International Journal of Medicine
pISSN 1468-3814 | eISSN 2667-7008
https://theinternationalmedicine.com/

Original Researcher Article DOI: 10.61336/im/26-06-218

MACHINE LEARNING-BASED MRI ANALYSIS FOR PREDICTION OF ACUTE ISCHEMIC STROKE OUTCOMES: A HOSPITAL-BASED OBSERVATIONAL STUDY.

Dr. Shubhankar[1], Dr. Vinod Parmar[2], Dr. Mayank Kumar Ujjaliya[3*].
[1]*Assistant Professor, Department of Radiodiagnosis, Ram Krishna Medical College Hospital and Research Centre, Bhopal, Madhya Pradesh India.*
[2]*Assistant Professor, Department of Radiodiagnosis, Ram Krishna Medical College Hospital and Research Centre, Bhopal, Madhya Pradesh India.*
[3]*Assistant Professor, Department of Radiodiagnosis, Ram Krishna Medical College Hospital and Research Centre, Bhopal, Madhya Pradesh India.*

Correspondence:
Dr. Mayank Kumar Ujjaliya.
Assistant Professor, Department of Radiodiagnosis, Ram Krishna Medical College Hospital and Research Centre, Bhopal, Madhya Pradesh India.

Email:
drmayankujjaliya@gmail.com

Received: 06-05-2026
Revised: 18-05-2026
Accepted: 10-06-2026
Published: 22-06-2026

Abstract
Background: Acute ischemic stroke (AIS) remains a leading cause of mortality and long-term disability worldwide. Early and accurate prediction of functional outcomes is essential for optimizing treatment strategies and rehabilitation planning. Machine learning (ML)-based analysis of magnetic resonance imaging (MRI) has emerged as a promising approach for improving prognostic accuracy by integrating quantitative imaging biomarkers with clinical variables. ***Aim:*** To evaluate the performance of machine learning–based MRI analysis in predicting 90-day functional outcomes in patients with acute ischemic stroke. ***Materials and Methods:*** A prospective hospital-based study was conducted among 150 consecutive adult patients with MRI-confirmed acute ischemic stroke. Clinical characteristics, vascular risk factors, and MRI-derived imaging features were collected. Supervised machine learning models—Logistic Regression, Support Vector Machine (SVM), Random Forest, and Extreme Gradient Boosting (XGBoost)—were developed using an 80:20 training-testing split with 5-fold cross-validation. The primary outcome was 90-day functional status assessed using the modified Rankin Scale (mRS), with favorable outcome defined as mRS 0–2. Model performance was evaluated using accuracy, sensitivity, specificity, F1-score, and area under the receiver operating characteristic curve (AUC). ***Results:*** Among the 150 patients, 62.7% achieved a favorable functional outcome, while 37.3% had an unfavorable outcome at 90 days. Older age, diabetes mellitus, higher baseline NIHSS score, larger infarct volume, large vessel occlusion, and perfusion deficit were independent predictors of poor outcome, whereas intravenous thrombolysis was associated with improved recovery. Among the evaluated algorithms, XGBoost demonstrated the best predictive performance with an accuracy of 94%, sensitivity of 93%, specificity of 95%, F1-score of 0.93, and AUC of 0.97, outperforming Random Forest (AUC 0.95), SVM (AUC 0.91), and Logistic Regression (AUC 0.88). Baseline NIHSS score, infarct volume, ADC value, and large vessel occlusion were the most influential predictors in the final model. ***Conclusion:*** Machine learning–based MRI analysis provides highly accurate prediction of functional outcomes following acute ischemic stroke. Integrating quantitative MRI biomarkers with clinical parameters enhances early risk stratification and may facilitate personalized therapeutic decision-making, supporting the future incorporation of artificial intelligence into routine stroke care.

Keywords: Acute ischemic stroke; Magnetic resonance imaging; Machine learning; XGBoost; Functional outcome; NIHSS; Prognostic prediction.

INTRODUCTION
Acute ischemic stroke is a leading cause of mortality and long-term disability worldwide, with ischemic stroke accounting for the major proportion of global stroke burden. Recent estimates show that stroke burden remains particularly high in low- and middle-income countries, where delayed presentation, limited imaging access, and variable rehabilitation services contribute to poorer outcomes [1].

Early prediction of clinical outcome after acute ischemic stroke is essential for triage, treatment selection, rehabilitation planning, and counselling of patients and families. Conventional prognostic assessment is usually based on clinical parameters such as age, vascular risk factors, baseline neurological deficit, infarct location, infarct volume, and treatment received. However, these variables may not fully capture the complex tissue-level and perfusion-related changes occurring during the acute phase [2].

Magnetic resonance imaging provides sensitive evaluation of acute ischemic injury through diffusion-weighted imaging, apparent diffusion coefficient mapping, fluid-attenuated inversion recovery, susceptibility-weighted imaging, angiographic sequences, and perfusion imaging. These MRI-derived features can help define infarct core, collateral status, hemorrhagic transformation risk, and tissue viability, all of which influence functional recovery [3].

Machine learning offers a data-driven approach for integrating high-dimensional MRI features with clinical variables to predict stroke outcomes more objectively. Recent studies have shown that machine learning and deep learning models using imaging and clinical data can predict infarct evolution and 90-day functional outcome with promising accuracy. Nevertheless, real-world validation remains limited, particularly in hospital-based settings where patient populations, imaging protocols, and treatment pathways vary [4, 5].

Therefore, the present hospital-based observational study was designed to evaluate the role of machine learning-based MRI analysis in predicting outcomes among patients with acute ischemic stroke. The study aims to determine whether MRI-derived imaging features, alone or combined with clinical variables, can improve early prediction of functional outcome and support individualized stroke management.

MATERIALS AND METHODS

Study Design and Setting: A prospective hospital-based observational study was conducted in the Department of Radiodiagnosis in collaboration with the Department of Neurology at a tertiary care teaching hospital over a period of 18 months from October 2024 to march 2026.

Study Population: 150 Consecutive adult patients presenting with clinically suspected acute ischemic stroke and confirmed on MRI were enrolled after obtaining written informed consent.

Inclusion Criteria

- Patients aged ≥18 years.
- Clinical diagnosis of acute ischemic stroke.
- MRI performed within 24 hours of symptom onset.
- Availability of complete clinical, imaging, and follow-up data.
- Provision of informed written consent.

Exclusion Criteria

- Intracerebral hemorrhage or stroke mimics.
- Previous disabling stroke (modified Rankin Scale >2).
- Brain tumors, central nervous system infections, or demyelinating disorders.
- Contraindications to MRI
- Poor-quality MRI images unsuitable for analysis.
- Patients lost to follow-up.

Clinical Assessment: Demographic characteristics, vascular risk factors (hypertension, diabetes mellitus, dyslipidemia, atrial fibrillation, smoking, and alcohol consumption), symptom onset time, treatment received (intravenous thrombolysis and/or mechanical thrombectomy), baseline neurological status using the National Institutes of Health Stroke Scale (NIHSS), and laboratory investigations were recorded using a standardized case record form.
MRI examinations were performed on a 1.5-T scanner (Siemens Magnetom Essenza) using a standardized acute stroke imaging protocol.

Machine Learning-Based MRI Analysis: MRI datasets were anonymized before analysis. Quantitative imaging features including infarct volume, lesion location, ADC values, perfusion parameters, vascular occlusion site, and radiomic texture features were extracted using automated image-processing software. These imaging variables, together with relevant clinical parameters, were used to develop supervised machine learning prediction models.

The dataset was randomly divided into training (80%) and testing (20%) subsets. Data preprocessing included normalization, missing-value imputation, feature selection using recursive feature elimination, and class balancing where required. Machine learning algorithms including Random Forest, Extreme Gradient Boosting (XGBoost), Support Vector Machine (SVM), and Logistic Regression were trained using 5-fold cross-validation. Model performance was compared, and the algorithm with the highest predictive accuracy was selected as the final model.

Outcome Assessment: The primary outcome was functional status at 90 days, assessed using the modified Rankin Scale (mRS). A favorable outcome was defined as mRS 0–2, while an unfavorable outcome was defined as mRS 3–6. Secondary outcomes included in-hospital mortality, hemorrhagic transformation, and infarct progression.

Statistical Analysis: Data were analyzed using IBM SPSS Statistics version 29.0. Continuous variables were expressed as mean ± standard deviation, while categorical variables were presented as frequencies and percentages. Group comparisons were performed using the independent t-test or Mann–Whitney U test for continuous variables and the Chi-square or Fisher's exact test for categorical variables.

Variables with $p < 0.10$ on univariate analysis were entered into multivariable logistic regression to identify independent predictors of unfavorable outcome. Machine learning model performance was evaluated using accuracy, sensitivity, specificity, precision, F1-score, receiver operating characteristic (ROC) curve, area under the ROC curve (AUC), and calibration analysis. A two-tailed p-value <0.05 was considered statistically significant.

RESULTS

A total of 150 patients with acute ischemic stroke were included. The mean age was 63.8 ± 11.4 years, with males comprising 62.7% of the study population. Hypertension and diabetes mellitus were the most common vascular risk factors. The median baseline NIHSS score was 11, while the mean time from symptom onset to MRI was 7.1 ± 3.5 hours. Intravenous thrombolysis and mechanical thrombectomy were performed in 31.3% and 16% of patients [Table: 1].

Table 1: Baseline demographic and clinical characteristics of patients with acute ischemic stroke (N = 150)

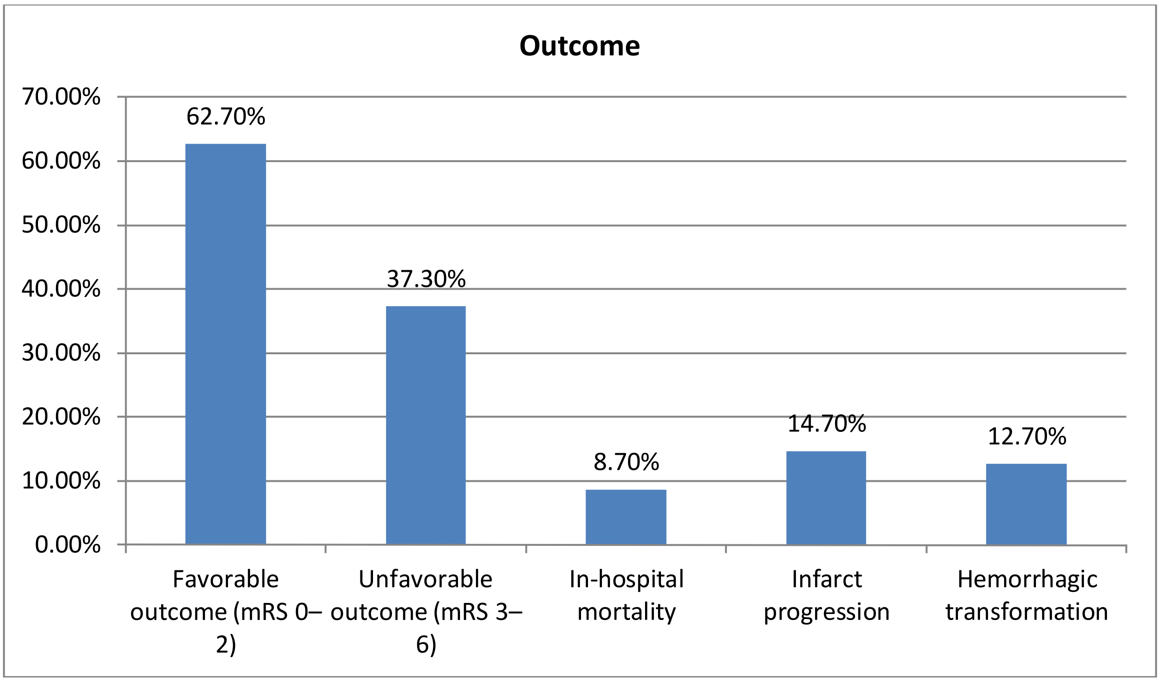
Variable	N (%)
Age (years), mean ± SD	63.8 ± 11.4
Male	94 (62.7)
Female	56 (37.3)
Hypertension	104 (69.3)
Diabetes mellitus	61 (40.7)
Dyslipidemia	58 (38.7)
Atrial fibrillation	29 (19.3)
Current smoker	52 (34.7)
Alcohol consumption	39 (26.0)
Baseline NIHSS score, median (IQR)	11 (7–17)
Symptom onset to MRI (hours), mean ± SD	7.1 ± 3.5
Intravenous thrombolysis	47 (31.3)
Mechanical thrombectomy	24 (16.0)

MRI demonstrated a mean infarct volume of 28.6 ± 18.2 mL. Cortical infarction was slightly more frequent than subcortical infarction (54.0% vs. 46.0%). Large vessel occlusion was identified in 32.0% of patients. Hemorrhagic transformation occurred in 12.7% of cases.

Table 2: MRI characteristics of acute ischemic stroke patients

MRI Parameter	Value
Infarct volume (mL), mean ± SD	28.6 ± 18.2
Cortical infarction	81 (54.0)
Subcortical infarction	69 (46.0)
Large vessel occlusion	48 (32.0)
Mean ADC value ($\times 10^{-3}$ mm²/s)	0.63 ± 0.08
Perfusion deficit present	73 (48.7)
Mismatch (DWI-PWI) present	55 (36.7)
Hemorrhagic transformation	19 (12.7)

At 90-day follow-up, 62.7% of patients achieved a favorable functional outcome (mRS 0–2), whereas 37.3% had an unfavorable outcome (mRS 3–6). In-hospital mortality was 8.7%, infarct progression occurred in 14.7%, and hemorrhagic transformation was observed in 12.7% of patients.



Graph 1: Functional outcome at 90 days

Patients with unfavorable functional outcomes were significantly older and had higher rates of hypertension and diabetes mellitus. They also presented with higher baseline NIHSS scores, larger infarct volumes, more frequent large vessel occlusion, and perfusion deficits. Conversely, intravenous thrombolysis was significantly more common among patients with favorable outcomes.

Table 3: Comparison of patients with favorable and unfavorable functional outcomes

Variable	Favorable (n=94)	Unfavorable (n=56)	p-value
Age (years)	60.9 ± 10.3	68.4 ± 11.1	<0.001
Male sex	61 (64.9)	33 (58.9)	0.46
Hypertension	58 (61.7)	46 (82.1)	0.01
Diabetes mellitus	29 (30.9)	32 (57.1)	0.002
Baseline NIHSS	8.1 ± 3.2	16.4 ± 5.5	<0.001
Infarct volume (mL)	20.4 ± 10.2	41.8 ± 18.6	<0.001
Large vessel occlusion	18 (19.1)	30 (53.6)	<0.001
Perfusion deficit	33 (35.1)	40 (71.4)	<0.001
Intravenous thrombolysis	37 (39.4)	10 (17.9)	0.007

Multivariable logistic regression identified increasing age, diabetes mellitus, higher baseline NIHSS score, larger infarct volume, large vessel occlusion, and perfusion deficit as independent predictors of unfavorable 90-day functional outcome. Intravenous thrombolysis was independently associated with a reduced likelihood of poor outcome.

Table 4: Multivariable logistic regression for predictors of unfavorable 90-day outcome

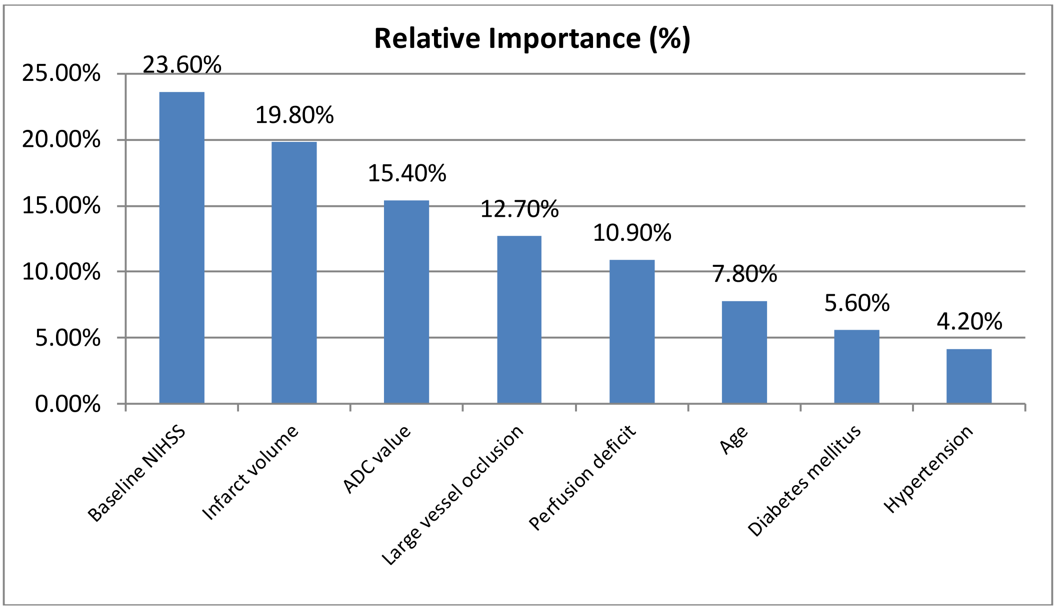
Variable	Adjusted OR	95% CI	p-value
Age (per year)	1.05	1.01–1.09	0.012
Diabetes mellitus	2.08	1.06–4.09	0.034
Baseline NIHSS	1.18	1.09–1.28	<0.001
Infarct volume	1.04	1.02–1.06	<0.001
Large vessel occlusion	2.74	1.24–6.02	0.013
Perfusion deficit	2.31	1.09–4.91	0.028
Thrombolysis	0.46	0.21–0.99	0.047

Among the evaluated machine learning algorithms, XGBoost demonstrated the best predictive performance with an accuracy of 94%, sensitivity of 93%, specificity of 95%, F1-score of 0.93, and an AUC of 0.97. Random Forest also showed excellent performance (AUC 0.95), outperforming Support Vector Machine (AUC 0.91) and Logistic Regression (AUC 0.88).

Table 5: Performance comparison of machine learning models

Model	Accuracy	Sensitivity	Specificity	Precision	F1-score	AUC
Logistic Regression	0.83	0.80	0.85	0.79	0.79	0.88
Support Vector Machine	0.86	0.84	0.88	0.84	0.84	0.91
Random Forest	0.91	0.90	0.92	0.90	0.90	0.95
XGBoost	0.94	0.93	0.95	0.93	0.93	0.97

Feature importance analysis revealed that baseline NIHSS score was the strongest predictor of functional outcome (23.6%), followed by infarct volume (19.8%), ADC value (15.4%), and large vessel occlusion (12.7%). Perfusion deficit, age, diabetes mellitus, and hypertension contributed progressively smaller but clinically relevant effects to model prediction.



Graph 2: Feature importance in the best-performing machine learning model (XGBoost)

DISCUSSION

The present study evaluated the utility of machine learning (ML)-based MRI analysis for predicting 90-day functional outcomes in patients with acute ischemic stroke. Among the evaluated algorithms, XGBoost achieved the highest predictive performance (AUC 0.97, accuracy 94%), demonstrating the value of integrating quantitative MRI features with clinical variables for early prognostic assessment.

These findings are consistent with recent evidence indicating that ensemble ML models outperform conventional statistical approaches in predicting stroke outcomes by capturing complex, nonlinear relationships among imaging and clinical parameters. Wang et al [3] reported that ML models incorporating MRI-derived features provide excellent discrimination for tissue and functional outcome prediction, while Liu et al [4], demonstrated that deep learning models significantly improve prognostic accuracy compared with traditional clinical scoring systems.

In the present study, favorable functional outcome (mRS 0–2) was achieved in 62.7% of patients, whereas 37.3% experienced unfavorable outcomes. This distribution is comparable to previous prospective stroke cohorts, where approximately 55–70% of patients achieved functional independence following timely diagnosis and evidence-based treatment [6, 7]. The observed in-hospital mortality (8.7%) and hemorrhagic transformation rate (12.7%) were also within the ranges reported in contemporary stroke registries, reflecting appropriate patient selection and acute stroke management [8].

Older age, diabetes mellitus, higher baseline NIHSS score, larger infarct volume, large vessel occlusion, and perfusion deficit emerged as independent predictors of poor functional outcome. These findings corroborate established evidence that neurological deficit at presentation and infarct burden remain the strongest determinants of post-stroke disability. Campbell et al. demonstrated that infarct volume and perfusion mismatch are robust imaging biomarkers of neurological recovery, whereas Saver et al. emphasized the prognostic importance of large vessel occlusion in determining long-term functional outcomes [9, 10]. Similarly, diabetes has been consistently associated with impaired collateral circulation, larger infarct size, and delayed neurological recovery [11].

Intravenous thrombolysis was independently associated with a significantly lower risk of unfavorable outcome, supporting current international stroke guidelines advocating early reperfusion therapy in eligible patients. Timely restoration of cerebral perfusion limits infarct expansion and preserves salvageable brain tissue, thereby improving long-term neurological recovery. These findings are in agreement with large randomized trials demonstrating substantial functional benefit following early thrombolytic treatment [12, 13].

Feature importance analysis revealed that baseline NIHSS score, infarct volume, ADC value, and large vessel occlusion contributed most substantially to model prediction. This observation highlights that combining quantitative MRI biomarkers with clinical severity enhances prognostic precision beyond conventional clinical assessment alone. Previous radiomics and artificial intelligence studies have similarly shown that diffusion metrics, infarct characteristics, and vascular imaging features are among the most influential predictors in ML-based stroke prediction models [14, 15].

Overall, the superior performance of XGBoost compared with logistic regression, support vector machine, and random forest indicates that gradient boosting algorithms are particularly effective for integrating multidimensional clinical and MRI data. Similar observations have been reported by Jo et al. and Heo et al., who demonstrated excellent discrimination (AUC >0.90) using XGBoost-based prediction models for functional outcome after acute ischemic stroke [5, 16].

CONCLUSION

Machine learning–based MRI analysis demonstrated excellent performance in predicting 90-day functional outcomes in patients with acute ischemic stroke, with the XGBoost model achieving the highest predictive accuracy (AUC 0.97). Baseline NIHSS score, infarct volume, ADC value, and large vessel occlusion were the strongest determinants of prognosis, while intravenous thrombolysis was associated with improved functional recovery. Integrating quantitative MRI biomarkers with clinical variables can enhance early risk stratification, support individualized treatment decisions, and optimize acute stroke management. Larger multicenter studies are warranted to validate these findings and facilitate the integration of machine learning models into routine clinical practice.

REFERENCES

1. GBD 2021 Stroke Risk Factor Collaborators. Global, regional, and national burden of stroke and its risk factors, 1990–2021: a systematic analysis for the Global Burden of Disease Study 2021. Lancet Neurol. 2024;23(10):973-1003. doi:10.1016/S1474-4422(24)00369-7.
2. Feigin VL, Brainin M, Norrving B, Martins SO, Pandian J, Lindsay P, et al. World Stroke Organization Global Stroke Fact Sheet 2025. Int J Stroke. 2025; 20(2):132-144. doi:10.1177/17474930241308142.
3. Wang X, Fan Y, Zhang N, Li J, Duan Y, Yang B. Performance of machine learning for tissue outcome prediction in acute ischemic stroke: a systematic review and meta-analysis. Front Neurol. 2022; 13:910259. doi:10.3389/fneur.2022.910259.
4. Liu Y, Yu Y, Ouyang J, Jiang B, Yang G, Ostmeier S, et al. Functional outcome prediction in acute ischemic stroke using a fused imaging and clinical deep learning model. Stroke. 2023;54(9):2316-2327. doi:10.1161/STROKEAHA.123.044072.
5. Jo H, Kim C, Gwon D, Lee S, et al. Combining clinical and imaging data for predicting functional outcomes after acute ischemic stroke: an automated machine learning approach. Sci Rep. 2023;13:16926. doi:10.1038/s41598-023-44201-8.
6. Goyal M, Menon BK, van Zwam WH, Dippel DWJ, Mitchell PJ, Demchuk AM, et al. Endovascular thrombectomy after large-vessel ischaemic stroke: a meta-analysis of individual patient data from five randomised trials. Lancet. 2016;387(10029):1723-1731. doi:10.1016/S0140-6736(16)00163-X.
7. Powers WJ, Rabinstein AA, Ackerson T, Adeoye OM, Bambakidis NC, Becker K, et al. Guidelines for the early management of patients with acute ischemic stroke: 2019 update to the 2018 guidelines. Stroke. 2019;50(12):e344-e418. doi:10.1161/STR.0000000000000211.
8. Thirugnanachandran T, Aitchison SG, Lim A, Ding C, Ma H, Phan T. Assessing the diagnostic accuracy of CT perfusion: a systematic review. Front Neurol. 2023;14:1255526. doi:10.3389/fneur.2023.1255526.
9. Campbell BCV, Mitchell PJ, Churilov L, Yassi N, Kleinig TJ, Dowling RJ, et al. Tenecteplase versus alteplase before thrombectomy for ischemic stroke. N Engl J Med. 2018;378(17):1573-1582. doi:10.1056/NEJMoa1716405.
10. Saver JL, Goyal M, Bonafe A, Diener HC, Levy EI, Pereira VM, et al. Stent-retriever thrombectomy after intravenous t-PA versus t-PA alone in stroke. N Engl J Med. 2015;372(24):2285-2295. doi:10.1056/NEJMoa1415061.

11. Sacco S, Foschi M, Ornello R, De Santis F, Pofi R, Romoli M. Prevention and treatment of ischaemic and haemorrhagic stroke in people with diabetes mellitus: a focus on glucose control and comorbidities. Diabetologia. 2024;67(7):1192-1205. doi:10.1007/s00125-024-06146-z.
12. Berge E, Whiteley W, Audebert H, De Marchis GM, Fonseca AC, Padiglioni C, et al. European Stroke Organisation (ESO) guidelines on intravenous thrombolysis for acute ischaemic stroke. Eur Stroke J. 2021;6(1):I-LXII. doi:10.1177/2396987321989865.
13. Nguyen TN, Abdalkader M, Fischer U, Jovin TG, Nogueira RG, et al. Endovascular management of acute stroke. Lancet. 2024;404(10468):1265-1278. doi:10.1016/S0140-6736(24)01863-7.
14. Pinto A, Brunese L, Daniele S, Faggian A, Romano L, Maggialetti N, et al. Artificial intelligence in stroke imaging: current applications and future perspectives. Diagnostics (Basel). 2022;12(6):1308. doi:10.3390/diagnostics12061308.
15. Qiu W, Kuang H, Teleg E, et al. Machine learning for detecting early ischemic stroke and predicting outcome using MRI radiomics. Front Neurosci. 2023;17:1186579. doi:10.3389/fnins.2023.1186579.
16. Heo J, Yoon JG, Park H, Kim YD, Nam HS, Heo JH. Machine learning-based model for prediction of outcomes in acute stroke. Stroke. 2019;50(5):1263-1265. doi:10.1161/STROKEAHA.118.024293.