



Access this article online

Quick Response Code:



Website:

www.turkjemergmed.com

DOI:

10.4103/tjem.tjem\_33\_23

# Predictors of 7-day symptomatic hemorrhagic transformation in patients with acute ischemic stroke and proposal of a novel screening tool: A retrospective cohort study

Mehmet Muzaffer Islam\*, Cemrenur Uygun, Melike Delipoyraz, Merve Osoydan Satici, Servan Kurt, Enis Ademoglu, Serkan Emre Eroglu

Department of Emergency Medicine, Umraniye Training and Research Hospital, University of Health Sciences, Istanbul, Turkey

\*Corresponding author

Submitted: 05-02-2023

Revised: 27-03-2023

Accepted: 03-05-2023

Published: 26-06-2023

ORCID:

MMI: 0000-0001-6928-2307

CU: 0000-0002-3780-6660

MD: 0000-0001-7347-1629

MOS: 0000-0002-3169-0724

SK: 0000-0001-6464-9931

EA: 0000-0002-6330-666X

SEE: 0000-0002-3183-3713

Address for  
correspondence:

Dr. Mehmet Muzaffer

Islam,

Department of Emergency

Medicine, Umraniye

Training and Research

Hospital, University

of Health Sciences,

Adem Yavuz Street,

No: 1, Elmalikent,

Istanbul, Turkey.

E-mail: mehmetislam1988

@gmail.com

## Abstract:

**OBJECTIVES:** Hemorrhagic transformation (HT) is significantly related to poor neurological outcomes and mortality. Although variables and models that predict HT have been reported in the literature, the need for a model with high diagnostic performance continues. We aimed to propose a model that can accurately predict symptomatic HT within 7 days of acute ischemic stroke (AIS).

**METHODS:** Patients with AIS admitted to the emergency department of a tertiary training and research hospital between November 07, 2021, and August 26, 2022, were included in this single-center retrospective study. For the model, binary logistics with the forced-entry method was used and the model was validated with 3-fold cross-validation. After the final model was created, the optimal cutoff point was determined with Youden's index. Another cut-off point was determined at which the sensitivity was the highest.

**RESULTS:** The mean age of the 423 patients included in the study was 70 (60–81) and 53.7% ( $n=227$ ) of the patients were male. Symptomatic HT was present in 31 (7.3%) patients. Mechanical thrombectomy, atrial fibrillation, and diabetes mellitus were the independent predictors ( $P < 0.001$ ,  $P = 0.003$ ,  $P = 0.006$ , respectively). The mean area under the curve of the receiver operating characteristics of the model was 0.916 (95% confidence interval [CI] = 0.876–0.957). The sensitivity for the optimal cut-off point was 90.3% (95% CI = 74.3%–97.9%) and specificity was 80.6% (95% CI = 76.4%–84.4%). For the second cutoff point where the sensitivity was 100%, the specificity was 60.5% (95% CI = 55.4%–65.3%).

**CONCLUSION:** The diagnostic performance of our model was satisfactory and it seems to be promising for symptomatic HT. External validation studies are required to implement our results into clinical use.

## Keywords:

Hemorrhage, stroke, thrombectomy, thrombolysis

## Introduction

With the introduction of the systematic approach in the treatment of

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

For reprints contact: WKHLRPMedknow\_reprints@wolterskluwer.com

**How to cite this article:** Islam MM, Uygun C, Delipoyraz M, Satici MO, Kurt S, Ademoglu E, *et al.* Predictors of 7-day symptomatic hemorrhagic transformation in patients with acute ischemic stroke and proposal of a novel screening tool: A retrospective cohort study. Turk J Emerg Med 2023;23:176-83.

**Box-ED section****What is already known on the study topic?**

- Intracranial hemorrhage after acute ischemic stroke, also known as hemorrhagic transformation (HT), is reported to occur in approximately 1.8% of ischemic stroke patients. HT is reported to be the most important cause of nonfavorable outcomes in patients with acute ischemic stroke.

**What is the conflict on the issue? Has it importance for readers?**

- Risk-stratifying models have been described to predict the various adverse outcomes in patients with acute ischemic stroke. However, the performances of these models are often reported to be insufficient.

**How is this study structured?**

- This is a retrospective cohort study.

**What does this study tell us?**

- Predictors such as administration of mechanical thrombectomy, history of atrial fibrillation or diabetes mellitus, and white blood cell count can be used to derive high-performing models for predicting 7-day symptomatic HT.
- Our model can be used as a tool for selecting patients with high risk and can aid clinicians to decide further treatment options and the need for monitorization.

ischemic stroke, the mortality of the patients decreased significantly and the focus of current studies has shifted toward the complications of stroke rather than the mortality.<sup>[1-3]</sup> Poststroke intracranial hemorrhage, also known as hemorrhagic transformation (HT), is reported to occur in approximately 1.8% of these patients and is the most important cause of nonfavorable outcomes and in-hospital mortality.<sup>[2]</sup> Therefore, predicting HT in ischemic stroke patients is crucial for preventing both mortality and morbidity.

Many studies have achieved significant results in predicting HT, in ischemic stroke patients as a whole, or in the subgroups of patients undergoing mechanical thrombectomy or intravenous tissue plasminogen activators (IV rtPA). Various prediction models have been described over time, which were created with predictors selected from patient histories, clinical features, imaging characteristics, laboratory results, and demographic features.<sup>[3-13]</sup> However, the diagnostic performance of these models could not reach the desired level and the area under the curve (AUC) of the receiver operating characteristics (ROC) analysis in most of the studies ranged from 0.550 to 0.850.<sup>[2-10,13]</sup> Therefore, the need for a high-performing and robust diagnostic tool still continues.<sup>[2]</sup>

Regarding this demand, the primary outcome of this retrospective study is to propose a regression model with high diagnostic performance to predict symptomatic HT in patients diagnosed with acute ischemic stroke in emergency medicine settings.

**Methods****Study setting**

This retrospective cohort study was conducted in the emergency medicine department of Umraniye Training and Research Hospita (*reduced due to blinding*). Our emergency medicine department has a capacity of 600000 patients annually and serves as a stroke center. This study is carried out after the approval of the local review board (stands for Umraniye Training and Research Hospital Ethical Committee *reduced due to blinding*, 20.12.2021, E-54132726-000-27300).

In our hospital, the stroke team which consists of emergency medicine, neurology, and interventional radiology physicians, is alerted for every patient suspected of stroke, and these patients are managed in a multidisciplinary manner according to the Guidelines for the Early Management of Patients With Acute Ischemic Stroke: 2019 Update.<sup>[14]</sup>

**Study population**

Patients older than 18 years of age with a diagnosis of acute ischemic stroke between the dates November 07, 2021 and August 26, 2022, were included in the study. Patients with trauma, who refused the treatment, and with missing outcome information were excluded from the study.

**Definitions and the selection of the potential predictors**

Symptomatic HT is defined as any HT accompanied with an increase of at least 4 points in total National Institutes of Health Stroke Scale (NIHSS), an increase of at least 2 points in a single NIHSS category, or intubation, mortality, hemicraniectomy, or external ventricular drain placement.<sup>[15,16]</sup> Patients with symptomatic HT within the first 7 days after admission were considered as positive events.

A literature review was performed in PubMed with the MESH terms specified in the Supplementary Material to identify the potential predictors for HT after ischemic stroke. Variables reported to be significant for any hemorrhage outcome in these studies were noted, and the potential predictors that could be collected retrospectively were included in the study.

**Study protocol**

The diagnosis of acute ischemic stroke was made based on

evidence of acute ischemia and the absence of hemorrhage on contrast-enhanced brain computed tomography angiography or diffusion magnetic resonance imaging. The diagnostic process of the potential stroke patients was carried out by the stroke team collectively.

In our hospital, the diagnosis and treatment processes of acute ischemic stroke patients are carried out according to the “Guidelines for the Early Management of Patients with Acute Ischemic Stroke: 2019 Update” published by the American Heart Association in 2019. Accordingly, patients who were candidates for recombinant tissue plasminogen activator administration received rTPA up to 4.5 hours after symptom onset.<sup>[14]</sup>

After hospitalization, stroke patients underwent a control brain CT for HT only in case of worsening in consciousness (increase of at least 4 points in total NIHSS or an increase of at least 2 points in a single NIHSS category) or vital follow-up. Routine control CT was not performed in the patients who did not have any worsening in their follow-up. Even if HT developed in these patients, they were not considered as symptomatic HT. Patients whose condition worsened during follow-up and died before CT was performed were accepted as symptomatic HT, as in previous studies.<sup>[15,16]</sup>

After applying the inclusion and exclusion criteria, a univariate analysis was performed. Afterward, the generalizability of the logistic regression method and the possibility of overfitting were tested with 3-fold cross-validation. After summarizing the cross-validation results, a regression model was proposed with the selected predictors of all patients, and the diagnostic performance of the final model was reported. The flow chart of the study method is summarized in Figure 1.

### Statistical analysis

SPSS (IBM Corp. Released 2019 IBM SPSS Statistics for Windows, version 29.0. Armonk, NY, USA: IBM Corp) program was used for the statistical analysis of the data. Categorical data were expressed as frequency and percentage and Chi-square was used for pairwise comparisons. Shapiro–Wilk test was used to determine the distribution of continuous data. All of the continuous data had nonnormal distribution and were expressed as median (25% to 75% percentile). For the pairwise comparison of the continuous data, the Mann–Whitney U-test was used.

In the literature, it has been stated that the number of predictors that should be included in the logistic regression model, especially for sensitivity studies, can be calculated according to the “5-9 events per predictor” formula.<sup>[17]</sup> According to this reference, we determined the to include four predictors in our model.

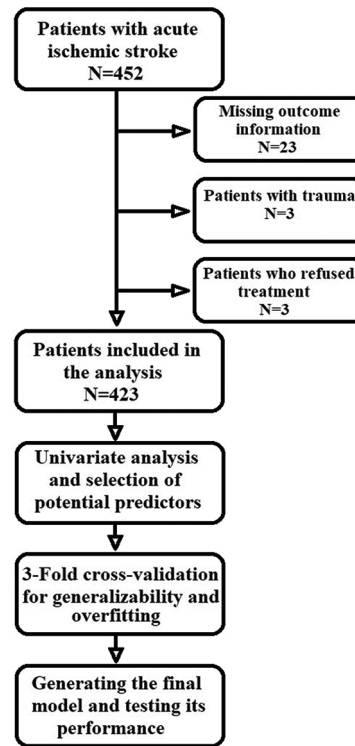


Figure 1: Patient flow chart and the summary of the method

For the selection of the predictors, variables that had a  $P < 0.05$  were included in the regression model with the forced entry method. The first four predictors which had the highest Wald statistics were included in the model.

Afterward, 3-fold cross-validation was applied to test the generalizability of this method and the possibility of overfitting. The dataset was divided into three equal folds. The model, which was trained with two folds at a time, was validated with the remaining fold. Thus, each fold was used for training and for validation. Cross-validation results were evaluated with the ROC test, reporting AUC results and accuracy. After the cross-validation, the final model was created using the same logistic regression method with the selected predictors of all patients. Multicollinearity was checked and the assumption of goodness of fit was tested with Hosmer–Lemeshow. The  $\beta$ -coefficients, Wald statistics,  $P$  values, and odds ratios of the variables in the regression model were summarized. AUC was calculated for the final model and the diagnostic test performance measures were summarized. An optimal cutoff value was set where the sum of the sensitivity and the specificity was the highest, using Youden’s index. Two additional cutoff values were defined for 100% sensitivity and for relatively balanced performance.

### Outcome measures

The primary outcome of the study is to propose a regression model with high diagnostic performance to predict the development of symptomatic HT within

7 days in patients diagnosed with acute ischemic stroke, in emergency settings.

## Results

### Basic characteristics

A total of 452 patients were diagnosed with acute ischemic stroke. Of those patients, 23 were excluded due to missing outcome variable, 3 due to trauma, and 3 due to refusal of treatment, and 423 patients were included in the final analysis [Figure 1]. The median age of the patients was 70 (60 to 81) and 227 (53.7%) patients were male. Twenty-eight (6.6%) patients were eligible for IV rtPA and 72 (17%) were eligible for mechanical thrombectomy. Thirty-one (7.3%) patients were positive for the outcome variable, HT, and 37 (8.7%) patients had in-hospital mortality [Table 1].

### Univariate analysis

In the univariate analysis, age, white blood cell count, history of atrial fibrillation, history of congestive heart failure, anterior circulation infarction, IV rtPA, and mechanical thrombectomy were significantly higher and duration between the symptom onset to treatment and history of diabetes mellitus were significantly lower in the HT group ( $P = 0.046$ ,  $P = 0.017$ ,  $P = 0.002$ ,  $P = 0.005$ ,  $P = 0.024$ ,  $P = 0.011$ ,  $P < 0.001$ ,  $P = 0.009$ ,  $P = 0.003$ , respectively) [Table 1].

### Cross-validation

For folds 1, 2, and 3, the AUC values of the models generated with the training cohort were calculated as 0.934 (0.889 to 0.978), 0.913 (0.860 to 0.967), and 0.902 (0.846 to 0.958), respectively. The AUC values of these regression functions when applied to the validation cohorts were 0.863 (0.757–0.969), 0.914 (0.843–0.985), and 0.945 (0.898–0.992). The mean AUC values of the training folds were 0.916 (0.876–0.957) and the validation folds were 0.907 (0.804–1.000). Those results were found to be consistent and summarized in Table 2.

### Regression model

In the results of the preliminary analysis, the assumption of multicollinearity was met (tolerance  $>0.1$  for all predictors) and no strong correlation was detected between any of the predictors. Standardized residuals indicated that there were 9 outliers but Cook's distances for all of the outliers were  $<1$  so none of the cases were excluded from the model. The assumption of goodness of fit was met (Hosmer–Lemeshow,  $P = 0.283$ ). The regression model was able to correctly classify 93.4% of all the cases ( $n = 432$ ,  $P < 0.001$ ) and was able to explain 43.1% of all the variance (Nagelkerke  $R^2 = 0.431$ ).

Mechanical thrombectomy, atrial fibrillation, and diabetes mellitus were found to be independent predictors of HT ( $P < 0.001$ ,  $P = 0.003$ ,  $P = 0.006$ , respectively).

**Table 1: Basic characteristics of the study population and the univariate analysis**

	All patients	HT (-)	HT (+)	P
Age	70 (60–81)	70 (59–80)	80 (63–85)	0.046
Gender (male) (%)	227 (53.7)	215 (54.8)	12 (38.7)	0.083
Symptom onset to treatment (hours)	6 (2–12)	6 (2–12)	4 (1–8)	0.009
Systolic BP (mmHg)	151 (131–171)	151 (131–171)	150 (137–180)	0.378
Diastolic BP (mmHg)	80 (71–92)	80 (70–92)	85 (73–100)	0.092
mAP (mmHg)	104 (93–117)	104 (93–116)	109 (98–123)	0.158
Pulse (bpm)	81 (74–88)	81 (73–88)	84 (78–93)	0.116
SpO <sub>2</sub> (%)	96 (95–98)	96 (95–98)	95 (93–98)	0.067
Blood urea nitrogen (mg/dl)	37 (30–47)	37 (29–47)	41 (32–58)	0.092
Creatinine (mg/dl)	0.9 (0.7–1.1)	0.9 (0.7–1.1)	0.9 (0.7–1.1)	0.569
White blood cell (10 <sup>3</sup> /μL)	8.3 (7–10.4)	8.2 (7–10.3)	9.3 (7.7–11.1)	0.017
Hemoglobin (g/dl)	13.1 (11.7–14.3)	13.1 (11.7–14.4)	13.5 (12.1–13.9)	0.701
Platelet (10 <sup>3</sup> /μL)	241 (200–300)	241 (201–300)	232 (178–285)	0.463
Diabetes mellitus (%)	175 (41.4)	170 (43.4)	5 (16.1)	0.003
Hypertension (%)	309 (73)	287 (73.2)	22 (71)	0.786
Atrial fibrillation (%)	86 (20.3)	73 (18.6)	13 (41.9)	0.002
Stroke history (%)	105 (24.8)	98 (25)	7 (22.6)	0.764
Coronary artery disease (%)	126 (29.8)	119 (30.4)	7 (22.6)	0.362
Congestive heart failure (%)	40 (9.5)	32 (8.2)	8 (25.8)	0.005
Chronic renal failure (%)	19 (4.5)	19 (4.8)	0 (0)	0.228
IV rtPA (%)	28 (6.6)	22 (5.6)	6 (19.4)	0.011
Mechanical thrombectomy (%)	72 (17)	48 (12.2)	24 (77.4)	$<0.001$
Anterior circulation infarction (%)	261 (61.7)	236 (60.2)	25 (80.6)	0.024
HT (%)	31 (7.3)	NA	NA	NA
In-hospital mortality (%)	37 (8.7)	20 (5.1)	17 (54.8)	$<0.001$

HT: Hemorrhagic transformation, mAP: Mean arterial pressure, IV rtPA: Intravenous recombinant tissue plasminogen activator, BP: Blood pressure



The variable that contributed the most to the model was found to be mechanical thrombectomy (Wald statistics = 44.944). Results of the regression model were summarized in Table 3.

The AUC of the final model for predicting HT in patients with acute ischemic stroke was 0.917 (95% CI = 0.876–0.959,  $P < 0.001$ ) [Figure 2]. Setting the optimal cutoff value as  $-1.1633$ , where the sum of the sensitivity and specificity of the regression function is the highest, the model's sensitivity was 90.3% (74.3%–97.9%), specificity was 80.6% (76.4% to 84.4%), and accuracy was 81.3% (77.3%–84.9%). A second cutoff value was set ( $-2.39$ ) for 100% sensitivity at the specificity of

60.5% (95% CI = (55.4%–65.3%) and the accuracy was 63.36% (58.57%–67.96%). The third cutoff value was defined (1.5139) for a specificity dominant performance, where the sensitivity was 25.8% (11.9%–44.6%), the specificity was 98.7% (97.1%–99.6%), and the accuracy was 93.38% (90.57%–95.56%). The diagnostic test performances of the regression model for these cutoff values are summarized in Table 4.

## Discussion

We found that the AUC value and diagnostic performance of our final model, which we proposed to predict symptomatic HT within 7 days in acute ischemic stroke patients resulted quite well. The regression function of our model showed good discrimination at the optimal cut-off value with high sensitivity and moderate-high specificity. Moreover, at a sensitivity of 100%, our model had a considerable specificity. These results indicate that our regression model can be used as a screening tool to predict symptomatic HT within 7 days in patients with acute ischemic stroke. In addition, the consistency we observed in the AUC values and accuracies in the cross-validation indicates the overall generalizability of our method and that there is no critical overfitting risk in our model.

One of the most recent studies by Bonkhoff *et al.* proposed a series of prediction models derived from approximately 75,000 patients retrospectively. The study reported that the different models, derived using various classifiers, had an AUC of 0.780–0.800 for secondary intracerebral hemorrhage.<sup>[3]</sup> In this study, for the logistic regression model, the backward stepwise method was utilized. The most valuable predictor was reported to be the administration of thrombolysis. Although many predictors could be included in this model due to the large number of events in this study, we had a more

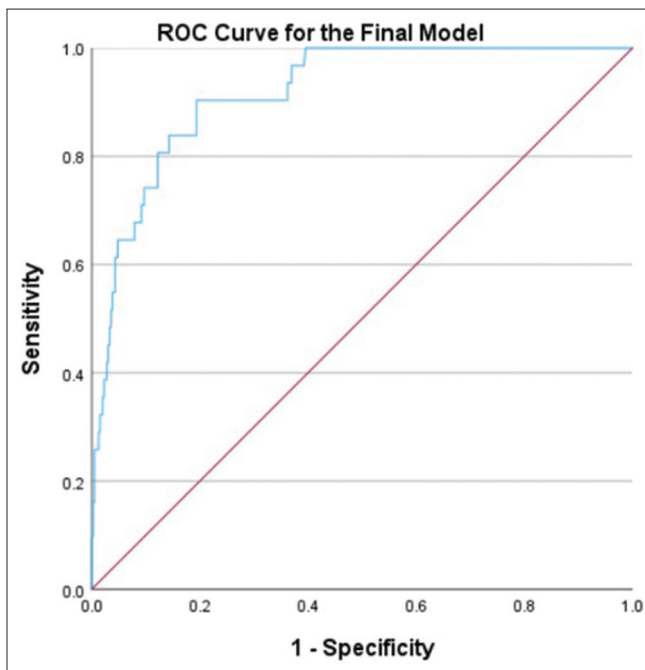


Figure 2: Receiver operating characteristics curve of the final regression model

Table 2: Summary of the results of 3-fold cross-validation

	Training		Validation	
	AUC (95% CI)	Accuracy (95% CI)	AUC (95% CI)	Accuracy (95% CI)
Fold 1	0.934 (0.889–0.978)	93.3 (89.7–95.9)	0.863 (0.757–0.969)	93.6 (88.2–97)
Fold 2	0.913 (0.860–0.967)	94.7 (91.4–97)	0.914 (0.843–0.985)	91.5 (85.6–95.5)
Fold 3	0.902 (0.846–0.958)	92.6 (88.8–95.3)	0.945 (0.898–0.992)	95 (90–98)
Mean	0.916 (0.876–0.957)	93.5 (90.9–96.2)	0.907 (0.804–1.000)	93.4 (89–97.7)

AUC: Area under the curve, CI: Confidence interval

Table 3: Results of the regression model in predicting hemorrhagic transformation

	$\beta$ -coefficients	Wald statistics	P	OR	95% CI
Mechanical thrombectomy*	3.296	44.944	<0.001	27.02	10.31–70.83
Atrial fibrillation*	1.494	9.031	0.003	4.45	1.68–11.8
Diabetes Mellitus*	-1.571	7.557	0.006	0.21	0.07–0.64
White blood cell ( $10^3/\mu\text{L}$ )	0.116	3.058	0.080	1.12	0.99–1.28
Constant (intercept)	-3.427	25.843	<0.001	0.032	NA

\*RF =  $-3.427 + (\text{mechanical thrombectomy} \times 3.296) + (\text{diabetes mellitus} \times -1.571) + (\text{atrial fibrillation} \times 1.494) + (\text{white blood cell} \times 0.116)$ , \* $\beta$ -Coefficients to be multiplied with 1 if present, 0 if absent for nominal variables. OR: Odds ratio, CI: Confidence interval, RF: Regression function

**Table 4: Diagnostic performance of the regression model for different cutoff values**

	95% CI		
	RF* cut-off=-1.1633	RF* cut-off=-2.39	RF* cut-off=1.5139
Sensitivity	90.3 (74.3–97.9)	100 (88.8–100)	25.8 (11.9–44.6)
Specificity	80.6 (76.4–84.4)	60.5 (55.4–65.3)	98.7 (97.1–99.6)
PLR	4.66 (3.69–5.88)	2.53 (2.24–2.86)	20.23 (7.04–58.15)
NLR	0.12 (0.04–0.35)	0	0.75 (0.61–0.92)
PPV	26.9 (22.6–31.7)	16.67 (15.04–18.44)	61.54 (35.76–82.14)
NPV	99.1 (97.3–99.7)	100	94.39 (93.18–95.39)
Accuracy	81.3 (77.3–84.9)	63.36 (58.57–67.96)	93.38 (90.57–95.56)

\*RF=-3.427 + (mechanical thrombectomy × 3.296) + (diabetes mellitus × -1.571) + (atrial fibrillation × 1.494) + (white blood cell × 0.116). PLR: Positive likelihood ratio, PPV: Positive predictive value, NLR: Negative likelihood ratio, NPV: Negative predictive value, RF: Regression function, CI: Confidence interval

limited choice in our model. Therefore, although there was a significant difference between the groups in the univariable analysis, thrombolysis did not significantly contribute to our model.

In the literature, there have been many proposed models examining the same outcome; however, their performance has rarely reached the desired accuracy. In particular, older studies have been found to have somewhat low performance.<sup>[4-10,13]</sup> In addition to the high AUC value of our model, its sensitivity and negative likelihood ratio were good enough to be used as a successful screening tool. Furthermore, the variables in our model (diabetes, atrial fibrillation, thrombectomy, and white blood cell count) enhance its applicability not only in the emergency department but also in other services and intensive care units, as they can be easily obtained. Due to the nature of the emergency room setting where our study was conducted, we were unable to examine variables such as Low density lipoprotein or Hemoglobin A1c (LDL or HbA1c), which were found to be valuable predictors in previous studies.<sup>[12]</sup> As a result, we were unable to assess the contribution of these variables to our model.

Although our model can be used to predict the possibility of symptomatic HT after acute ischemic stroke, it does not justify the decision to withhold thrombectomy or IV rtPA treatment, because the efficiency of these procedures outweighs the potential risk of complications. In addition, it has been emphasized that the patients who are most likely to benefit from IV rtPA treatment are also those with the highest risk of HT.<sup>[18]</sup> However, it can provide measures such as closer follow-up, better blood pressure monitoring, and performing more frequent brain computed tomography in these selected patients. Moreover, the high negative predictive value of our model can be used as a guide for transferring the selected patients to the ward follow-up from the intensive care unit sooner and help the health-care providers with the decision-making process.

There are a few studies that report a model with high performance in this outcome. One of the most recent

studies by Wei *et al.* proposed a model (SHAIS score) with five predictors consisting of patient history and imaging characteristics and found the AUC value of the model to be 0.890 for overall HT.<sup>[11]</sup> The predictors in the model were listed as history of atrial fibrillation, NIHSS, hypodensity greater than 1/3 in the middle cerebral artery territory, hyperdense artery sign, and anterior circulation infarction. The sensitivity and specificity were reported to be 70.3% and 87.5%, respectively. The authors have indicated that the SHAIS score performed better than six other previous models. However, it can be said that this study has critical limitations because it was conducted only on Chinese patients where the treatment of stroke was not performed in accordance with the current guidelines and was combined with Chinese medicine. Despite all these limitations, it would be safe to note that this model is promising.

Another recent study by Chung *et al.* has proposed a model using artificial neural networks, an advanced machine-learning method that is just beginning to make its debut in clinical practice.<sup>[12]</sup> In this model, using only five predictors (diastolic blood pressure, low-density lipoprotein, history of hyperlipidemia, history of atrial fibrillation, and history of heart disease), the authors have achieved an impressive AUC of 0.941, with 85.7% sensitivity, and 92.5% specificity. In addition, 5-fold cross-validation results have been reported quite consistent. Despite these positive aspects, this model has not yet been validated by other researchers.

The study population of almost every prediction model in the literature for HT after stroke is limited only to the subgroups of patients who underwent mechanical thrombectomy or IV rTPA. However, we did not choose such a subgroup when generating this model because we wanted to test the effect of these variables on symptomatic HT, and we thought that the presence of many factors affecting the application of thrombectomy or administration of IV rtPA could potentially cause a selection bias. Mechanical thrombectomy appeared as the variable that contributed the most to the model in our study, indicating that including this variable in future models is critical.

## Limitations

There are a number of limitations to our study. First, the retrospective nature of our study and patients with missing outcome data may have caused a potential selection bias. In addition, although the logistic regression as a method of choice was found to be generalizable for this outcome and for our dataset according to the cross-validation results, the single-center nature of our study poses a question mark on the generalizability of this model in other populations. Therefore, our model needs to be validated externally in different patient cohorts.

Although our model showed a fairly high AUC in ROC analysis, it has failed to classify nine outliers at the optimal cutoff value. We did not mention it in the results section to keep the manuscript brief, but it may be useful to declare that we tried many logistic regression methods, including the stepwise and hierarchical methods, to classify these nine cases correctly, but a misclassification problem arose in the same patients with every method. To overcome this problem, we have determined a secondary and a tertiary cutoff value for higher sensitivity at the expense of decreased specificity and accuracy. To solve this classification problem effectively, it can be considered to use more advanced classifiers (for example, artificial neural networks, deep learning, or support vector machine) that can more accurately identify the complex relationship between the predictors and the outcome.

At the beginning of our study, we determined the aim of the use of this model to predict the risk of symptomatic HT in all stroke patients with or without rTPA or with or without thrombectomy. However, we wanted to point out that this may have caused a confounding factor, as rTPA administration increases the risk of bleeding significantly as stated in the literature.

## Conclusion

There is still a need for a model in the literature to predict symptomatic HT with high performance in patients diagnosed with acute ischemic stroke. The model we proposed can play an effective role in the selection of these patients and is one of the highest-performing models in the literature. To make a stronger emphasis, our model needs to be validated in different patient cohorts.

### Author contributions

Conceptualization: Mehmet Muzaffer İslam, Enis Ademoğlu, Methodology: Mehmet Muzaffer İslam, Cemrenur Uygun, Melike Delipoyraz, Validation: Mehmet Muzaffer İslam, Cemrenur Uygun, Merve Osoydan Satıcı, Şervan Kurt, Formal analysis and investigation: Mehmet Muzaffer İslam, Enis Ademoğlu, Cemrenur Uygun, Melike Delipoyraz, Merve Osoydan Satıcı, Şervan Kurt, Writing - original draft: Mehmet Muzaffer İslam, Cemrenur Uygun, Melike Delipoyraz,

Merve Osoydan Satıcı, Şervan Kurt, Enis Ademoğlu, Serkan Emre Eroğlu, Writing - review and editing: Mehmet Muzaffer İslam, Enis Ademoğlu, Serkan Emre Eroğlu., Visualization: Mehmet Muzaffer İslam, Şervan Kurt, Melike Delipoyraz, Merve Osoydan Satıcı, Supervision: Mehmet Muzaffer İslam, Serkan Emre Eroğlu, Project administration: Serkan Emre Eroğlu, Software: None, Funding acquisition: None, Resources: None

### Conflicts of interest

None Declared.

### Ethics approval

This study was approved by the following review board: Ümraniye Eğitim ve Araştırma Hastanesi Etik Kurulu. Date of approval: December 20, 2021, Approval serial: E-54132726-000-27300.

### Funding

None.

## References

1. Feigin VL, Forouzanfar MH, Krishnamurthi R, Mensah GA, Connor M, Bennett DA, *et al.* Global and regional burden of stroke during 1990-2010: Findings from the global burden of disease study 2010. *Lancet* 2014;383:245-54.
2. Siniscalchi A. Use of stroke scales in clinical practice: Current concepts. *Turk J Emerg Med* 2022;22:119-24.
3. Bonkhoff AK, Rübsamen N, Grefkes C, Rost NS, Berger K, Karch A. Development and validation of prediction models for severe complications after acute ischemic stroke: A study based on the stroke registry of Northwestern Germany. *J Am Heart Assoc* 2022;11:e023175.
4. Cucchiara B, Tanne D, Levine SR, Demchuk AM, Kasner S. A risk score to predict intracranial hemorrhage after recombinant tissue plasminogen activator for acute ischemic stroke. *J Stroke Cerebrovasc Dis* 2008;17:331-3.
5. Lou M, Safdar A, Mehdiratta M, Kumar S, Schlaug G, Caplan L, *et al.* The HAT Score: A simple grading scale for predicting hemorrhage after thrombolysis. *Neurology* 2008;71:1417-23.
6. Menon BK, Saver JL, Prabhakaran S, Reeves M, Liang L, Olson DM, *et al.* Risk score for intracranial hemorrhage in patients with acute ischemic stroke treated with intravenous tissue-type plasminogen activator. *Stroke* 2012;43:2293-9.
7. Strbian D, Engelter S, Michel P, Meretoja A, Sekoranja L, Ahlhelm FJ, *et al.* Symptomatic intracranial hemorrhage after stroke thrombolysis: The SEDAN score. *Ann Neurol* 2012;71:634-41.
8. Mazya M, Egidio JA, Ford GA, Lees KR, Mikulik R, Toni D, *et al.* Predicting the risk of symptomatic intracerebral hemorrhage in ischemic stroke treated with intravenous alteplase: Safe Implementation of Treatments in Stroke (SITS) symptomatic intracerebral hemorrhage risk score. *Stroke* 2012;43:1524-31.
9. Flint AC, Faigles BS, Cullen SP, Kamel H, Rao VA, Gupta R, *et al.* THRIVE score predicts ischemic stroke outcomes and thrombolytic hemorrhage risk in VISTA. *Stroke* 2013;44:3365-9.
10. Saposnik G, Guzik AK, Reeves M, Ovbiagele B, Johnston SC. Stroke prognostication using age and NIH stroke scale: SPAN-100. *Neurology* 2013;80:21-8.
11. Wei C, Liu J, Guo W, Jin Y, Song Q, Wang Y, *et al.* Development and validation of a predictive model for spontaneous hemorrhagic transformation after ischemic stroke. *Front Neurol* 2021;12:747026.
12. Chung CC, Chan L, Bamodu OA, Hong CT, Chiu HW. Artificial neural network based prediction of postthrombolysis intracerebral hemorrhage and death. *Sci Rep* 2020;10:20501.
13. Wu Y, Chen H, Liu X, Cai X, Kong Y, Wang H, *et al.* A new nomogram for individualized prediction of the probability of hemorrhagic transformation after intravenous thrombolysis for

- ischemic stroke patients. *BMC Neurol* 2020;20:426.
14. 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 for the early management of acute ischemic stroke: A guideline for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke* 2019;50:e344-418.
  15. Peng Q, Hou J, Wang S, Zhou F, EY, Wang W, *et al.* Hypersensitive C-reactive protein-albumin ratio predicts symptomatic intracranial hemorrhage after endovascular therapy in acute ischemic stroke patients. *BMC Neurol* 2021;21:47.
  16. von Kummer R, Broderick JP, Campbell BC, Demchuk A, Goyal M, Hill MD, *et al.* The heidelberg bleeding classification: Classification of bleeding events after ischemic stroke and reperfusion therapy. *Stroke* 2015;46:2981-6.
  17. Vittinghoff E, McCulloch CE. Relaxing the rule of ten events per variable in logistic and Cox regression. *Am J Epidemiol* 2007;165:710-8.
  18. Whiteley WN, Thompson D, Murray G, Cohen G, Lindley RI, Wardlaw J, *et al.* Targeting recombinant tissue-type plasminogen activator in acute ischemic stroke based on risk of intracranial hemorrhage or poor functional outcome: An analysis of the third international stroke trial. *Stroke* 2014;45:1000-6.



## Supplementary Material

### Advanced search query in the PubMed using MESH Terms:

(((((predictor[Title/ Abstract]) OR (predictors[Title/ Abstract])) OR (model[Title/ Abstract])) OR (predicting[Title/ Abstract])) AND (((((((((((((((((((((((Ischemic Strokes[Title/ Abstract]) OR (Stroke, Ischemic[Title/ Abstract])) OR (Ischaemic Stroke[Title/ Abstract])) OR (Ischaemic Strokes[Title/ Abstract])) OR (Stroke, Ischaemic[Title/ Abstract])) OR (Cryptogenic Ischemic Stroke[Title/ Abstract])) OR (Cryptogenic Ischemic Strokes[Title/ Abstract])) OR (Ischemic Stroke, Cryptogenic[Title/ Abstract])) OR (Stroke, Cryptogenic Ischemic[Title/ Abstract])) OR (Cryptogenic Stroke[Title/ Abstract])) OR (Cryptogenic Strokes[Title/ Abstract])) OR (Stroke, Cryptogenic[Title/ Abstract])) OR (Cryptogenic Embolism Stroke[Title/ Abstract])) OR (Cryptogenic Embolism Strokes[Title/ Abstract])) OR (Embolism Stroke, Cryptogenic[Title/ Abstract])) OR (Stroke, Cryptogenic Embolism[Title/ Abstract])) OR (Wake-up Stroke[Title/ Abstract])) OR (Stroke, Wake-up[Title/ Abstract])) OR (Wake up Stroke[Title/ Abstract])) OR (Wake-up Strokes[Title/ Abstract])) OR (Acute Ischemic Stroke[Title/ Abstract])) OR (Acute Ischemic Strokes[Title/ Abstract])) OR (Ischemic Stroke, Acute[Title/ Abstract])) OR (Stroke, Acute Ischemic[Title/ Abstract])))) AND (((((((((((((((((((((((bleeding[Title/ Abstract]) OR (Hemorrhage, Cerebrum[Title/ Abstract])) OR (Cerebrum Hemorrhage[Title/ Abstract])) OR (Cerebrum Hemorrhages[Title/ Abstract])) OR (Hemorrhages, Cerebrum[Title/ Abstract])) OR (Cerebral Parenchymal Hemorrhage[Title/ Abstract])) OR (Cerebral Parenchymal Hemorrhages[Title/ Abstract])) OR (Hemorrhage, Cerebral Parenchymal[Title/ Abstract])) OR (Hemorrhages, Cerebral Parenchymal[Title/ Abstract])) OR (Parenchymal Hemorrhage, Cerebral[Title/ Abstract])) OR (Parenchymal Hemorrhages, Cerebral[Title/ Abstract])) OR (Intracerebral Hemorrhage[Title/ Abstract])) OR (Hemorrhage, Intracerebral[Title/ Abstract])) OR (Hemorrhages, Intracerebral[Title/ Abstract])) OR (Intracerebral Hemorrhages[Title/ Abstract])) OR (Hemorrhage, Cerebral[Title/ Abstract])) OR (Cerebral Hemorrhages[Title/ Abstract])) OR (Hemorrhages, Cerebral[Title/ Abstract])) OR (Brain Hemorrhage, Cerebral[Title/ Abstract])) OR (Brain Hemorrhages, Cerebral[Title/ Abstract])) OR (Cerebral Brain Hemorrhage[Title/ Abstract])) OR (Cerebral Brain Hemorrhages[Title/ Abstract])) OR (Hemorrhage, Cerebral Brain[Title/ Abstract])) OR (Hemorrhages, Cerebral Brain[Title/ Abstract]))