenous Alteplase Plus Endovascular Treatment on Functional Independence in Patients With Acute Ischemic Stroke: The DEVT Randomized Clinical Trial. *Jama* 2021;325:234-43. https://doi.org/10.1001/jama.2020.23523

6.Harker P, Aziz YN, Vranic J, et al. Asymptomatic Intracerebral Hemorrhage Following Endovascular Stroke Therapy Is Not Benign: A Systematic Review and Meta-Analysis. *Journal of the American Heart Association* 2024;13:e031749. https://doi.org/10.1161/jaha.123.031749

7.Huang X, Yang Q, Shi X, et al. Predictors of malignant brain edema after mechanical thrombectomy for acute ischemic stroke. *Journal of neurointerventional surgery* 2019;11:994-98. https://doi.org/10.1136/neurintsurg-2018-014650

8.Thorén M, Dixit A, Escudero-Martínez I, et al. Effect of Recanalization on Cerebral Edema in Ischemic Stroke Treated With Thrombolysis and/or Endovascular Therapy. *Stroke* 2020;51:216-23. https://doi.org/10.1161/strokeaha.119.026692

9.Merlino G, Tereshko Y, Pez S, et al. Hyperdense middle cerebral artery sign predicts favorable outcome in patients undergoing mechanical thrombectomy. *Journal of thrombosis and thrombolysis* 2023;55:312-21. https://doi.org/10.1007/s11239-022-02731-4

10.Lu ZJ, Lai JX, Huang JR, et al. Predictive value of intracranial high-density areas in neurological function. *World journal of psychiatry* 2024;14:1080-86. https://doi.org/10.5498/wjp.v14.i7.1080

11.Beyeler M, Pohle F, Weber L, et al. Long-Term Effect of Mechanical Thrombectomy in Stroke Patients According to Advanced Imaging Characteristics. *Clinical neuroradiology* 2024;34:105-14. https://doi.org/10.1007/s00062-023-01337-4

12.Olivot JM, Albucher JF, Guenego A, et al. Mismatch Profile Influences Outcome After Mechanical Thrombectomy. *Stroke* 2021;52:232-40. https://doi.org/10.1161/strokeaha.120.031929

13.Xiong X, Wang J, Ke J, et al. Radiomics-based intracranial thrombus features on preoperative noncontrast CT predicts successful recanalization of mechanical thrombectomy in acute ischemic stroke. *Quantitative imaging in medicine and surgery* 2023;13:682-94. https://doi.org/10.21037/qims-22-599

14.Wang C, Li T, Jia Z, et al. Radiomics features on computed tomography reflect thrombus histological age prior to endovascular treatment of acute ischemic stroke. *Journal of stroke and cerebrovascular diseases : the official journal of National Stroke Association* 2023;32:107358. https://doi.org/10.1016/j.jstrokecerebrovasdis.2023.107358

15.Jiang J, Wei J, Zhu Y, et al. Clot-based radiomics model for cardioembolic stroke prediction with CT imaging before recanalization: a multicenter study. *European radiology* 2023;33:970-80. https://doi.org/10.1007/s00330-022-09116-4

16.Wang C, Hang Y, Cao Y, et al. A nomogram for predicting thrombus composition in stroke patients with large vessel occlusion: combination of thrombus density and perviousness with clinical features. *Neuroradiology* 2023;65:371-80. https://doi.org/10.1007/s00234-022-03046-0

17.Qiu W, Kuang H, Nair J, et al. Radiomics-Based Intracranial Thrombus Features on CT and CTA Predict Recanalization with Intravenous Alteplase in Patients with Acute Ischemic Stroke. *AJNR American journal of neuroradiology* 2019;40:39-44. https://doi.org/10.3174/ajnr.A5918

18.Cheng Y, Wan S, Wu W, et al. Computed Tomography Angiography-Based Thrombus Radiomics for Predicting the Time Since Stroke Onset. *Academic radiology* 2023;30:2469-76. https://doi.org/10.1016/j.acra.2022.12.032

19.Alawieh A, Vargas J, Fargen KM, et al. Impact of Procedure Time on Outcomes of Thrombectomy for Stroke. *Journal of the American College of Cardiology* 2019;73:879-90. https://doi.org/10.1016/j.jacc.2018.11.052

20.Bourcier R, Saleme S, Labreuche J, et al. More than three passes of stent retriever is an independent predictor of parenchymal hematoma in acute ischemic stroke. *Journal of neurointerventional surgery* 2019;11:625-29. https://doi.org/10.1136/neurintsurg-2018-014380

21.Abbasi M, Arturo Larco J, Mereuta MO, et al. Diverse thrombus composition in thrombectomy stroke patients with longer time to recanalization. *Thrombosis research* 2022;209:99-104. https://doi.org/10.1016/j.thromres.2021.11.018

22.Maekawa K, Shibata M, Nakajima H, et al. Erythrocyte-Rich Thrombus Is Associated with Reduced Number of Maneuvers and Procedure Time in Patients with Acute Ischemic Stroke Undergoing Mechanical Thrombectomy. *Cerebrovascular diseases extra* 2018;8:39-49. https://doi.org/10.1159/000486042

23.Funatsu N, Hayakawa M, Hashimoto T, et al. Vascular wall components in thrombi obtained by acute stroke thrombectomy: clinical significance and related factors. *Journal of neurointerventional surgery* 2019;11:232-36. https://doi.org/10.1136/neurintsurg-2018-014041

24.Koge J, Kato S, Hashimoto T, et al. Vessel Wall Injury After Stent Retriever Thrombectomy for Internal Carotid Artery Occlusion with Duplicated Middle Cerebral Artery. *World neurosurgery* 2019;123:54-58. https://doi.org/10.1016/j.wneu.2018.11.223

25.Obusez EC, Hui F, Hajj-Ali RA, et al. High-resolution MRI vessel wall imaging: spatial and temporal patterns of reversible cerebral vasoconstriction syndrome and central nervous system vasculitis. AJNR American journal of neuroradiology 2014;35:1527-32. https://doi.org/10.3174/ajnr.A3909

26.Zhu C, Haraldsson H, Tian B, et al. High resolution imaging of the intracranial vessel wall at 3 and 7 T using 3D fast spin echo MRI. Magma (New York, NY) 2016;29:559-70. https://doi.org/10.1007/s10334-016-0531-x

27.Hom J, Dankbaar JW, Soares BP, et al. Blood-brain barrier permeability assessed by perfusion CT predicts symptomatic hemorrhagic transformation and malignant edema in acute ischemic stroke. *AJNR American journal of neuroradiology* 2011;32:41-8. https://doi.org/10.3174/ajnr.A2244

28.Renú A, Laredo C, Lopez-Rueda A, et al. Vessel Wall Enhancement and Blood-Cerebrospinal Fluid Barrier Disruption After Mechanical Thrombectomy in Acute Ischemic Stroke. *Stroke* 2017;48:651-57. https://doi.org/10.1161/strokeaha.116.015648

29.Shi ZS, Duckwiler GR, Jahan R, et al. Early Blood-Brain Barrier Disruption after Mechanical Thrombectomy in Acute Ischemic Stroke. *Journal of neuroimaging : official journal of the American Society of Neuroimaging* 2018;28:283-88. https://doi.org/10.1111/jon.12504

30.Gu H, Zhang X, di Russo P, et al. The Current State of Radiomics for Meningiomas: Promises and Challenges. *Frontiers in oncology* 2020;10:567736. https://doi.org/10.3389/fonc.2020.567736

31.Xie G, Li T, Ren Y, et al. Radiomics-based infarct features on CT predict hemorrhagic transformation in patients with acute ischemic stroke. *Frontiers in neuroscience* 2022;16:1002717. https://doi.org/10.3389/fnins.2022.1002717

32.Bisaso KR, Karungi SA, Kiragga A, et al. A comparative study of logistic regression based machine learning techniques for prediction of early virological suppression in antiretroviral initiating HIV patients. *BMC medical informatics and decision making* 2018;18:77. https://doi.org/10.1186/s12911-018-0659-x

33.van Kranendonk KR, Treurniet KM, Boers AMM, et al. Clinical and Imaging Markers Associated With Hemorrhagic Transformation in Patients With Acute Ischemic Stroke. *Stroke* 2019;50:2037-43. https://doi.org/10.1161/strokeaha.118.024255
34.Xu J, Chen XY, Wang HY, et al. Hemodynamic predictors of early neurological deterioration and clinical outcome after endovascular treatment in large artery occlusion. *Heliyon* 2024;10:e24746. https://doi.org/10.1016/j.heliyon.2024.e24746

SUPPLEMENTAL FILES

Online Supplemental Data:

Online Supplemental Data: Performance of eight classifiers in training and validation cohort

Classifiers	Groups	Accuracy	AUC	95% CI	Sensitivity	Specificity	PPV	NPV	F1 Score
Intra-thrombus									
LR	Training	0.78	0.81	0.73- 0.88	0.69	0.82	0.64	0.85	0.66
	validation	0.70	0.70	0.62- 0.78	0.44	0.91	0.79	0.67	0.57
SVM	Training	0.81	0.82	0.73- 0.90	0.73	0.86	0.70	0.87	0.71
	validation	0.65	0.63	0.55- 0.72	0.49	0.77	0.62	0.66	0.55
KNN	Training	0.86	0.91	0.87- 0.96	0.65	0.96	0.89	0.86	0.75
	validation	0.61	0.61	0.52- 0.69	0.13	0.99	0.91	0.59	0.23
RF	Training	0.88	0.87	0.81- 0.94	0.71	0.96	0.88	0.88	0.78
	validation	0.65	0.63	0.54- 0.72	0.51	0.77	0.63	0.66	0.56
ExtraTrees	Training	0.75	0.81	0.73- 0.88	0.73	0.76	0.59	0.86	0.65
	validation	0.66	0.70	0.62- 0.78	0.65	0.67	0.61	0.71	0.63
XGBoost	Training	0.91	0.92	0.87- 0.98	0.78	0.97	0.93	0.91	0.85
	validation	0.65	0.64	0.55- 0.72	0.31	0.92	0.75	0.63	0.44
LightGBM	Training	0.80	0.86	0.80- 0.92	0.77	0.82	0.66	0.88	0.71
	validation	0.67	0.67	0.59- 0.75	0.55	0.78	0.66	0.69	0.60
MLP	Training	0.78	0.81	0.74- 0.89	0.71	0.82	0.64	0.86	0.67
	validation	0.68	0.69	0.61- 0.77	0.51	0.82	0.68	0.68	0.58
Peri-thrombus									
LR	Training	0.73	0.83	0.76 - 0.90	0.75	0.73	0.56	0.86	0.64
	validation	0.77	0.80	0.73 - 0.87	0.68	0.85	0.78	0.77	0.72
SVM	Training	0.75	0.83	0.76 - 0.91	0.84	0.70	0.57	0.91	0.68
	validation	0.74	0.72	0.64 - 0.80	0.66	0.80	0.72	0.75	0.69
KNN	Training	0.83	0.88	0.83 - 0.93	0.55	0.96	0.88	0.82	0.68
	validation	0.73	0.81	0.75 - 0.88	0.47	0.94	0.86	0.69	0.61
RF	Training	0.73	0.86	0.80 - 0.92	0.82	0.69	0.55	0.89	0.66
	validation	0.74	0.76	0.69 - 0.84	0.69	0.78	0.71	0.76	0.70
ExtraTrees	Training	0.68	0.81	0.74 - 0.88	0.82	0.61	0.49	0.88	0.62
	validation	0.68	0.72	0.64 - 0.79	0.74	0.63	0.61	0.76	0.67
XGBoost	Training	0.90	0.96	0.93 - 0.99	0.94	0.88	0.79	0.97	0.86

LightGBM	validation	0.68	0.75	0.68 - 0.82	0.71	0.65	0.62	0.74	0.66
MLP	Training	0.78	0.88	0.83 - 0.93	0.71	0.82	0.64	0.86	0.67
	validation	0.66	0.74	0.67 - 0.82	0.68	0.65	0.61	0.72	0.64
	Training	0.67	0.82	0.75 - 0.88	0.90	0.56	0.49	0.93	0.63
	validation	0.72	0.77	0.69 - 0.84	0.57	0.84	0.73	0.71	0.64
Combined									
LR	Training	0.85	0.91	0.86 - 0.96	0.84	0.86	0.73	0.92	0.78
	validation	0.81	0.87	0.81 - 0.92	0.81	0.82	0.78	0.84	0.79
SVM	Training	0.91	0.91	0.85 - 0.98	0.82	0.96	0.89	0.92	0.86
	validation	0.77	0.81	0.75 - 0.88	0.70	0.82	0.75	0.78	0.73
KNN	Training	0.81	0.91	0.87 - 0.96	0.78	0.82	0.67	0.89	0.72
	validation	0.74	0.80	0.73 - 0.87	0.56	0.90	0.81	0.72	0.65
RF	Training	0.87	0.94	0.91 - 0.98	0.86	0.87	0.76	0.93	0.81
	validation	0.74	0.77	0.70 - 0.84	0.56	0.89	0.80	0.72	0.66
ExtraTrees	Training	0.87	0.89	0.83 - 0.95	0.69	0.96	0.88	0.87	0.77
	validation	0.75	0.79	0.72 - 0.86	0.71	0.78	0.71	0.78	0.71
XGBoost	Training	0.95	0.98	0.95 - 1.00	0.90	0.97	0.94	0.96	0.92
	validation	0.74	0.77	0.70 - 0.84	0.62	0.84	0.75	0.74	0.68
LightGBM	Training	0.85	0.93	0.88 - 0.97	0.86	0.85	0.72	0.93	0.79
	validation	0.73	0.74	0.67 - 0.82	0.58	0.85	0.75	0.72	0.66
MLP	Training	0.87	0.90	0.85 - 0.96	0.80	0.90	0.79	0.91	0.80
	validation	0.77	0.82	0.76 - 0.89	0.70	0.83	0.76	0.78	0.73

AUC, area under curve; PPV, positive predictive value; NPV, negative predictive value; RF, random forest; LR, logistic regression, SVM, support vector machine, KNN, k nearest neighbors; Light GBM, light gradient boosting machine; MLP, multi-layer perceptron; XGBoost, extreme gradient boosting.

Online Supplemental Table 1: The detailed description of the selected radiomics features

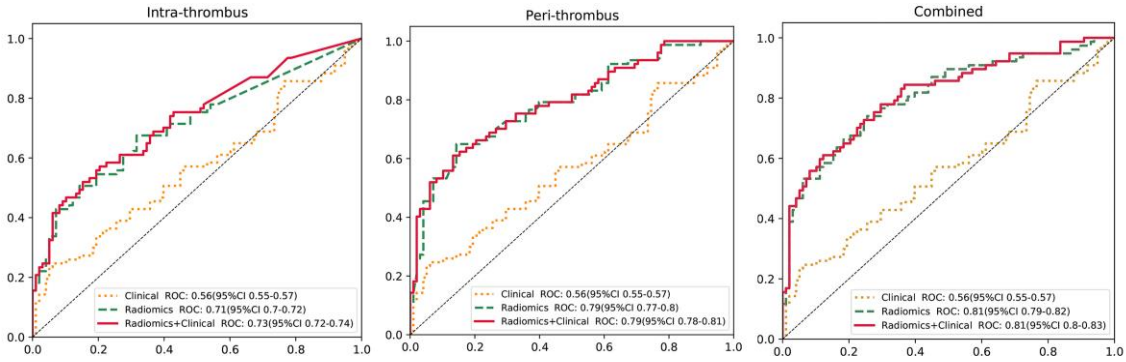
Radiomics features classification	Radiomics features	Description of radiomics characteristics
Intra-thrombus Model		
firstorder	TotalEnergy	Represents the sum of squared voxel intensities throughout the image volume. Provides insight into the total magnitude of voxel values, which could reflect the mass and potential energy of the thrombus.
	Median	A measure of central tendency, it is the middle value of voxel intensities when arranged in ascending order. Ideal for understanding the typical attenuation value, the median can relate to the composition of the thrombus.
Neighbouring Gray Tone Difference Matrix(ngtdm)	Contrast	"Contrast" measures the local variations in image intensity. High contrast may indicate heterogeneity in the thrombus, potentially associated with clot composition.

	Busyness	Evaluates the rate of change in intensity values between a voxel and its neighbours. May indicate the complexity and the degree of structural change within the thrombus.
Gray Level Size Zone Matrix (glszm)	GrayLevelNonUniformity	Measures the non-uniformity of gray-level intensity values, which reflects zones with similar gray level values. A higher non-uniformity suggests varying densities within the thrombus, potentially corresponding to different components like fibrin or red blood cells.
shape	Flatness	Quantifies the flatness of the thrombus shape, calculated from the eigenvalues of a shape-based matrix. Helps identify irregular thrombus geometries.
Peri-thrombus		
firstorder	Minimum	The lowest intensity value within the region of interest. Can indicate the presence of different regions within the peri-thrombus.
	TotalEnergy	The sum of squared voxel intensities, indicating the overall magnitude of voxel values. May reflect on the peri-thrombus density and structure.
	Kurtosis	A statistical measure that describes the distribution shape concerning the tails. Higher kurtosis indicates a more outlier-prone distribution. Provides insights into tissue heterogeneity.
	Skewness	Reflects asymmetry in the distribution of voxel intensity values. Helps predict asymmetrical or irregular attributes.
Gray Level Run Length Matrix(glrlm)	RunEntropy	Measures the randomness or complexity in the distribution of runs of pixels. A higher value could indicate a complex structure.
	LongRunHighGrayLevelEmphasis	Gives preference to the runs of high-intensity values that are longer. May suggest areas with dense calcifications within the peri-thrombus.
	RunLengthNonUniformity	Quantifies the non-uniformity of run lengths, assessing the texture of the structures. Uneven run lengths might point to peri-thrombus heterogeneity.
Gray Level Co-occurrence Matrix(glcm)	Contrast	Measures pixel intensity contrast between pixel pairs across the image, reflecting texture roughness. Helpful in understanding the level of density.
	Imc1	A correlate of the homogeneity of textures across an image. Lower values may correlate with more uniform structures.
Gray Level Dependence Matrix(gldm)	LargeDependenceHighGrayLevelEmphasis	Emphasizes larger dependencies with higher gray-level values. Can suggest peri-thrombus with sections of dense matter.
ngtdm	Strength	Evaluates the strength of patterns or textures within the image. Strong patterns could indicate a consistent structural makeup that impacts treatment decisions.
glszm	ZoneVariance	Measures the distribution of zone size and gray level variations. Variance in zone sizes can reflect the presence of different structure characteristics.
Combined		
Intra-thrombus	Median	The same as above.
firstorder	TotalEnergy	The same as above.
	Skewness	"Skewness" reflect the asymmetry of the intensity histogram. It might provide insights into the physical characteristics of the thrombus.
ngtdm	Contrast	The same as above.
gldm	LargeDependenceHighGrayLevelEmphasis	This feature measures the joint distribution of large dependence with higher gray-level values. Clinically, it could be related to more dense or calcified areas within the thrombus.
	GrayLevelNonUniformity	The same as above.
glszm	LargeAreaLowGrayLevelEmphasis	Highlights the presence of extensive, low-intensity areas, potentially indicating a softer thrombus composition.

		ZoneVariance	These attributes assess the variability and uniformity of gray levels. Variations in gray level might suggest different textures within the thrombus.
Peri-thrombus	shape	Flatness	The same as above.
	firstorder	Minimum	The same as above.
		Skewness	The same as above.
	glszm	ZoneVariance	The same as above.
	glcm	DifferenceVariance	Can assess the variation in intensity differences, suggesting textural complexity.
	glrlm	LongRunHighGrayLevelEmphasis	The same as above.
		RunEntropy	The same as above.

Online Supplemental Table 2: Performance of clinical variables in validation cohort

Models	Accuracy	AUC	95% CI	Sensitivity	Specificity	PPV	NPV	F1 Score	P value
Intra-thrombus models									
Clinical	0.60	0.56	0.55-0.57	0.10	0.99	0.89	0.58	0.19	
Radiomics	0.67	0.71	0.70-0.72	0.51	0.81	0.67	0.68	0.58	0.18
Radiomics+ Clinical	0.67	0.73	0.80-0.83	0.32	0.94	0.81	0.64	0.46	
Peri-thrombus models									
Clinical	0.60	0.56	0.55-0.57	0.10	0.99	0.89	0.58	0.19	
Radiomics	0.76	0.79	0.77-0.80	0.65	0.85	0.77	0.75	0.70	0.82
Radiomics+ Clinical	0.69	0.79	0.78-0.81	0.32	0.98	0.93	0.65	0.48	
Combined models									
Clinical	0.60	0.56	0.55-0.57	0.10	0.99	0.89	0.58	0.19	
Radiomics	0.73	0.81	0.79-0.82	0.68	0.77	0.69	0.75	0.68	0.49
Radiomics+ Clinical	0.74	0.81	0.72-0.74	0.45	0.96	0.90	0.69	0.60	



Online Supplemental Fig 1: The receiver operating characteristic curves for all three clinical models.

STARD 2015

Section & Topic	Item No	Item	Reported on Page Number/ Line Number	Reported on Section/ Paragraph
TITLE OR ABSTRACT				
	1	Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC)	1/17,19	ABSTRACT/2
ABSTRACT				
	2	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)	1/10-23	ABSTRACT/1-4
INTRODUCTION				
	3	Scientific and clinical background, including the intended use and clinical role of the index test	2/54-74	INTRODUCTION/1-3
	4	Study objectives and hypotheses	2/75-77	INTRODUCTION/4
METHODS				
Study design	5	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)	2/82-83	METHODS/1
Participants	6	Eligibility criteria	2/83-86	METHODS/1
	7	On what basis potentially eligible participants were identified (such as symptoms, results from previous tests, inclusion in registry)	2/83-86	METHODS/1
	8	Where and when potentially eligible participants were identified (setting, location and dates)	2/82-83	METHODS/1
	9	Whether participants formed a consecutive, random or convenience series	2/90-93	METHODS/1
Test methods	10a	Index test, insufficient detail to allow replication	2/96-117	METHODS/2-3
	10b	Reference standard, insufficient detail to allow replication	4/131-132	METHODS/4
	11	Rationale for choosing the reference standard (if alternatives exist)	none	none
	12a	Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory	3-4/119-132	METHODS/4-5
	12b	Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory	3-4/119-132	METHODS/4-5
	13a	Whether clinical information and reference standard results were available to the performers/readers of the index test	2/103-106	METHODS/2
	13b	Whether clinical information and index test results were available to the assessors of the reference standard	2/87-91	METHODS/1

Analysis	14	Methods for estimating or comparing measures of diagnostic accuracy	4/136-141	METHODS/5
	15	How indeterminate index test or reference standard results were handled	none	none
	16	How missing data on the index test and reference standard were handled	none	none
	17	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory	3-4/121-132	METHODS/4-5
	18	Intended sample size and how it was determined	2/81-85	METHODS/1
RESULTS				
Participants	19	Flow of participants, using a diagram	2/92-93	METHODS/1
	20	Baseline demographic and clinical characteristics of participants	4/154-159	RESULTS/1
	21a	Distribution of severity of disease in those with the target condition	4/147-152	RESULTS/1
	21b	Distribution of alternative diagnoses in those without the target condition	4/147-152	RESULTS/1
	22	Time interval and any clinical interventions between index test and reference standard	2/82-88	METHODS/1
Test results	23	Cross tabulation of the index test results (or their distribution) by the results of the reference standard	4-6/155-194	RESULTS/2-7
	24	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)	4-6/155-194	RESULTS/2-7
	25	Any adverse events from performing the index test or the reference standard	none	none
DISCUSSION				
	26	Study limitations, including sources of potential bias, statistical uncertainty, and generalisability	7/239-246	DISCUSSION/5
	27	Implications for practice, including the intended use and clinical role of the index test	9/250-253	Conclusion/1
OTHER INFORMATION				
	28	Registration number and name of registry	2/80-81	METHODS/1
	29	Where the full study protocol can be accessed	2/93-94	METHODS/1
	30	Sources of funding and other support; role offunders	9/256-258	Acknowledgments/1

AIM

STARD stands for “Standards for Reporting Diagnostic accuracy studies”. This list of items was developed to contribute to the completeness and transparency of reporting of diagnostic accuracy studies. Authors can use the list to write informative study reports. Editors and peer-reviewers can use it to evaluate whether the information has been included in manuscripts submitted for publication.

Explanation

A **diagnostic accuracy study** evaluates the ability of one or more medical tests to correctly classify study participants as having a **target condition**. This can be a disease, a disease stage, response or benefit from therapy, or an event or condition in the future. A medical test can be an imaging procedure, a laboratory test, elements from history and physical examination, a combination of these, or any other method for collecting information about the current health status of a patient.

The test whose accuracy is evaluated is called **index test**. A study can evaluate the accuracy of one or more index tests. Evaluating the ability of a medical test to correctly classify patients is typically done by comparing the distribution of the index test results with those of the **reference standard**. The reference standard is the best available method for establishing the presence or absence of the target condition. An accuracy study can rely on one or more reference standards.

If test results are categorized as either positive or negative, the cross tabulation of the index test results against those of the reference standard can be used to estimate the **sensitivity** of the index test (the proportion of participants with the target condition who have a positive index test), and its **specificity** (the proportion without the target condition who have a negative index test). From this cross tabulation (sometimes referred to as the contingency or “2x2” table), several other accuracy statistics can be estimated, such as the positive and negative **predictive values** of the test. Confidence intervals around estimates of accuracy can then be calculated to quantify the statistical **precision** of the measurements.

If the index test results can take more than two values, categorization of test results as positive or negative requires a **test positivity cut-off**. When multiple such cut-offs can be defined, authors can report a receiver operating characteristic (ROC) curve which graphically represents the combination of sensitivity and specificity for each possible test positivity cut-off. The **area under the ROC curve** informs in a single numerical value about the overall diagnostic accuracy of the index test.

The **intended use** of a medical test can be diagnosis, screening, staging, monitoring, surveillance, prediction or prognosis. The **clinical role** of a test explains its position relative to existing tests in the clinical pathway. A replacement test, for example, replaces an existing test. A triage testis used before an existing test; an add-on testis used after an existing test.

Besides diagnostic accuracy, several other outcomes and statistics may be relevant in the evaluation of medical tests. Medical tests can also be used to classify patients for purposes other than diagnosis, such as staging or prognosis. The STARD list was not explicitly developed for these other outcomes, statistics, and study types, although most STARD items would still apply.

DEVELOPMENT

This STARD list was released in 2015. The 30 items were identified by an international expert group of methodologists, researchers, and editors. The guiding principle in the development of STARD was to select items that, when reported, would help readers to judge the potential for bias in the study, to appraise the applicability of the study findings and the validity of conclusions and recommendations. The list represents an update of the first version, which was published in 2003. More information can be found on http://www.equator-network.org/reporting-guidelines/stard.

Please leave this space alone as it will be supplemented by the editorial office when needed.