Original Article

External Validation and Updating of Published Models for Predicting 7-day Risk of Symptomatic Intracranial Hemorrhage after Receiving Alteplase for Acute Ischemic Stroke: A Retrospective Cohort Study

Phaweesa Chawalitpongpun[1,2], Paveena Sonthisombat[2,3], Napacha Piriyachananusorn[4], Natthakarn Manoyana[5]

[1]Department of Pharmaceutical Care, School of Pharmaceutical Sciences, University of Phayao, Phayao, [2]The College of Pharmacotherapy of Thailand, The Pharmacy Council of Thailand, Nonthaburi, [3]Department of Pharmacy Practice, Faculty of Pharmaceutical Sciences, Naresuan University, Phitsanulok, Departments of [4]Pharmacy and [5]Medicine, Phrae Hospital, Phrae, Thailand

Abstract

Background: Prediction scores for symptomatic intracranial hemorrhage (sICH) in acute ischemic stroke patients receiving thrombolytic therapy have been widely developed, but the external validation of these scores, especially in the Thai population, is lacking. This study aims to externally validate existing models and update the selected model to enhance its performance in our specific context. **Methods:** This cohort study retrospectively collected data from medical records between 2013 and 2022. Acute ischemic stroke patients who received thrombolysis were included. All predictors were gathered at admission. External validation was performed on eight published prediction models; in addition, the observed and expected probabilities of sICH were compared. The most effective model for discrimination was then chosen for further updating using multivariable logistic regression and was bootstrapped for internal validation. Finally, a points-based system for clinical practice was developed from the optimism-corrected model. **Results:** Fifty patients (10% of the 502 included cohort members) experienced sICH after undergoing thrombolysis. The SICH score outperformed the other seven models in terms of discrimination (area under the receiver operating characteristic [AuROC] curve = 0.74 [95% confidence interval {CI} 0.67 to 0.81]), but it still overstated risk (expected-to-observed outcomes [E/O] ratio = 1.7). Once updated, the optimism-corrected revised SICH model showed somewhat better calibration (E/O = 1 and calibration-in-the-large = 0), slightly worse underprediction in the moderate-to-high risk group (calibration slope = 1.152), and marginally better discrimination (AuROC = 0.78). The points-based system also demonstrated substantial agreement (88.1%) with the risk groups predicted by the logistic regression model (kappa statistic = 0.78). **Conclusion:** Since the SICH score outperformed seven models in terms of discrimination, it was then modified to the Revised-SICH score, which predicted that patients with at least 5.5 points were at high risk of having sICH.

Keywords: Acute ischemic stroke, external validation, intravenous thrombolysis, predictive risk score, symptomatic intracranial hemorrhage

Introduction

Symptomatic intracranial hemorrhage (sICH), a severe side effect of intravenous thrombolysis (IVT) therapy, has an incidence that ranges from 2.4% to 11.1% worldwide.[1,2] It is linked to a significantly higher risk of 24-h deterioration and 3-month mortality (odds ratio [OR] 32.3 and 18.0, respectively).[3]

Recently, many different predictive scoring systems for calculating the risk of sICH following thrombolytic therapy have been established, including the Post-thrombolysis Risk Score like Multicenter Stroke Survey (MSS), Hemorrhage After Thrombolysis (HAT) score,[5] blood Sugar, Early infarct signs and hyperDense cerebral artery sign, Age, and NIHSS (SEDAN) score,[6] Safe Implementation of Treatment in Stroke – Symptomatic Intracranial Hemorrhage (SITS-SICH) risk score,[7] Glucose Race Age Sex Pressure Stroke severity (GRASPS) score,[8] the Dense cerebral artery prestroke modified Rankin scale Age Glucose Onset-to-treatment time (DRAGON) score.[9] Stroke Prognostication using Age and NIH stroke scale (SPAN-100 index),[10] Totaled Health Risks in Vascular Events (THRIVE) score,[11] iScore,[12] Thrombolysis risk Using mRS and NIHSS (TURN) score,[13] and Symptomatic Intracranial Hemorrhage (SICH) score.[14] Due to the fact that many of these models were primarily constructed using data from Western populations, they

Address for correspondence: Dr. Paveena Sonthisombat, Department of Pharmacy Practice, Faculty of Pharmaceutical Sciences, Naresuan University, 99 Moo 9, Thapo Sub-district, Muang District, Phitsanulok 65000, Thailand.
E-mail: paveenas@nu.ac.th

Submitted: 19-Sep-2023 **Revised:** 22-Nov-2023 **Accepted:** 17-Dec-2023
Published: 06-Feb-2024

DOI: 10.4103/aian.aian_837_23

underperformed in Thai populations (area under the receiver operating characteristic curve (AuROC) 0.59–0.68 by National Institute of Neurological Disorders and Stroke (NINDS) definition and 0.53–0.68 by European Cooperative Acute Stroke Study II [ECASS II] definition).[2]

This study aimed to compare the performance of eight original predicted risk scores for sICH in patients with acute ischemic stroke (AIS) following IVT, and to update the selected discrimination model to suit our patient data for predicting the high-risk individuals.

Methods

This study's report adhered to the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) guideline [Supplementary Table S1].[15]

Study design, setting, and participants

Patients' records from Phrae hospital, a 500-bed secondary hospital in northern Thailand, were collected retrospectively between January 2013 and March 2022. The included patients were at least 18 years of age, had an AIS diagnosis, and received IVT. Those who were transferred out within the first 24 h of admission or who had at least one component of the clinical risk scores missing were excluded from our study.

Ethics approval

This study was conducted in accordance with the principles of the Declaration of Helsinki. Approval was granted by the Phrae Hospital Ethical Committee for Clinical Research on January 27, 2022, with the approval number 1/2565.

Data collection and data source

Our study's data source was the medical records' extraction of the demographic information, medical history, and clinical information from admission. The severity of stroke was determined according to the National Institutes of Health Stroke Scale (NIHSS) and the Glasgow coma score (GCS). Blood glucose level, platelet count, serum creatinine, and estimated glomerular filtration rate (eGFR) based on the Chronic kidney disease epidemiology collaboration (CKD-EPI) equation were the laboratory measures included in the study. Brain computed tomography (CT) imaging was conducted at the time of admission and again within 24 h of initiating IVT, or whenever neurologic deterioration was noticed. Only cases without any missing data were included in the study.

Outcome

The study outcome was the patient experiencing sICH within 7 days after thrombolytic therapy. To diagnose sICH, individuals had to meet the ECASS-II criteria, in addition to experiencing clinical deterioration (defined as an increase in NIHSS of at least 4 points from baseline or from the lowest score during the first 7 days of admission) and bleeding at any site in the brain region as seen on brain CT.[16]

Predictors

All predictors, collected at the time of hospital admission, were derived from existing sICH prediction scoring systems. Supplementary Table S2 provides information on the predictors for each scoring system.

Sample size calculation

We calculated the sample size using the following formula:[17]

$$N = [\frac{Z^2_{\alpha/2}\ V(AUC)}{d^2}]$$

where $\alpha = 0.05$, d = ±10%, AUC = 80%,[18] and the estimated incidence of sICH is 10.1%.[19] Therefore, 468 patients represented the minimum number of subjects needed for this investigation.

External validation model

We conducted a systematic search in the Embase, PubMed, and Google Scholar databases using the following search terms: symptomatic intracranial hemorrhage, intravenous thrombolysis, predictive risk score, and alteplase. Eleven published models for predicting sICH were found, which included MSS,[4] HAT score,[5] SEDAN score,[6] SITS-SICH risk score,[7] GRASPS score,[8] SPAN-100 index,[10] THRIVE score,[11] SICH score,[14] DRAGON score,[9] TURN score,[13] and iScore.[12] However, the last three models incorporated variables which were not regularly collected in our setting. As a result, just eight models received external validation.

To verify the models externally, we first calculated the probability of sICH for individual participants using the original scoring system. Then, we computed and compared the performance parameters of each model. These parameters included the AuROC curve, expected-to-observed outcomes (E/O) ratio, calibration-in-the-large (CITL), and calibration slope (C-slope). In addition, we determined each scoring system's sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio (LR+), negative likelihood ratio (LR−), and correct classification rate (CCR).

Model updating, further internal validation, and creating a points-based system

The model with the best discrimination during the external validation process was selected for further updating and internal validation. To enhance model calibration, we modified the model's intercept and the beta-coefficients of each predictor. Further internal validation using a bootstrapping method with 500 cycles was performed, and the model was adjusted for optimism, derived from bootstrapping, accordingly. The full details of the external validation, model updating, and internal validation process can be found in the supplementary materials.

Finally, we applied a method established by Sullivan *et al*.[20] to the optimism-corrected updated model to produce a points-based system (PBS). The rationale for assigning different scores to each predictor can be found in the supplementary materials. Following discussions among experts, a cut-off score was used to create four risk groups based on the probability of sICH: very low risk (<5.1%; 0–1 point), low risk (≤35.3%; 1.5–3.5 points), moderate risk (≥46.5%; 4–5 points), and high risk (≥77.7%; 5.5–6 points). To validate PBS with the logistic model, the percentage agreement and kappa statistic were calculated.[21]

The full details of the external validation process are presented in Supplementary S1.

This study's analyses were all carried out utilizing a complete case approach. For the statistical analysis, we used STATA version 14 (StataCorp LLC, College Station, TX, USA), and the level of statistical significance was set at a two-sided alpha error of 0.05.

Results

Patient characteristics

Figure 1 shows the procedure for including and excluding patients from the study. The study included 502 AIS patients who received IVT between January 2013 and March 2022. According to the ECASS-II criteria, 50 patients (or about 10%) developed sICH.

The general population's average age was 66.4 years, and 51.6% of them were male. It included 211 (42%), 102 (20.3%), and 95 (18.9%) cases of hypertension, dyslipidemia, and diabetes mellitus, respectively. Furthermore, 83 (16.5%) people reported using aspirin. The mean duration from the onset of symptoms to the administration of alteplase was 186.5 minutes, with the average dose being 0.87 mg/kg. Table 1 presents the baseline characteristics of the patients. In terms of sex distribution, age, weight, alteplase dosage, period from onset to treatment, and the majority of underlying disorders, the two groups exhibited similarities ($P > 0.05$). On the initial brain CT, patients with sICH showed early ischemia alterations, in contrast to those without sICH. These changes included hypodensity ≥1/3 of the middle cerebral artery (MCA) territory ($P < 0.001$), a sign of hyperdense cerebral artery ($P = 0.001$), and a sign of early infarction ($P < 0.001$). Moreover, patients with sICH had a higher NIHSS (14.5 vs. 9.0, $P < 0.001$) and a lower GCS (11.5 vs. 14.0, $P = 0.01$). Before the introduction of alteplase, more sICH patients (53.1% vs. 23.9%, $P < 0.001$) took intravenous (IV) antihypertensive drugs, and their baseline systolic blood pressure (SBP) was higher (158.8 vs. 150.6 mmHg, $P = 0.03$) than in those without sICH.

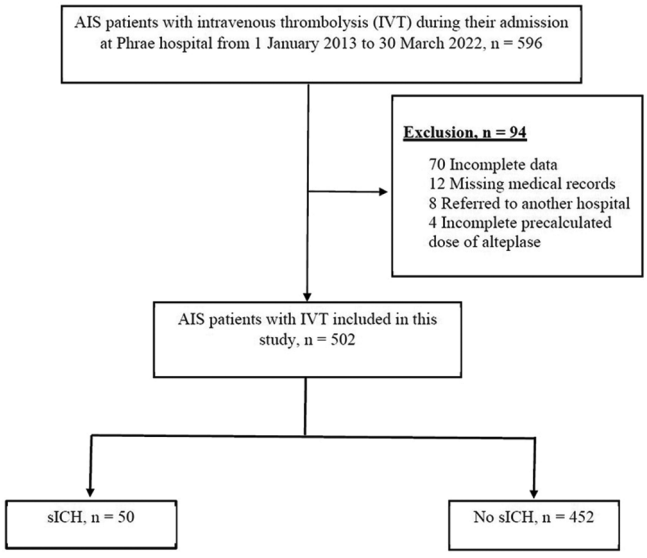


Figure 1: Patient selection flowchart

External validation of existing sICH prediction models

Table 2 provides a summary of the eight models' abilities to predict sICH. The top three models for predicting sICH with the highest discrimination scores were SICH (AuROC = 0.74), SEDAN (AuROC = 0.73), and HAT (AuROC = 0.72). The SPAN-100 index, however, displayed weak discrimination (AuROC = 0.49) [Table 2, Supplementary Figure S1].

Regarding model calibration, the GRASPS, HAT, SITS-SICH, THRIVE, and SEDAN scores tended to underestimate the incidence of sICH (E/O = 0.40, 0.44, 0.45, 0.52, and 0.65, respectively), whereas the SPAN-100 index, SICH score, and MSS tended to overestimate the risk (E/O = 1.31, 1.70, and 2.47, respectively) [Table 2, Supplementary Figure S2].

The SICH score demonstrated the best discrimination among the eight prediction models currently in use. This scoring model included a number of variables, including valvular heart disease (VHD), aspirin use, SBP before thrombolysis treatment, NIHSS, platelet count, and IV antihypertensive medication administration during thrombolysis. This score has the benefit of not requiring a brain CT scan, a parameter with poor interpretative accuracy, especially in the early stages of AIS, and frequently requiring the interpretation of an expert and a trained radiologist. Despite its strong discriminatory power, the SICH score was found to overstate the risk. We then went ahead and updated the model to improve its functionality.

The updated and optimism-corrected updated model

The Revised-SICH score was the name given to the modified SICH model. An improvement of performance was indicated as AuROC = 0.78 (95% confidence interval [CI]: 0.71 to 0.84), CITL = 0.00, C-slope = 1.00, and E/O = 1.00. Only two of the six predictors of original SICH models significantly predicted the likelihood of sICH in our cohorts: IV antihypertensive drugs during thrombolysis (beta-coefficient 1.067 [95% CI: 0.429 to 1.704], $P = 0.001$), NIHSS 10–20 (beta-coefficient 1.775 [95% CI: 0.931 to 2.620], $P < 0.001$), and NIHSS >20 (beta-coefficient 2.761 [95% CI: 1.583 to 3.939], $P < 0.001$) [Table 3]. Also, none of the predictors showed any signs of multicollinearity (mean variance inflation factor (VIF) = 1.05 [range: 1.01–1.07]).

After internal validation using bootstrapping, the models' AuROC and CITL remained constant. The C-slope did, however, decline by 13.2% from the initial value, indicating probable optimism necessitating additional correction [Supplementary Table S3]. AuROC, CITL, and E/O stayed unchanged after the model was shrunk; however, the C-slope changed to 1.152, indicating a small underprediction. Figure 2 shows a comparison of the calibration plots for the externally validated original SICH model, the updated SICH model, and the updated optimism-corrected SICH model. Table 3 gives a

Table 1: Clinical and demographic characteristics of the study patients

Characteristics	sICH (*n*=50)	No sICH (*n*=452)	*P*
Male sex, *n* (%)	24 (48.0)	235 (52.0)	0.59
Age (years), mean (±SD)	68.2 (±12.6)	66.3 (±11.8)	0.28
Weight (kg), mean (±SD)	56.9 (±11.8)	56.1 (±11.5)	0.64
Alteplase dosing (mg/kg), mean (±SD)	0.9 (±0.1)	0.9 (±0.1)	1.00
Onset-to-treatment time (min), mean (±SD)	187.2 (±55.6)	186.4 (±51.3)	0.92
IV antihypertensive drugs, *n* (%)	26 (52.0)	103 (22.8)	<0.001
Medical history, *n* (%)			
Hypertension	23 (46.0)	188 (41.6)	0.55
Atrial fibrillation	13 (26.0)	118 (26.1)	0.99
Dyslipidemia	14 (28.0)	88 (19.5)	0.16
Diabetes mellitus	11 (22.0)	84 (18.6)	0.56
Previous stroke	7 (14.0)	34 (7.5)	0.11
Valvular heart disease	1 (2.0)	6 (1.3)	0.52
Chronic kidney disease	7 (14.0)	32 (7.1)	0.08
Concurrent medication, *n* (%)			
Aspirin	12 (24.0)	71 (15.7)	0.13
Clopidogrel	0 (0.0)	8 (1.8)	1.00
Aspirin and clopidogrel	0 (0.0)	10 (2.2)	0.61
Warfarin	0 (0.0)	19 (4.2)	0.24
Brain CT finding, *n* (%)			
Hypodensity			
<1/3 of MCA territory	11 (22.0)	54 (12)	<0.001
≥1/3 of MCA territory	8 (16.0)	5 (1.1)	
Sign of hyperdense cerebral artery	19 (38.0)	82 (18.2)	0.001
Sign of early infarction	27 (54.0)	99 (21.9)	<0.001
Severity of stroke, median (IQR)			
NIHSS	14.5 (11–18)	9 (6–14)	<0.001
GCS	11.5 (10–15)	14 (11–15)	0.01
Baseline clinical data			
SBP (mmHg), mean (±SD)	158.8 (±25.3)	150.6 (±24.8)	0.03
DBP (mmHg), mean (±SD)	85.0 (±13.8)	85.2 (±24.8)	0.93
Blood glucose (mg/dL), median (IQR)	125 (105–153)	118 (100.5–151)	0.31
Platelets ($\times 10^3$/mm^3), mean (±SD)	236.6 (±87.5)	246.4 (±74.6)	0.38

CT=Computerized tomography, DBP=Diastolic blood pressure, GCS=Glasgow coma scale, IQR=Interquartile range, MCA=Middle cerebral artery, NIHSS=National Institutes of Health Stroke Scale, SBP=Systolic blood pressure, SD=standard deviation, sICH=Symptomatic intracranial hemorrhage

comparison of the performance of the updated SICH model and the updated SICH model with optimism correction.

Each predictor in the optimism-corrected updated SICH model was given the following number of points: 0.5 points for VHD, aspirin use, SBP less than 140 mmHg before thrombolysis, and platelet count less than 250,000 cells/mm^3; 1 point for IV antihypertensive medication while undergoing thrombolysis; 2 points for NIHSS scores of 10–20; and 3 points for NIHSS scores greater than 20 [Table 3, Figure 3]. Higher scores indicated a higher risk of sICH within 7 days after taking alteplase, with the total score ranging from 0 to 6 [Table 3, Figure 3].

To facilitate the use of the optimism-corrected updated SICH model in clinical practice, patients were categorized into four risk groups based on their total points: very low risk (≤5.1%; 0–1 point), low risk (≤35.3%; 1.5–3.5 points), moderate risk (≥46.5%; 4–5 points), and high risk (≥77.7%; 5.5–6 points). When compared to the predicted risk from logistic regression, PBS with four risk groups demonstrated substantial agreement (88.1%) with a kappa statistic of 0.78 ($P < 0.001$). However, compared to the risk predicted by the regression model, PBS was likely to overestimate patients who were at moderate-to-high risk [Supplementary Table S4].

Discussion

This study found that most participants were over 60 years old and male, both of which are significant risk factors in AIS patients. In addition, all patients received IVT at doses and treatment timelines consistent with current treatment guidelines.[22] According to these results, the severity of AIS, the presence of early ischemic alterations on baseline brain CT, the use of IV antihypertensives during thrombolysis administration, and a higher SBP at the baseline were all statistically significant predictors of sICH.

Table 2: Performance of each score in predicting sICH in AIS patients receiving IVT treatment

Model	AuROC	95% CI	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	LR+	LR−	CCR (%)	E/O
MSS ≥2	0.59	0.52–0.66	96	16.59	10.99	97.53	1.15	2.24	24.5	2.47
HAT ≥1	0.72	0.65–0.79	74	64.82	18.41	95.88	2.10	0.40	65.74	0.44
SEDAN ≥2	0.73	0.64–0.79	82	58.45	17.47	96.80	1.97	0.31	60.99	0.65
SITS-SICH ≥4	0.65	0.57–0.73	72	52.43	13.97	94.59	1.51	0.53	54.38	0.45
GRASPS ≥72	0.63	0.58–0.69	92	41.15	14.36	97.96	1.56	0.94	41.15	0.40
SPAN-100 positive	0.49	0.47–0.51	2	96.24	5.40	90.16	0.53	1.02	86.85	1.31
SICH ≥3	0.74	0.67–0.81	82	55.31	16.45	96.63	1.83	0.33	57.97	1.70
THRIVE ≥3	0.62	0.56–0.68	76	50.22	14.08	95.13	1.53	0.48	52.79	0.52
Very good performance		Good performance			Moderate performance				Poor performance	
AuROC 0.80–0.89		AuROC 0.70–0.79			AuROC 0.50–0.69				AuROC <0.5	
Sensitivity 80.00%–100.00%		Sensitivity 70.00%–79.99%			Sensitivity 50.00%–69.99%				Sensitivity <50.00%	
Specificity 80.00%–100.00%		Specificity 70.00%–79.99%			Specificity 50.00%–69.99%				Specificity <50.00%	
PPV 80.00%–100.00%		PPV 70.00%–79.99%			PPV 50.00%–69.99%				PPV <50.00%	
NPV 80.00%–100.00%		NPV 70.00%–79.99%			NPV 50.00%–69.99%				NPV <50.00%	
LR+ >10.00		LR +5.00–10.00			LR +2.00–4.99				LR+ 1.00–1.99	
LR− <0.1		LR −0.1–0.19			LR −0.2–0.49				LR −0.5–1	
CCR 80.00%–100.00%		CCR 70.00%–79.99%			CCR 50.00%–69.99%				CCR <50.00%	
E/O=1									E/O <1 or >1	

AIS=acute ischemic stroke, AuROC=area under the receiver operating characteristic, CI=confidence interval, CCR=correct classification rate, E/O=expected to observed outcomes, GRASPS=Glucose Race Age Sex Pressure Stroke severity score, HAT=Hemorrhage After Thrombolysis score, IVT=intravenous thrombolysis, LR− = negative likelihood ratio, LR+ = positive likelihood ratio, MSS=Post-thrombolysis Risk Score, NIHSS=National Institutes of Health Stroke Scale, NPV=negative predictive value, PPV=positive predictive value, SBP=systolic blood pressure, SEDAN=blood Sugar, Early infarct signs and hyperDense cerebral artery sign, Age, and NIHSS score, sICH=symptomatic intracranial hemorrhage, SICH=Symptomatic Intracranial Hemorrhage score, SITS-SICH=Safe Implementation of Treatment in Stroke – Symptomatic Intracranial Hemorrhage risk score, SPAN=Stroke Prognostication using Age and NIH stroke scale, THRIVE=Totaled Health Risks In Vascular Events

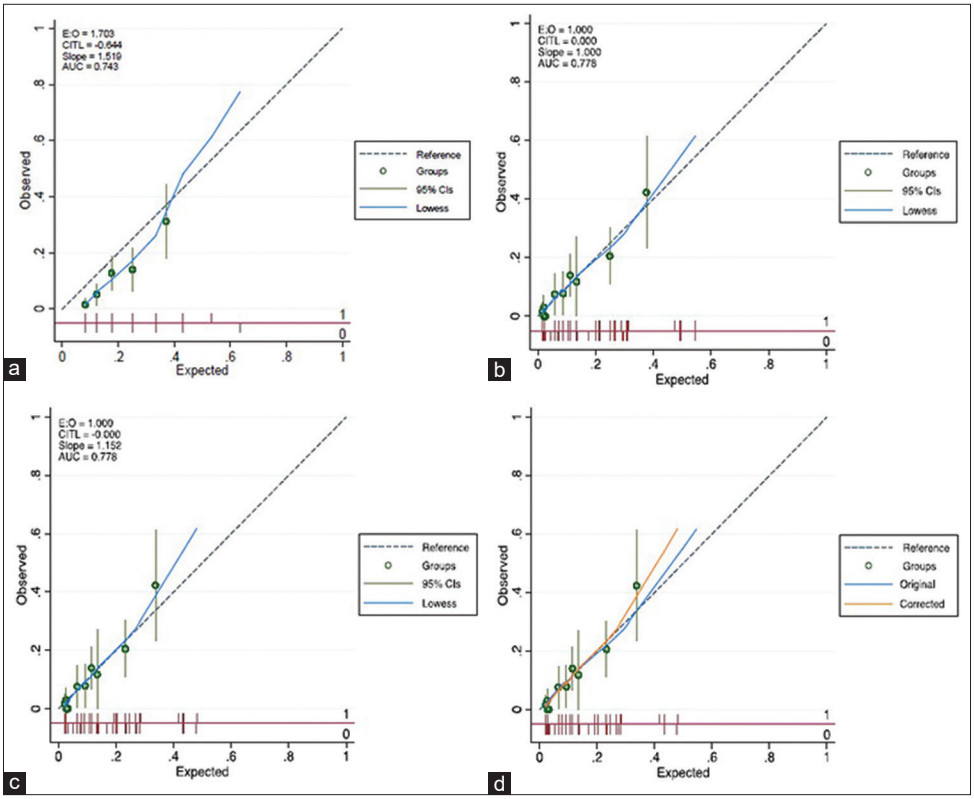


Figure 2: Calibration plots for (a) original SICH score, (b) Revised-SICH score, (c) optimism-corrected Revised-SICH score. (d) Comparison of calibration plots between the Revised-SICH model and optimism-corrected Revised-SICH model

NIHSS, a tool for assessing the severity of strokes, was also found to be associated with sICH.[23,24] Patients who have experienced a severe stroke (NIHSS more than 20) are at an elevated risk of developing sICH by 5.06 times.[24] Therefore,

Table 3: Comparisons between updated and optimism-adjusted updated SICH models with corresponding points

Predictors	Updated model			Optimism-adjusted updated model	
	OR (95% CI)	Beta	95% CI, *P*	Beta[a]	Corresponding points
NIHSS					
<10	1	0		0	0
10–20	5.90 (2.54 to 13.74)	1.775	0.931 to 2.620, <0.001	1.541	2
>20	15.82 (4.87-51.37)	2.761	1.583 to 3.939, <0.001	2.396	3
IV antihypertensive drugs during thrombolysis					
No	1	0		0	0
Yes	2.91 (1.54 to 5.50)	1.067	0.429 to 1.704, 0.001	0.926	1
Valvular heart disease					
No	1	0		0	0
Yes	1.69 (0.17 to 17.24)	0.526	-1.794 to 2.847, 0.66	0.457	0.5
Use of aspirin					
No	1	0		0	0
Yes	1.23 (0.58 to 2.63)	0.210	-0.549 to 0.968, 0.59	0.182	0.5
SBP before thrombolysis (mmHg)					
<140	1	0		0	0
≥140	1.34 (0.62 to 2.87)	0.292	-0.471 to 1.055, 0.45	0.254	0.5
Platelet count (cells/mm^3)					
≥250,000	1	0		0	0
<250,000	1.01 (0.53 to 1.93)	0.011	-0.635 to 0.657, 0.97	0.010	0.5
Model's intercept		−4.154		-3.846	
Model's performance					
Discrimination					
AuROC		0.778	(95% CI: 0.713-0.844)	0.778	(95% CI: 0.713 to 0.844)
Calibration					
E/O		1.000		1.000	
CITL		0.000		0.000	
C-slope		1.000		1.152	

[a]Optimism-adjusted betas were derived by multiplying updated betas by 0.868. AuROC=area under the receiver operating characteristic, CI=confidence interval, CILT=calibration-in-the-large, C-slope=calibration slope, E/O=expected to observed outcomes, IV=intravenous, NIHSS=National Institutes of Health Stroke Scale, OR=odds ratio, SBP=systolic blood pressure, SICH=Symptomatic Intracranial Hemorrhage score

among the eight predictive risk scores for sICH in individuals with AIS after IVT, NIHSS was a crucial predictor and component. The study's findings were in line with those of Tanne *et al*.,[25] who discovered that early ischemia alterations in brain CT were related to sICH in both minor (33% of MCA territory) and major (>33% of MCA territory) forms. Early cytotoxic edema, persistent hypoperfusion, and irreversible tissue damage are the hallmarks of these ischemic changes. The chance of developing sICH increases by 2.6 when either baseline SBP or diastolic blood pressure (DBP) is more than 185 and 110 mmHg, respectively.[26] As a result, the current recommendation is for people with blood pressure below 185/110 mmHg to use alteplase.[22]

With AuROC values of 0.74, 0.73, and 0.72, respectively, this study demonstrated that the SICH score, SEDAN score, and HAT score exhibited good discrimination for predicting sICH after receiving IVT. These three models contain at least two significant parameters, linked to sICH in this study, including high NIHSS, high SBP, early infarct signs, MCA territory hypodensity/hyperdensity on brain CT, or the use of IV antihypertensive medications during IVT. The SPAN-100 index, on the other hand, lacked the ability to predict sICH because it relied on just two factors, one of which, age, was not statistically significant in connection to sICH in this investigation. As a result, this model's AuROC fell.

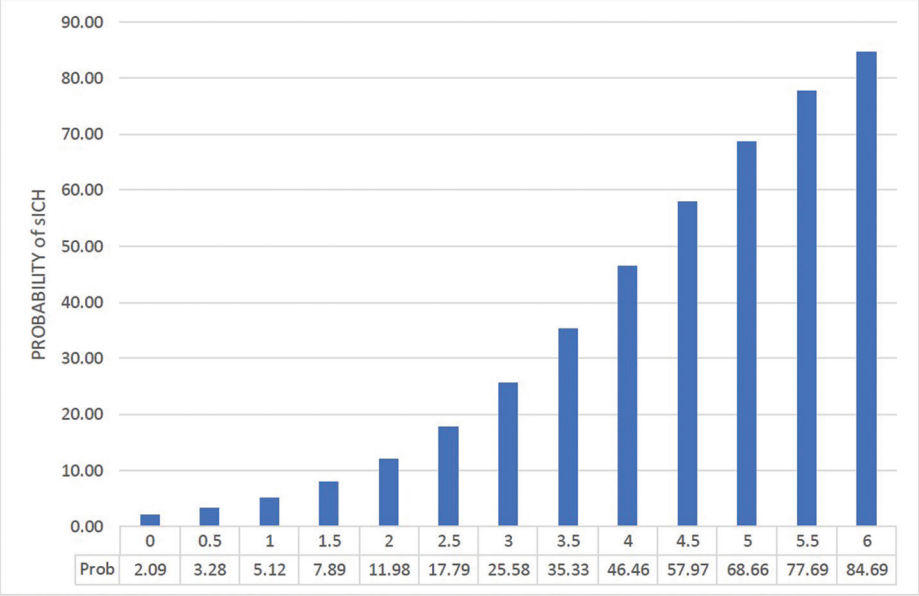
Based on earlier studies, the predictive performance of the sICH score was inconsistent. A study of a Western population by Asuzu *et al*.[27] discovered that four predictive models, the DRAGON score (AuROC = 0.76), Stroke-Thrombolytic Predictive Instrument (Stroke-TPI) (AuROC = 0.74), Acute Stroke Registry and Analysis of Lausanne (ASTRAL) score (AuROC = 0.72), and HAT score (AuROC = 0.70), possessed good discrimination power. However, a different retrospective cohort study found that the SEDAN score had the highest predictive power (AuROC = 0.70), whereas the SPAN-100 index consistently had the lowest predictive power (AuROC = 0.56).[28] Despite having the highest predictive value in an Asian population, the HAT score's discriminatory power was poor, according to a Taiwanese retrospective cohort study (AuROC = 0.69).[29] A multicenter prospective investigation in a Chinese population found that MSS had the highest predictive performance for sICH when

Calculation tool for predicting symptomatic intracerebral hemorrhage after intravenous thrombolysis in acute ischemic stroke

Please answer the questions below to calculate.

Predictors	No		Yes	
Valvular heart diseases	○	0 point	○	0.5 point
Use of aspirin	○	0 point	○	0.5 point
SBP prior to thrombolysis ≥ 140 mmHg	○	0 point	○	0.5 point
NIHSS				
< 10 points	○	0 point	○	0 point
10–20 points	○	0 point	○	2 points
> 20 points	○	0 point	○	3 points
Platelet count < 250 (× 10^3 cells/mm^3)	○	0 point	○	0.5 point
IV anti-hypertensive drugs	○	0 point	○	1 point
Total points				

*** IV, Intravenous; NIHSS, National Institutes of Health Stroke Scale; SBP, Systolic blood pressure.**



Very Low risk	Low risk	Moderate risk	High risk

Figure 3: The Revised-SICH score and corresponding probability of sICH. sICH = symptomatic intracranial hemorrhage

using the NINDS and ECASS-II definitions, but when using the Safe Implementation of Thrombolysis in Stroke-Monitoring Study (SITS-MOST) definition, the GRASPS score was the best predictor of sICH.[30] Results from two research studies on the Thai population were likewise inconsistent. Meanwhile, another retrospective cohort study indicated that the DRAGON score was the most accurate predictor, with an AuROC of 0.74, followed by the SEDAN score (AuROC = 0.73) and the HAT score (AuROC = 0.70),[31] while Suengtaworn *et al*.[2] did not discover a model with good discrimination power for predicting sICH. It is evident from the previous studies that the models' capacity to forecast the prevalence of sICH differs between studies for various reasons discussed below.

First, various scoring systems have been created using different definitions of sICH, leading to varying rates of sICH and different capacities to predict sICH outcomes following IVT. In our study, the ECASS-II definition was used, which was found to have the highest inter-rater agreement,[32] and a consensus had to be reached by at least two of the three blinded sICH evaluators: one neurologist, one radiologist, and one emergency physician.

Second, brain CT images during early infarction changes in AIS patients may not show evident abnormalities,[33] necessitating the expertise and experience of radiologists for interpretation. The prediction model may become uncertain as a result of these restrictions.

Third, altering the alteplase dose to 0.9 mg/kg and timing of delivery to 3–4.5 h after the onset of AIS, respectively, are recent recommendations for treating AIS.[22] This might make it easier to forecast sICH in the future. According to a retrospective cohort study carried out in Thailand, the length of time between the diagnosis and the start of treatment is substantially related to sICH.[34] Therefore, studies involving participants who received 0.6 mg/kg of alteplase or treatment within 3 h may have demonstrated a diminished ability to predict the occurrence of sICH.[2,29]

Fourth, Asian ethnicity is a significant risk factor for sICH, according to the findings of numerous research studies.[7,35] Genetic variations and disparities in health-care access related to the prevention of vascular risk factors also play a role in this outcome among these populations.[36] Consequently, the model's capacity to forecast the development of sICH may vary by ethnic groups. Furthermore, because the SICH score was developed from a study of Thai patients, it takes into account patients who are genetically similar to those in this study and have a similar public health system and way of life. This might improve the model's potential for prediction.

The SICH score was a useful discriminating model for predicting sICH without consulting experts, despite its calibration exaggerating risk. As a result, the Revised-SICH score has good calibration (CITL = 0.000, C-slope = 1.00, and E/O = 1) and discrimination (AuROC = 0.78). Accordingly, high-risk IVT patients who have a Revised-SICH score of more than or equal to 5.5 points merit close observation in the intensive care unit, and risk factors for sICH, like an elevated SBP, ought to be treated before IVT administration. However, this model includes factors such as a history of VHD and aspirin use, which can be a drawback for patients with obscure medical histories. In addition, PBS tended to overstate the risk for patients at moderate-to-high risk compared to logistic regression estimates. This contrast with the regression model's estimates might stem from this study's findings. The study discovered a strong link between aberrant brain CT imaging and the development of sICH. Despite this, the Revised-SICH scoring system does not include CT scan factors. As a result, PBS's estimate of risk deviated from expectations as a result of this exclusion.

Strengths and limitations

To our knowledge, this is the first study that has externally validated and updated a prediction model for sICH in the Thai population. Given the substantial number of existing predictive tools, our work underscores the significance of the external validation process before their application across diverse settings. We have demonstrated that even models created for particular subsets of the Thai population exhibit suboptimal performance when applied to various subgroups of the same population.

There are a few limitations in our work that warrant attention, nevertheless. In the first place, the Revised-SICH score had elements including a history of VHD and aspirin use, which can provide challenges for patients with unreported medical histories. Second, our regression model performed poorly at identifying the high-risk group. This can be attributable, in part, to the removal of predictors like anomalies in brain CT imaging that have shown strong links with sICH. Due to the difficulty of interpreting such imaging in the early phases of sICH, a process prone to high variability, we have chosen not to use brain CT imaging as supporting evidence. The patients receiving IVT during the window of 3–4.5 h following the start of AIS are also a special focus of our model. Consequently, the model's generalizability might be constrained for individuals with ambiguous onset periods or those admitted beyond this time frame. Finally, the lack of statistical significance demonstrated for several predictors in our analysis may be attributed to the very small number of occurrences, evident in the wide and inaccurate confidence ranges. This could be partly due to the retrospectively collected data, with approximately 13.2% of total patients receiving IVT during their admission having missing data and not being included in our analysis. This might introduce some selection bias and result in a reduced sample size, impacting patients who could potentially develop sICH. While taking into account the various contexts and patient populations involved, our study continues to be important in offering insights into the applicability and generalizability of the risk prediction model within our particular setting, despite these acknowledged limitations.

Implications

Regarding the clinical implications of our research, we advise that high-risk IVT patients who have a Revised-SICH score of 5.5 points or more merit close observation in the critical care unit and that risk factors for sICH, like an elevated SBP, should be treated before IVT administration. Further studies are necessary to confirm the relationship between the abnormality of the CT brain findings in the early infarction phase and the occurrence of sICH. The effectiveness of the instrument can also be increased by tweaking the Revised-SICH score, especially when predicting a high-risk group.

Conclusion

The SICH score was a more accurate predictor of the occurrence of sICH than the other seven scores, although its calibration overstated risk. The Revised-SICH score incorporated evaluable variables for all medical professionals and exhibited good discrimination and calibration. Patients at high risk for sICH predicted by the Revised-SICH score must have their signs and symptoms closely monitored for at least 24 h after IVT initiation in the critical care unit, and concurrently decreasing risk factors, particularly blood pressure, may help minimize the risk of sICH.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors used ChatGPT, an AI language model developed by OpenAI, to improve readability and language. After using this tool, the authors reviewed and edited the content as needed, and they take full responsibility for the content of the publication.

Availability of data and materials

The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

Acknowledgements

We would like to thank the Phrae hospital staff for providing patient data and Dr. Nat Na-ek (School of Pharmaceutical Sciences, Phayao University) for assistance with the statistical analysis.

Financial support and sponsorship
Nil.

Conflicts of interest
There are no conflicts of interest.

References
1. National Institute of Neurological Disorders and Stroke rt-PA Stroke Study Group. Tissue plasminogen activator for acute ischemic stroke. N Engl J Med 1995;333:1581–7.
2. Suengtaworn A, Saposnik G, Hurst C, Poungvarin N, Nilanont Y. Accuracy of clinical risk scores in predicting post-rtPA intracerebral hemorrhage in a Thai cohort. J Med Assoc Thai 2019;102:1–7.
3. Fiorelli M, Bastianello S, Von Kummer R, Del Zoppo GJ, Larrue V, Lesaffre E, *et al*. Hemorrhagic transformation within 36 hours of a cerebral infarct relationships with early clinical deterioration and 3-month outcome in the European Cooperative Acute Stroke Study I (ECASS I) Cohort. Stroke 1999;30:2280–4.
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 PA, 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. 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.
7. Mazya M, Egido 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.
8. Menon BK, Saver JL, Prabhakaran S, Reeves M, Liang L, Daiwai M, *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.
9. Strbian D, Meretoja A, Ahlhelm FJ, Pitkäniemi J, Lyrer P, Kaste M, *et al*. Predicting outcome of IV thrombolysis-Treated ischemic stroke patients: The DRAGON score. Neurology 2012;78:427–32.
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. Flint A, Faigeles B, Cullen S, 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.
12. Park TH, Park SS, Ko Y, Lee SJ, Lee KB, Lee J, *et al*. The iScore predicts clinical response to tissue plasminogen activator in Korean stroke patients. J Stroke Cerebrovasc Dis 2014;23:367–73.
13. Asuzu D, Nyström K, Sreekrishnan A, Schindler J, Wira C, Greer D, *et al*. TURN Score Predicts 24-hour cerebral edema after IV thrombolysis. Neurocrit Care 2016;24:381-8.
14. Lokeskrawee T, Muengtaweepongsa S, Patumanond J, Tiamkao S, Thamangraksat T, Phankhian P, *et al*. Prediction of symptomatic intracranial hemorrhage after intravenous thrombolysis in acute ischemic stroke: The Symptomatic intracranial hemorrhage score. J Stroke Cerebrovasc Dis 2017;26:2622–9.
15. Moons KGM, Altman DG, Reitsma JB, Ioannidis JPA, Macaskill P, Steyerberg EW, *et al*. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): Explanation and elaboration. Ann Intern Med 2015;162:W1–73.
16. Larrue V, Von Kummer R, Müller A, Bluhmki E. Risk factors for severe hemorrhagic transformation in ischemic stroke patients treated with recombinant tissue plasminogen activator. A secondary analysis of the European-Australasian Acute Stroke Study (ECASS II). Stroke 2001;32:438–41.
17. Hajian-Tilaki K. Sample size estimation in diagnostic test studies of biomedical informatics. J Biomed Inform 2014;48:193–204.
18. Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. J Thorac Oncol 2010;5:1315–6.
19. Promkumtan P. External validation of the symptomatic intracranial hemorrhage score predicts symptomatic intracranial hemorrhage after intravenous thrombolysis in acute ischemic stroke patients at Nakornping Hospital. J Nakornping Hosp 2018;9:13–22.
20. Sullivan LM, Massaro JM, D'Agostino RB. Presentation of multivariate data for clinical use: The Framingham Study risk score functions. Stat Med 2004;23:1631–60.
21. Landis JR, Koch GG. The Measurement of observer agreement for categorical data. Biometrics 1977;33:159–74.
22. 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.
23. Whiteley WN, Emberson J, Lees KR, Blackwell L, Albers G, Bluhmki E, *et al*. Risk of intracerebral haemorrhage with alteplase after acute ischaemic stroke: A secondary analysis of an individual patient data meta-analysis. Lancet Neurol 2016;15:925–33.
24. Liu M, Pan Y, Zhou L, Wang Y. Predictors of post-thrombolysis symptomatic intracranial hemorrhage in Chinese patients with acute ischemic stroke. PLoS One 2017;12:e0184646.
25. Tanne D, Kasner SE, Demchuk AM, Koren-Morag N, Hanson S, Grond M, *et al*. Markers of increased risk of intracerebral hemorrhage after intravenous recombinant tissue plasminogen activator therapy for acute ischemic stroke in clinical practice: The multicenter rt-PA acute stroke survey. Circulation 2002;105:1679–85.
26. Tsivgoulis G, Frey JL, Flaster M, Sharma VK, Lao AY, Hoover SL, *et al*. Pre-tissue plasminogen activator blood pressure levels and risk of symptomatic intracerebral hemorrhage. Stroke 2009;40:3631–4.
27. Asuzu D, Nystrom K, Amin H, Schindler J, Wira C, Greer D, *et al*. Comparison of 8 scores for predicting symptomatic intracerebral hemorrhage after IV thrombolysis. Neurocrit Care 2015;22:229–33.
28. Strbian D, Michel P, Seiffge DJ, Saver JL, Numminen H, Meretoja A, *et al*. Symptomatic intracranial hemorrhage after stroke thrombolysis: Comparison of prediction scores. Stroke 2014;45:752–8.
29. Sung SF, Chen SCC, Lin HJ, Chen TW, Tseng MC, Chen Chih-Hung. Comparison of risk-scoring systems in predicting symptomatic intracerebral hemorrhage after intravenous thrombolysis. Stroke 2013;44:1561–6.
30. Li M, Wang-Qin RQ, Wang YL, Bin LL, Pan YS, Liao XL, *et al*. Symptomatic intracerebral hemorrhage after intravenous thrombolysis in Chinese patients: Comparison of prediction models. J Stroke Cerebrovasc Dis 2015;24:1235–43.
31. Kunadison S, Nganwai R. Comparison of accuracy among post-thrombolytic symptomatic intracerebral hemorrhage predictive scores in ischemic stroke patients in Khon Kaen Hospital. J Thai Stroke Soc 2020;19:17–30.
32. Gumbinger C, Gruschka P, Böttinger M, Heerlein K, Barrows R, Hacke W, *et al*. Improved prediction of poor outcome after thrombolysis using conservative definitions of symptomatic hemorrhage. Stroke 2012;43:240–2.
33. Sabarudin A, Subramaniam C, Sun Z. Cerebral CT angiography and CT perfusion in acute stroke detection: A systematic review of diagnostic value. Quant Imaging Med Surg 2014;4:282–90.
34. Lorwanich P, Limsuriyakan W. Factor related intracerebral hemorrhage of patients with acute ischemic stroke who receive rtPA at Ayutthaya Hospital. TUH Journal Online 2021;6:12-20.
35. Mehta RH, Cox M, Smith EE, Xian Y, Bhatt DL, Fonarow GC, *et al*. Race/ethnic differences in the risk of hemorrhagic complications among patients with ischemic stroke receiving thrombolytic therapy. Stroke 2014;45:2263–9.
36. Maïer B, Desilles JP, Mazighi M. Intracranial hemorrhage after reperfusion therapies in acute ischemic stroke patients. Front Neurol 2020;11:599908.

Supplementary S1: Full details of external validation, model updating, and internal validation process

We conducted a systematic search in the Embase, PubMed, and Google Scholar databases using the following search terms: symptomatic intracranial hemorrhage, intravenous thrombolysis, predictive risk score, and alteplase. Our search yielded 11 predictive scoring systems, and eight of them were selected for further external validation based on the availability of each component in our setting. We found 11 published models for predicting sICH, including MSS, HAT score, SEDAN score, SITS-SICH risk score, GRASPS score, SPAN-100 index, THRIVE score, SICH score, DRAGON score, TURN score, and iScore. The last three models, however, incorporated variables including the pre-stroke modified Rankin scale (mRS), and smoking history, which were not regularly gathered in our setting. Therefore, only eight models were externally validated.

To externally validate the models, we first calculated the probability of sICH for individual participants using the assigned scores from the original models. Second, we created a calibration plot for each model in which the observed probability (Y-axis) was plotted against the predicted probability (X-axis). In accordance with the predicted probability deciles, participants were divided into 10 categories. The 10 groups were connected by a smooth line made using the lowess smoothing method, and the performance parameters of each model were calculated and compared. These parameters included a discrimination parameter, such as the area under the receiver operating characteristic (AuROC) curve (which is identical to the C-index), and calibration parameters like the expected to observed outcomes (E/O) ratio, calibration-in-the-large (CITL), and calibration slope (C-slope). Detailed explanations of each parameter are available elsewhere[1].

We decided not to perform the Hosmer–Lemeshow test for model calibration due to criticisms regarding its limited power to detect poor calibration and sensitivity to grouping and sample size[1]. In addition, using a predetermined cut-off point for identifying patients at high risk of developing sICH, we determined each scoring system's sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio (LR+), negative likelihood ratio (LR–), and correct classification rate (CCR).

Model updating, further internal validation, and creating a points-based system

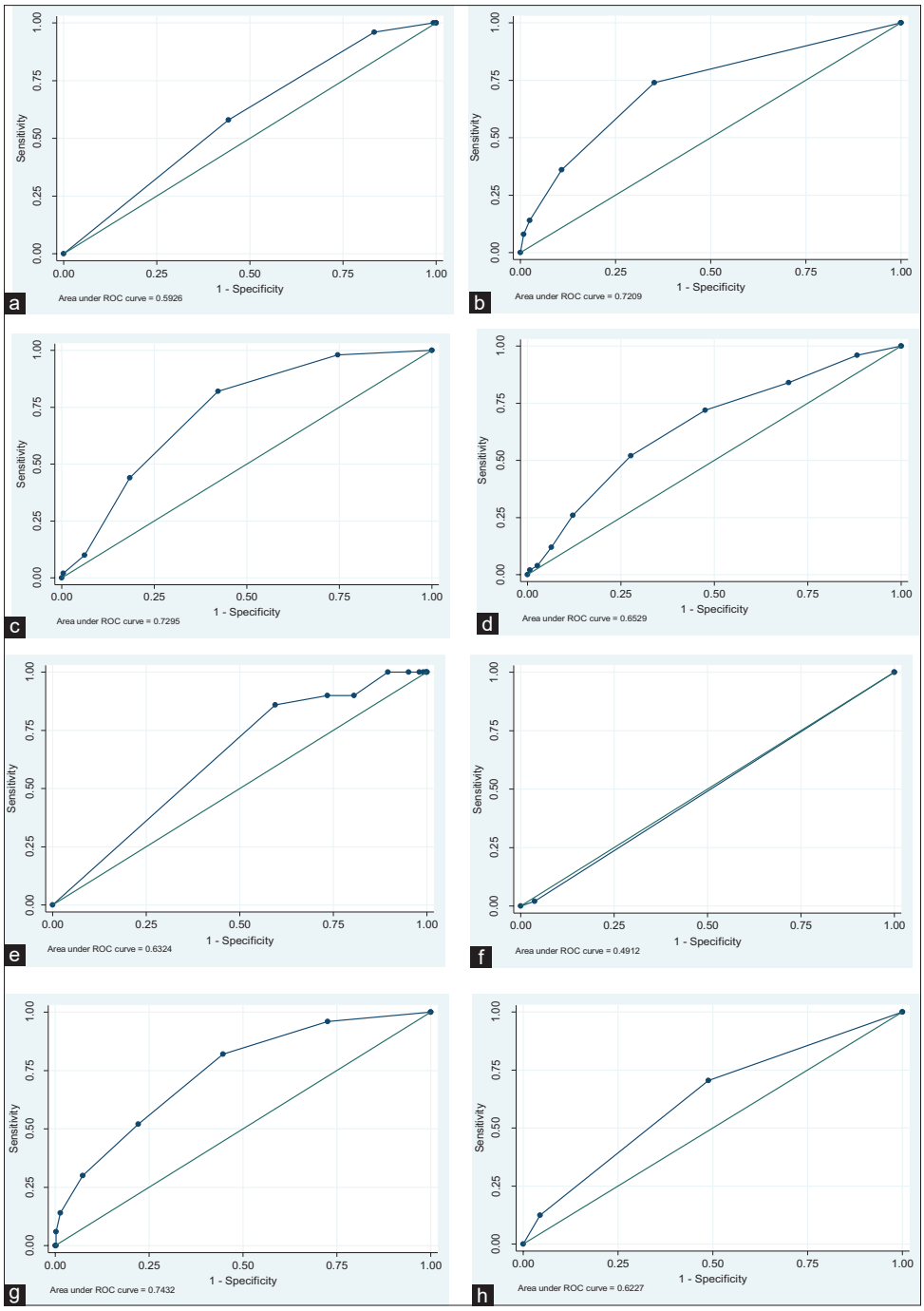
Based on the model's performance during external validation, we selected the model with the best discrimination in the groups, regardless of its calibration performance, for further updating and internal validation. The justification for selecting the model based on its discrimination parameter was that in comparison to calibration parameters, it was insensitive to systematic errors in calibration, which, in turn, make it less affected by updating (refitting) the model[2]. As a result, after sufficient updating, we ensured that we had the model that performed the best in terms of both calibration and discrimination. Furthermore, it was independent of the applied cut-off threshold.

Given that the selected model showed poor calibration, further model updating was necessary. Using binary multivariable logistic regression, we re-estimated the model's intercept and the beta-coefficients of each predictor in the same cohort used for external validation. As a result, each predictor was given a new beta-coefficient in the revised model, which kept the same predictors as the original model.

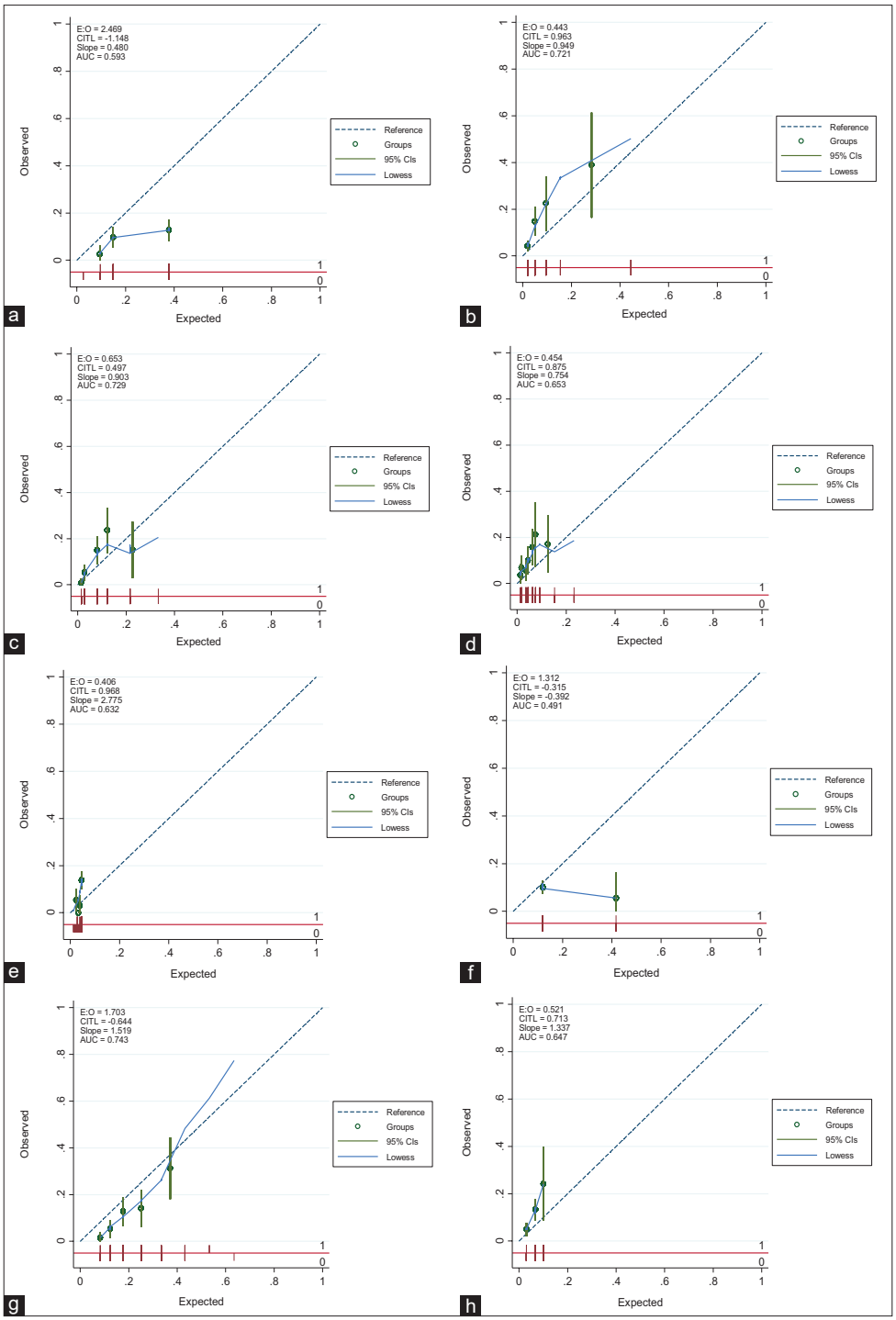
To assess the potential optimism of the updated model, we conducted an internal validation using a bootstrapping method, according to the transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) guideline[1]. We performed 500 cycles of bootstrapping and calculated the average of the bootstrap performance. The optimism-corrected estimate of performance was then determined as the difference between the bootstrap (bootstrap sample) and test (original sample) performance in terms of AuROC, CITL, and C-slope. The model was further adjusted for optimism by subtracting a linear predictor with the original intercept and then multiplying the new linear predictor (without intercept) with a bootstrapped calibration slope. The model's new intercept was estimated by forcing the linear predictor as an offset.

References

1. Moons KGM, Altman DG, Reitsma JB, Ioannidis JPA, Macaskill P, Steyerberg EW, *et al*. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): Explanation and elaboration. Ann Intern Med. 2015;162 (1):W1–73.
2. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, *et al*. Assessing the performance of prediction models: A framework for traditional and novel measures. Epidemiology. 2010;21 (1):128–38.



Supplementary Figure S1: Area under the receiver operating characteristic curve (AuROC) curve of each score: (a) Post-thrombolysis Risk Score (MSS); (b) Hemorrhage After Thrombolysis (HAT) score; (c) blood Sugar, Early infarct signs and hyperDense cerebral artery sign, Age, and NIHSS (SEDAN) score; (d) Safe Implementation of Treatment in Stroke – Symptomatic Intracranial Hemorrhage (SITS-SICH) risk score; (e) Glucose Race Age Sex Pressure Stroke severity (GRASPS) score; (f) Stroke Prognostication using Age and NIH stroke scale (SPAN-100) index; (g) Symptomatic Intracranial Hemorrhage (SICH) score; (h) Totaled Health Risks in Vascular Events (THRIVE) score



Supplementary Figure S2: a) Post-thrombolysis Risk Score- Multicenter Stroke Survey (MSS) (b) Hemorrhage After Thrombolysis (HAT) score (c) blood Sugar, Early infarct signs and hyperDense cerebral artery sign, Age, and NIHSS (SEDAN) score (d) Safe Implementation of Treatment in Stroke – Symptomatic Intracranial Hemorrhage (SITS-SICH) risk score (e) Glucose Race Age Sex Pressure Stroke severity (GRASPS) score (f) Stroke Prognostication using Age and NIH stroke scale (SPAN-100) index (g) Symptomatic Intracranial Hemorrhage (SICH) score (h) Totaled Health Risks in Vascular Events (THRIVE) score

Supplementary Table S1: TRIPOD checklist

Section/topic	Item	Checklist	Page
		Title and abstract	
Title	1	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted	1
Abstract	2	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcomes, statistical analysis, results, and conclusions	1
		Introduction	
Background and objectives	3a	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models	3
	3b	Specify the objectives, including whether the study describes development or validation of the model, or both	3
		Methods	
Source of data	4a	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable	4
	4b	Specify the key study dates, including start of accrual, end of accrual, and, if applicable, end of follow-up	4
Participants	5a	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centers	4
	5b	Describe the eligibility criteria for participants	4
	5c	Give details of treatments received, if relevant	4
Outcome	6a	Cleary define the outcome that is predicted by the prediction model, including how and when it is assessed	5
	6b	Report any actions to blind assessment of the outcome to be predicted	-
Predictions	7a	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured	4-5
	7b	Report any actions to blind assessment of predictors for the outcome and other predictors	-
Sample size	8	Explain how the study size was arrived at	5
Missing data	9	Describe how missing data were handled (e.g., complete case analysis, single imputation, multiple imputation) with details of any imputation methods	4
Statistical analysis methods	10a	Describe how predictions were handled in the analyses	5–6
	10b	Specify the type of model, all model-building procedures (including any predictor selection), and method for internal validation	5–6
	10d	Specify all measures used to assess model performance and, if relevant, to compare multiple models	5–6
Risk groups	11	Provide details on how risk groups were created, if done	6–7
		Results	
Participants	13a	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful	Figure 1
	13b	Describe the characteristics of the participants (basic demographics, clinical features, available predictions), including the number of participants with missing data for predictors and outcome	7–8, Table 1
Model development	14a	Specify the number of participants and outcome events in each analysis	7–8, Table 1
	14b	If done, report the unadjusted association between each candidate prediction and outcome	-
Model specification	15a	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point)	8–10, Tables 2, 3
	15b	Explain how to the use the prediction model	9–10, Figure 2
Model performance	16	Report performance measures for the prediction model	9, Figure 3, Table 3
		Discussion	
Limitations	18	Discuss any limitations of the study (such as nonrepresentative samples, few events per predictor, missing data)	14–15

Contd...

Supplementary Table S1: Contd...

Section/topic	Item	Checklist	Page
		Discussion	
Interpretation	19b	Give an overall interpretation of the results, considering objectives, limitations, and results from similar studies and other relevant evidence	11–15
Implications	20	Discuss the potential clinical use of the model and implications for future research	15
		Other information	
Supplementary information	21	Provide information about the availability of supplementary resources such as the study protocol, web calculator, and data sets	Supp.
Funding	22	Give the source of funding and the role of the funders for the present study	16

Supplementary Table S2: Summary of variables included in the different predictive sICH scoring systems

Variable	MSS	HAT	SEDAN	SITS-SICH	GRASPS	SPAN	THRIVE	SICH
Sex: male					✓			
Race: Asian					✓			
Age	✓		✓	✓	✓	✓	✓	
Weight				✓				
Underlying disease								
Hypertension							✓	✓
Diabetes							✓	
Atrial fibrillation							✓	
Valvular heart disease								
Comedication								
Aspirin				✓				✓
Clopidogrel				✓				
NIHSS	✓	✓	✓	✓	✓	✓	✓	✓
SBP	✓		✓	✓	✓			✓
Laboratory test								
Blood glucose	✓	✓	✓	✓	✓			✓
Platelet	✓							✓
Onset-to-treatment time				✓				
CT brain		✓	✓					
IV antihypertensive drug								✓

CT=computed tomography, GRASPS=Glucose Race Age Sex Pressure Stroke severity score, HAT=Hemorrhage After Thrombolysis score, IV=intravenous, MSS= Multicenter Stroke Survey, NIHSS=National Institutes of Health Stroke Scale, SBP=systolic blood pressure, SEDAN=blood Sugar, Early infarct signs and hyperDense cerebral artery sign, Age, and NIHSS score, sICH=symptomatic intracranial hemorrhage, SICH=Symptomatic Intracranial Hemorrhage score, SITS-SICH=Safe Implementation of Treatment in Stroke – Symptomatic Intracranial Hemorrhage risk score, SPAN=Stroke Prognostication using Age and NIH stroke scale, THRIVE=Totaled Health Risks In Vascular Events score

Supplementary Table S3: Performance statistics in original (updated SICH) and optimism-adjusted (optimism-adjusted updated SICH) models

Statistics	Performance		
	Original apparent	Test	Optimism adjusted
AuROC	0.794	0.768	-0.026
Calibration-in-the-large	0.000	0.007	0.007
Calibration slope	1.000	0.868	-0.132

Optimism-adjusted values were derived from test performance – original apparent performance

Supplementary Table S4: Agreement between group risk based on a logistic regression model and a points-based system

Logistic model	Points-based system				
	Very low	Low	Moderate	High	Total
Very low	195	10	0	0	205
Low	0	240	48	0	288
Moderate	0	0	7	2	9
High	0	0	0	0	0
Total	195	250	55	2	502