Derivation and validation of a novel and parsimonious prognosis prediction score for ischemic stroke patients: The S2AFI Score

Xiaochuan Liu (xiaochuanliu05@outlook.com)  
Wannan Medical College

Yunhe Luo  
Fudan University

Yang Liu  
Fudan University

Xiaofan Li  
Fudan University

Zhimin Yan  
Fudan University

Jing Zhao  
Fudan University

Research article

Keywords: ischemic stroke, prognosis prediction, prediction model, stroke scale

DOI: https://doi.org/10.21203/rs.3.rs-22750/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Background To develop and validate a novel and integer-based score to predict the unfavorable outcome at 3-month after ischemic stroke using a dataset from the Chinese population.

Methods We retrospectively included 394 patients presented to our stroke center within 24 hours after symptom onset. 70% of them were randomly selected as the model derivation dataset and the remaining 30% as the model validation dataset. Step-wise logistic regression were applied to identify the strong predictors of the unfavorable outcome, which was defined as modified Rankin Scale (mRS) >2.

Results We identified 4 strong predictors in the final prediction model: stroke severity, age, serum fibrinogen level and intravenous thrombolytic therapy. 2 points were assigned when patients with NIH stroke scale higher than 6 at admission. 1 point was assigned to patients older than 71 years old and serum fibrinogen level higher than 2.98g/L respectively. For patients miss the opportunity of intravenous thrombolysis therapy, 1 point was assigned. We called our score system as S2AFI score and it ranges from 0 to 5. The Youden index of the S2AFI score was 2 which means patients with 3 points or higher tend to have unfavorable outcome in the future. The discriminative ability of the S2AFI score as dichotomized predictor achieved an area under the curve of 0.72 (95% Confidence Internal 0.66-0.79) in the derivation dataset, 0.72 (95% Confidence Internal 0.61-0.83) in the validation dataset and 0.68 (95% Confidence Internal 0.63-0.74) in all patients. While for the S2AFI score as continuous predictor, the score achieved an area under the curve of 0.76 (95% Confidence Internal 0.69-0.83) in the derivation dataset, 0.74 (95% Confidence Internal 0.62-0.87) in the validation dataset and 0.74 (95% Confidence Internal 0.68-0.80) in all patients. There was a strong correlation between the S2AFI score and the risk of unfavorable outcome at 3-month (P for trend <0.001).

Conclusions Our novel and parsimonious S2AFI score had good performance in the prediction of unfavorable outcome at 3-month after ischemic stroke.

Background

Stroke has been ranked the leading cause of death and acquired disability in recent years for Chinese people and it is also associated with 43.7 million annual disability-adjusted life-years worldwide[1-3]. The situation is expected to be even worse in consideration of the population ageing. Ischemic stroke accounts for almost 80% of all strokes. As the advanced treatment like the intravenous thrombolytic therapy and the endovascular thrombectomy being applied increasingly in daily practice, the mortality rate of stroke patients has the inclination to be stable for last decades[3]. Unfortunately, not all the ischemic stroke patients are qualified of receiving those therapies. Patients dead in the hospital or with severe disability after ischemic stroke are unacceptable for their family members. Answering the questions asked by patients or their family members regarding the prognosis is quite hard when stroke physicians based only on their experiences. There are a considerable number of prognosis prediction models developed to assist decision making for neurologists in clinical practice in recent years, while few
of them are available for Chinese population[4-11]. The CoRisk score facilitate their application through online calculator, which the website development costs are quite high[4]. A prediction tool needs to be parsimonious like the ABCD2 score and with good predictive ability at the same time[12]. In this study, we aimed to develop and validate a novel and integer-based parsimonious prognosis prediction score using dataset from the Chinese population to assist stroke physicians in clinical practice.

**Methods**

Study design and population

This prediction score developed in accordance with the guidelines of the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis[13]. The study protocol was approved by the Institutional Review Board of Minhang Hospital affiliated to Fudan University. Written informed consent was obtained from all patients or their welfare guardians for data collection and subsequent analysis. The data that support the findings of this study could be available from the corresponding author under reasonable request.

This was a retrospective study using prospectively collected hospital-based dataset. We included all the ischemic stroke patients, confirmed by Computer Tomography (CT) or Magnetic Resonance Imaging (MRI), admitted to our stroke center within 24 hours after symptom onset from January, 2018 to July, 2019. 70% of them were randomly selected as the model development dataset and the remaining 30% of them as the model validation dataset. As for patients with unknown symptom onset time, we defined the time last known well as the onset time. Patients with hemorrhage stroke, transient ischemic stroke or pre-stroke disability were excluded from this study. The unfavorable outcome at 3-month was defined as modified Rankin Scale (mRS) >2 and it was evaluated through out-patient routine visit or structured telephone interview by experienced neurologists or trained nurses who were blind to the patients’ archives. Patients lost follow-up or with severe complications during hospitalization were also excluded. Ischemic stroke subtype was classified according to the Trial of Org 10172 in Acute Stroke Treatment (TOAST) criteria by the experienced neurologists at our center[14].

Predictors included in the multivariate logistic regression

Stroke severity was assessed on admission with the NIH Stroke Scale (NIHSS). Risk factors like alcohol consumption and smoking habit were defined as binary variable (0 No or 1 Yes). We gave 1 (Yes) to the patients who had those habits no matter whether they have quit it or not. Other risk factors like hypertension and atrial fibrillation (AF) were defined according to standard clinical criteria and confirmed within the 24 hours after admission. The blood sample was measured within 24 hours after admission. All laboratory tests (White Blood Cell, Low Density Lipoprotein, Homocysteine, Glycated hemoglobin, Creatinine, and Fibrinogen) were treated as binary variables in our model. Recanalization therapy including intravenous thrombolytic therapy (IVT) and endovascular thrombectomy (EVT) was also treated as binary variables (0 Yes 1 No).
Strategies for developing a novel and parsimonious prediction score

In order to facilitate the use of our prediction model, we transformed our model into a scoring system. All the continuous predictors were dichotomized into binary variables by calculating their Youden index. The optimal cut-point of the variable age was 71 years old, stroke severity was 6, White Blood Cell (WBC) was $6.83 \times 10^9/L$, Low Density Lipoprotein (LDL) was 2.93mmol/L, Homocysteine (HCY) was 13umol/L, Glycated hemoglobin (GH) was 6.2%, Creatinine was 81mmol/L, and Fibrinogen was 2.98g/L respectively. These dichotomized variables were put into the multivariate logistic regression. Forward and backward step-wise logistic regression were applied by using the likelihood ratio test with Akaike's information criterion (AIC) as the stopping rule. The P value of the selection was set at <0.1 and <0.2. The model with the smallest AIC value, indicating the best predictive ability, was our target model. The β-coefficients of the predictors in the target model was rounded to its nearest integer as the score of each predictor. The $S_2$AFI score was generated by calculating the sum of each score. The optimal threshold of the novel score was determined from receiver operating characteristic (ROC) curve analysis.

Statistical analysis

Since we dichotomized all the continuous variables, Fisher exact test was used to compare the differences between the groups. The analysis of the variables as continuous variable was shown in online supplement material.

All the predictors included in the multivariate logistic regression analysis were confirmed with no strong collinearity (variation inflation factors <2). The Youden index of those predictors was calculated by Liu Method[15]. The performance of the novel prediction score was tested with the ROC curve analysis. After the prediction score developed, we validated its performances in two datasets: one is the 30% validation dataset and the other is all the patients included in this study. The area under the curve (AUC) of the score as a continuous predictor and dichotomized predictor in these two datasets was calculated respectively. The calibration of the score was assessed for goodness of fit by plotting the estimated probability on the x-axis against the observed probability along the y-axis compared with the diagonal line, representing perfect calibration. P for trend test was applied to test whether the higher score is related to the higher risk of achieving an unfavorable outcome at 3-month. The statistical analysis was performed on STATA (Version 15.0 Stata Corp College Station, Texas, USA) and R software (R version 3.5.3 The R Foundation for Statistical Computing). Two tailed P value less than 0.05 was considered statistically significant.

Results

429 patients arrived at our stroke center within 24 hours after symptom onset during January, 2018 to July, 2019. There were 31 patients excluded from this study because of the discharge diagnosis is not ischemic stroke. Patients with cancer (1) and pre-stroke disability (2) were excluded. 1 patient died during hospitalization was also excluded. Finally, 394 patients were included in this study. Of them, 21.1% (83) had an unfavorable outcome at 3-month. The mean age of all the patients was 68.61 years.
In brief, the univariate and multivariate analysis of all the predictors in our dataset were shown in online supplement material. The $\beta$-coefficients of our novel prognosis prediction model were shown in Table 1. 2 points were assigned to the stroke severity when patients with NIHSS higher than 6 at admission. 1 point was assigned to patients older than 71 years old and serum fibrinogen level higher than 2.98g/L respectively. For patients miss the opportunity of intravenous thrombolysis therapy, 1 point was assigned. We called our score system as $S_2$AFI score and it ranges from 0 to 5. The performances of the $S_2$AFI score to predict the unfavorable outcome at 3-month were shown in Table 2. The Youden index of the $S_2$AFI score was 2 which means patients with 3 points or higher tend to have unfavorable outcome in the future. The discriminative ability of the $S_2$AFI score as dichotomized predictor achieved an area under the curve of 0.72 (95% Confidence Interval 0.66-0.79) in the derivation dataset, 0.72 (95% Confidence Interval 0.61-0.83) in the validation dataset and 0.68 (95% Confidence Interval 0.63-0.74) in all patients (Figure 1). While for the $S_2$AFI score as continuous predictor, the score achieved an area under the curve of 0.76 (95% Confidence Interval 0.69-0.83) in the derivation dataset, 0.74 (95% Confidence Interval 0.62-0.87) in the validation dataset and 0.74 (95% Confidence Interval 0.68-0.80) in all patients (Figure 2). The calibration plot was shown next to their ROC curve analysis. There was a strong correlation between the $S_2$AFI score and the risk of unfavorable outcome at 3-month as shown in Figure 3 (P for trend <0.001). Higher scores augmented the probability of achieving unfavorable outcome at 3-month.

**Discussion**

Our study showed the novel integer-based $S_2$AFI score achieved within 24 hours after admission had good performance in the prediction of unfavorable outcome at 3-month after acute ischemic stroke. Early identification of those patients who at high risk of death or severe disability becomes more and more important in recent years. For one reason, patients or their family members wouldn't understand the cause of the post-stroke disability after spending plenty of money and this would aggravate the doctor-patient intense relationship in China, so early communication between them for those patients at high risk becomes quite critical. Our 5-score prediction tool had the potential to be helpful in explaining the risk of the unfavorable outcome in the future to the patients or their family members. Another reason is that the importance of early rehabilitation for those patients at high risk of post-stroke disability during the hospitalization. Since there is an upsurge in the number of young stroke patients in the past decade[16], post-stroke disability may cause not only the loss of huge amount of money but also the loss of the hope of bright future for them. Sometimes, the decision regarding implementing early rehabilitation based only on stroke physicians’ experiences was not enough, a lean, easy-to-calculate, without human-biased $S_2$AFI score had the potential to assist them.

There have been a lot of prognosis prediction models published in the leading clinical journals for ischemic stroke patients such as the ASTRAL score, iScore and the CoRisk score[4-6]. In 2017, Dr Quinn compared the performances of 8 prognosis scales using a large, independent, clinical trials dataset and his team found that the ASTRAL score showed the superior utility in comparison of the other prediction models[7]. The ASTRAL score is also an integer-based score like our $S_2$AFI score, consisting of six items,
developed from the dataset in Switzerland. In 2013, Dr. Wang validated that the ASTRAL score is a reliable tool to predict long-term unfavorable outcome after ischemic stroke in the Chinese population[17]. Not like the ABCD2 score being widely used in the prediction of the risk of recurrent stroke, even with this good performance, the ASTRAL score still wasn’t included in stroke guidelines. In the clinical practice, it is always debatable whether the performance perceived as statistically acceptable are also clinically acceptable. The Youden index of the S$_2$AFI Score was 2 which means patients with 3 points or higher tend to have unfavorable outcome at 3-month, while it’s worth noting the specificity at this cut-point was insufficient, indicating that almost one of two patients would be wrongly assigned a poor prognosis. It’s acceptable when we utilize the score as a convenient tool to inform patients or their family members the risk of poor prognosis and the importance of early rehabilitation, while it’s not recommended for introducing the high-cost medicine based on the cut-point of the score. Since the performance of the score as continuous predictor was superior to the dichotomized predictor, the application of the score could be quite flexible.

Our novel and parsimonious S$_2$AFI score included four strong predictors, which are all easily available within 24 hours after admission, ranging from 0 to 5. We didn’t include the past medical history as potential predictors in the score development process as in clinical practice, many patients didn’t know whether they have it or not and sometimes the answers from the patients or their family members are contrary to the test results. We confirmed that all the potential predictors included in the study in the first 24 hours after admission. Including thrombolytic therapy in our S$_2$AFI score accounts for the progress made in the last 2 decades and it also promotes the concept of fast reaction when stroke attacks. As the hometown of Stroke 1-2-0 campaign[18], this score had the potential to be applied in the stroke knowledge promotion. For patients presented late to the stroke center and miss the opportunity of administering the thrombolytic therapy, the S$_2$AFI score would be higher than those patients acted immediately after symptom onset. Fibrinogen as a serum marker has been found to be associated with not only the development of stroke but also the outcome after stroke[19-21]. Early fibrinogen depletion after reperfusion therapy including IVT and EVT was also related to symptomatic intracranial hemorrhage[22]. We identified 2.98g/L as the optimal cut-point for the prediction of the unfavorable outcome after ischemic stroke which was consistent with the study published in 2009, showing that there were a considerable number of patients had an unfavorable outcome with the serum fibrinogen level lower than 3.0g/L[20]. Age and stroke severity as two strong predictors have been included in many prognosis prediction models, these two predictors alone were also the leanest prediction models[10, 23].

**Limitations**

The limitations of this study were that this was a single-center retrospective study and the number of the patients included in the development of the S$_2$AFI score was quite small. The low amount of the patients dead at 3-month also limited us to develop a model predicting the risk of death after the ischemic stroke. Although we developed the prediction model using four commonly available clinical variables at the early stage of ischemic stroke, the inclusion of neuroimaging predictors like the site of the occlusion and the
infarct size may improve the performances of our model. In order to get a parsimonious prognosis prediction score, we dichotomized the continuous predictor into binary predictor and this may sacrifice part of their predictive ability. Lacking external validation of our $S_2AFI$ score in a large multicenter dataset was also a flaw of this study, so multicenter collaboration is quite crucial in the future. An externally validated prediction model could also have the potential to be used in the evaluation of stroke care among different stroke centers in our country. Finally, further prospective studies assessing the prognostic accuracy of the $S_2AFI$ score was warranted in the future.

**Conclusion**

Our novel and parsimonious $S_2AFI$ score had good performance in the prediction of unfavorable outcome at 3-month after ischemic stroke. The score had the potential to be utilized in clinical practice as a convenient tool to inform patients or their family members the risk of poor prognosis and the importance of early rehabilitation after ischemic stroke.

**List Of Abbreviations**

NIHSS: NIH stroke scale; mRS: modified Rankin Scale; CI: Confidence Internal; CT: Computer Tomography; MRI: Magnetic Resonance Imaging; ROC: receiver operating characteristic; AUC: area under the curve; AF, atrial fibrillation; SD, standard difference; IQR, inter quartile range; WBC, white blood cell count; LDL, low density lipoprotein; HCY, homocysteine; NIHSS, National Institute of Health Stroke Scale; IVT, intravenous thrombolytic therapy; EVT, endovascular thrombectomy; LAA, Large-artery atherosclerosis; CE, Cardioembolism; SVO, Small-vessel occlusion;

**Declarations**

**Ethics approval and consent to participate**

The study protocol was approved by the Institutional Review Board of Minhang Hospital affiliated to Fudan University. Written informed consent was obtained from all patients or their welfare guardians for data collection and subsequent analysis.

**Consent for publication**

Not applicable.

**Availability of data and materials**

The data that support the findings of this study could be available from the corresponding author under reasonable request.

**Funding**
Acknowledgment

We thank all the doctors and nurses who contributed in the collection of the dataset.

Competing interests

None.

Authors' contributions

Xiaochuan Liu: Data acquisition, analysis and manuscript preparation

Yunhe Luo: Data acquisition, analysis and manuscript preparation

Yang Liu: Data acquisition, and manuscript preparation

Xiaofan Li: Data analysis, Data acquisition

Zhimin Yan: Study concept and design and manuscript preparation and revision

Jing Zhao: Study concept and design and manuscript preparation and revision

References

1. Liu S, Li Y, Zeng X, Wang H, Yin P, Wang L, et al. Burden of Cardiovascular Diseases in China, 1990-2016: Findings From the 2016 Global Burden of Disease Study. JAMA Cardiol. 2019. Epub 2019/03/14. doi: 10.1001/jamacardio.2019.0295. PubMed PMID: 30865215; PubMed Central PMCID: PMCPMC6484795.

2. Wang W, Jiang B, Sun H, Ru X, Sun D, Wang L, et al. Prevalence, Incidence, and Mortality of Stroke in China: Results from a Nationwide Population-Based Survey of 480 687 Adults. Circulation. 2017;135(8):759-71. Epub 2017/01/06. doi: 10.1161/CIRCULATIONAHA.116.025250. PubMed PMID: 28052979.

3. Wu S, Wu B, Liu M, Chen Z, Wang W, Anderson CS, et al. Stroke in China: advances and challenges in epidemiology, prevention, and management. Lancet Neurol. 2019;18(4):394-405. Epub 2019/03/18. doi: 10.1016/S1474-4422(18)30500-3. PubMed PMID: 30878104.

4. De Marchis GM, Dankowski T, Konig IR, Fladt J, Fluri F, Gensicke H, et al. A novel biomarker-based prognostic score in acute ischemic stroke: The CoRisk score. Neurology. 2019;92(13):e1517-e25. Epub 2019/03/03. doi: 10.1212/WNL.0000000000007177. PubMed PMID: 30824558.

5. Ntaios G, Faouzi M, Ferrari J, Lang W, Vemmos K, Michel P. An integer-based score to predict functional outcome in acute ischemic stroke: the ASTRAL score. Neurology. 2012;78(24):1916-22. Epub 2012/06/01. doi: 10.1212/WNL.0b013e318259e221. PubMed PMID: 22649218.
6. Saposnik G, Kapral MK, Liu Y, Hall R, O’Donnell M, Raptis S, et al. IScore: a risk score to predict death early after hospitalization for an acute ischemic stroke. Circulation. 2011;123(7):739-49. Epub 2011/02/09. doi: 10.1161/CIRCULATIONAHA.110.983353. PubMed PMID: 21300951.

7. Quinn TJ, Singh S, Lees KR, Bath PM, Myint PK, Collaborators V. Validating and comparing stroke prognosis scales. Neurology. 2017;89(10):997-1002. Epub 2017/08/11. doi: 10.1212/WNL.0000000000004332. PubMed PMID: 28794250.

8. Saposnik G, Fang J, Kapral MK, Tu JV, Mamdani M, Austin P, et al. The iScore predicts effectiveness of thrombolytic therapy for acute ischemic stroke. Stroke. 2012;43(5):1315-22. Epub 2012/02/07. doi: 10.1161/STROKEAHA.111.646265. PubMed PMID: 22308252.

9. Saposnik G, Raptis S, Kapral MK, Liu Y, Tu JV, Mamdani M, et al. The iScore predicts poor functional outcomes early after hospitalization for an acute ischemic stroke. Stroke. 2011;42(12):3421-8. Epub 2011/10/01. doi: 10.1161/STROKEAHA.111.623116. PubMed PMID: 21960583.

10. Saposnik G, Guzik AK, Reeves M, Ovbiagele B, Johnston SC. Stroke Prognostication using Age and NIH Stroke Scale: SPAN-100. Neurology. 2013;80(1):21-8. Epub 2012/11/24. doi: 10.1212/WNL.0b013e31827b1ace. PubMed PMID: 23175723; PubMed Central PMCID: PMCPMC3589202.

11. Weimar C, Konig IR, Kraywinkel K, Ziegler A, Diener HC, German Stroke Study C. Age and National Institutes of Health Stroke Scale Score within 6 hours after onset are accurate predictors of outcome after cerebral ischemia: development and external validation of prognostic models. Stroke. 2004;35(1):158-62. Epub 2003/12/20. doi: 10.1161/01.STR.0000106761.94985.8B. PubMed PMID: 14684776.

12. Kelly PJ, Albers GW, Chatzikostantinou A, De Marchis GM, Ferrari J, George P, et al. Validation and comparison of imaging-based scores for prediction of early stroke risk after transient ischaemic attack: a pooled analysis of individual-patient data from cohort studies. Lancet Neurol. 2016;15(12):1238-47. Epub 2016/10/19. doi: 10.1016/S1474-4422(16)30236-8. PubMed PMID: 27751555.

13. Collins GS, Reitsma JB, Altmann DG, Moons KGM, members of the Tg. Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD): The TRIPOD Statement. Eur Urol. 2015;67(6):1142-51. Epub 2015/01/13. doi: 10.1016/j.eururo.2014.11.025. PubMed PMID: 25572824.

14. Adams HP, Jr., Bendixen BH, Kappelle LJ, Biller J, Love BB, Gordon DL, et al. Classification of subtype of acute ischemic stroke. Definitions for use in a multicenter clinical trial. TOAST. Trial of Org 10172 in Acute Stroke Treatment. Stroke. 1993;24(1):35-41. Epub 1993/01/01. doi: 10.1161/01.str.24.1.35. PubMed PMID: 7678184.

15. Liu X. Classification accuracy and cut point selection. Stat Med. 2012;31(23):2676-86. Epub 2012/02/07. doi: 10.1002/sim.4509. PubMed PMID: 22307964.

16. Ekker MS, Boot EM, Singhal AB, Tan KS, Debette S, Tuladhar AM, et al. Epidemiology, aetiology, and management of ischaemic stroke in young adults. Lancet Neurol. 2018;17(9):790-801. Epub
Tables

Table 1 The β-coefficients of our novel prognosis prediction model
|                          | Coefficient | SE   | p value | 95% CI   |
|--------------------------|-------------|------|---------|----------|
| NIHSS at admission >= 7 | 1.711       | 0.339| <0.001  | 1.048    |
| Age >= 71 years old     | 1.082       | 0.326| 0.001   | 0.443    |
| Fibrinogen >= 2.98 g/L  | 0.754       | 0.328| 0.022   | 0.111    |
| Miss the opportunity of IVT | 0.974     | 0.451| 0.031   | 0.09     |
| Intercept               | -3.486      | 0.511| <0.001  | -4.487   |

NIHSS: NIH stroke scale

Table 2 The performances of S₂AFI score to predict the unfavorable outcome at 3-month

| Cut-point | Sensitivity% | Specificity% | Correctly Classified% |
|-----------|--------------|--------------|-----------------------|
| >=0       | 100%         | 0%           | 22.58%                |
| >=1       | 96.83%       | 6.94%        | 27.24%                |
| >=2       | 85.71%       | 43.06%       | 52.69%                |
| >=3       | 66.67%       | 77.31%       | 74.91%                |
| >=4       | 42.86%       | 92.13%       | 81.00%                |
| >=5       | 17.46%       | 99.54%       | 81.00%                |

Figures
Figure 1

The discriminative ability of the S2AFI score as dichotomized predictor. A: the derivation dataset, B: the validation dataset, C: the whole dataset. The calibration plot was shown next to their ROC curve analysis.

Figure 2

The discriminative ability of the S2AFI score as continuous predictor. A: the derivation dataset, B: the validation dataset, C: the whole dataset. The calibration plot was shown next to their ROC curve analysis.
Figure 3

The positive correlation between the S2AFI score and the risk of unfavorable outcome at 3-month after ischemic stroke.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- onlinesupplementmaterial.docx
- TRIPODChecklist.docx