External Validation of START nomogram to predict 3-Month unfavorable outcome in Chinese acute stroke patients

BaiLi Song, MS,*,†,[1] XiangLiang Chen, MD,‡,[1] Dan Tang, MS,†
Mako Ibrahim, MS,*,† YuKai Liu, MD,‡ Linda Nyame, MS,*,† Teng Jiang, MD,‡
Wei Wang, MD,‡ Xiang Li, MS,*,† Chao Sun, MS,*,† Zheng Zhao, MS,*
Jie Yang, MD,§ JunShan Zhou, MD,‡ and JianJun Zou, PhD*,[1]

Background: Recently, the NIHSS STroke Scale score, Age, pre-stroke mRS score, onset-to-treatment Time (START nomogram) predicts 3-month functional outcome after intravenous thrombolysis in ischemic stroke patients. However, this model has not yet been an external validation. We aim to validate the performance of START nomogram. *Methods:* Data were derived from the stroke center of the Nanjing First Hospital (China). Patients who lacked the necessary data to calculate the nomogram and missed 3-month modified Ranking scale scores were excluded. Modified Rankin Scale score more than 2 at 3-month was assessed as an unfavorable outcome. We used areas under the receiver operator characteristic curves (AUC-ROC) to quantify the prognostic value. Calibration was assessed by calibration plots and Hosmer-Lemeshow (HL) goodness of fit test. *Result:* The final cohort included 306 eligible patients. For 3-month unfavorable outcome, the AUC-ROC of the START nomogram was .766 (95%CI: .7013-.8304, $P <$ .0001), suggesting good discrimination in the START nomogram. It also showed good calibration (HL goodness of fit test $P = .1261$) in the external validation sample. *Conclusion:* The START nomogram with good predictive performance is a reliable and simple clinical instrument to predict unfavorable outcome after acute stroke.
Key Words: Stroke—START nomogram—prognosis—outcome—thrombolysis

Introduction

Stroke is the second leading global cause of death behind ischemic heart disease.[1] Among them, acute ischemic stroke (AIS) is the most common type of stroke.[2] The treatment of choice for most patients with AIS is intravenous thrombolysis (IVT) with recombinant tissue plasminogen activator (Alteplase).[3]

For making decisions in patient care management, most clinical physicians use their clinical experience to predict outcomes of patients. The accuracy of these informal predictions is unclear.[4] Prognostic models are good tools that may inform patients with a certain condition about their future outcome or help to judge subsequent treatment possibilities.[5] Several prognostic models had been developed to predict outcome after IVT for AIS.[6-11] Only 2 developed scores were designed to predict 3-month unfavorable functional outcome after IVT for AIS, such as (hyper) Dense cerebral artery sign/early infarct signs on admission CT scan, prestroke modified Rankin Scale (mRS) score, Age, Glucose level at baseline, Onset-to-treatment time, and baseline National Institutes of Health

From the *Department of Clinical Pharmacology, Nanjing First Hospital, Nanjing Medical University, Nanjing, Jiangsu, China; †School of Basic Medicine and Clinical Pharmacy, China Pharmaceutical University, Nanjing, Jiangsu, China; ‡Department of Neurology, Nanjing First Hospital, Nanjing Medical University, Nanjing, China; and §Department of Neurology, the First Affiliated Hospital of Chengdu Medical College, Chengdu, China.

Received December 18, 2018; revision received January 24, 2019; accepted February 23, 2019.

Funding Sources Statement: This study was supported by National Natural Science Foundation of China grant 81673511, Jiangsu key Research and Development Plan grant BE2017613, Jiangsu Six Talent Peaks Project grant WSN-151, and Nanjing Medical Science and Technique Development Foundation grant QRX17020 and ZKX15027.

Address correspondence to JianJun Zou PhD, Division of Clinical Pharmacology, Nanjing First Hospital Nanjing Medical University, Nanjing, China and JunShan Zhou, MD, Department of Neurology, Nanjing First Hospital, Nanjing Medical University, Nanjing, China. E-mails: zhjsh333@126.com, zoujianjun100@126.com.

[1]BaiLi Song and XiangLiang Chen contributed equally to this work.

1052-3057/$ - see front matter

https://doi.org/10.1016/j.jstrokecerebrovasdis.2019.02.032

Stroke Scale score (DRAGON score)[8] and Age (A), severity of stroke (S) measured by admission NIH Stroke Scale score, stroke onset to admission time (T), range of visual fields (R), acute glucose (A), and level of consciousness (L) (ASTRAL score),[9] which had been validated externally by European[12,13] and Chinese population.[14] However, these scores for individualized prediction of outcome after IVT are limited by the use of dichotomization/categorization of continuous variables such as NIHSS score and age, for the disadvantage of dichotomization is that it does not make use of within-category information and lead to the loss of information.[5]

Recently, NIHSS STroke Scale score, Age, pre-stroke mRS score, onset-to-treatment Time (START) nomogram, the first nomogram developed by using a continuous score and including 4 strongest continuous predictors in IVT stroke patients, was validated internally in multicenter Italy cohorts,[15] which reliably calculates the probability of unfavorable outcome. Nomogram is actually a visualization of a complex mathematical formula that uses traditional statistical methods such as multivariable logistic regression or Cox proportional hazards analysis to calculate the continuous probability of event of interest which is entirely based on the individual's disease characteristics, without averaging or combining within a category.[16,17] Nomogram which considers patients' clinical characteristics is widely used to predict outcome of patients in cancer, surgery, and other specialties to make important treatment decisions.[18-23] Compared with previous risk scores, the nomograms were more accurate and have better performance characteristics.[16,24,25] Moreover, the START nomogram did not require the interpretation of imaging or other precise measures to predict the functional outcome. Therefore, for busy clinicians, the START nomogram was simpler and easier to predict functional outcome in routine practice.

However, there has not yet been an external validation specifically in START nomogram whether in the European or Chinese population, therefore its application and promotion are limited. Especially in the Chinese population. This is the first paper to validate START nomogram externally. The purpose of this study is to evaluate the performance of START nomogram in Chinese AIS patients receiving IVT.

Methods

Study Participants and Data Collection

Patients included in this study were from the database of the stroke center (SC) of the Nanjing First Hospital (China). The scientific use of the data obtained from the Nanjing First Hospital Stroke Registry was approved by the Ethics Committees of Nanjing First Hospital, in accord with the Helsinki declaration in accordance to the internal protocol. Patients receiving IVT for AIS were collected. Modified Rankin Scale (mRS) more than 2 at 3-month was defined as unfavorable outcome. The inclusion criteria according to the START nomogram were: baseline National Institutes of Health (NIH) Stroke Scale (NIHSS) score: 0-25, age more than 18 years, pre-mRS score: 0-2, and onset-to-treatment time: 0-270 minutes. Patients who were diagnosed with hemorrhagic strokes, lacked baseline data, or without follow-up assessment and did not meet the START standards were excluded from the analysis. A total of 3419 acute stroke patients were admitted to the SC, leaving 306 eligible patients. Demographics, clinical parameters, laboratory results, and medical history of 306 subjects receiving IVT therapy are described in Table 1**.**

Statistical Analysis

Discrimination refers to how well the model differentiates those at higher risk of having an event from those at lower risk. Calibration refers to how similar the predicted absolute risk is to the true (observed) risk in groups of patients classified in different risk strata.[26]

Table 1. *Characteristics of patients in included in the analysis (n = 306)*

Characteristic	
Age, (mean ± SD), y	68.2 ± 12.3
Age, median (IQR), y	69 (60-78)
Male, n (%)	208 (68)
Female, n (%)	98 (32)
Baseline NIHSS (mean ± SD)	7.9 ± 5.7
Baseline NIHSS, median (IQR)	6 (3-12)
prestroke mRS, n (%)	
0	257 (84)
1	20 (6.5)
2	29 (9.5)
OTT time, (mean ± SD), min	152.3 ± 57.9
OTT time, median (IQR), min	150 (105-200)
Hypertension, n (%)	206 (67.3)
Diabetes mellitus, n (%)	74 (24.2)
Hyperlipidemia, n (%)	8 (2.6)
Coronary artery disease, n (%)	55 (18)
Atrial fibrillation, n (%)	64 (20.9)
Previous stroke, n (%)	73 (23.9)
Smoking, n (%)	
Never smoker	155 (50.7)
Current smoker	115 (37.6)
Former smoker	36 (11.8)
Body weight, median (IQR), Kg	70 (60-75)
Systolic BP, median (IQR), mmHg	148 (132-163)
Diastolic BP, median (IQR), mmHg	88 (79-96)
Creatinine, median (IQR), umol/L	73 (58.5-87)
Fasting blood glucose, median (IQR), mmol/L	5.7 (4.7-7.0)
3-month mRS, n (%)	
0-2	209 (68.3)
3-6	97 (31.7)

Values are listed as number (%), mean ± SD or median (interquartile range).

Abbreviation: Baseline NIHSS, National Institutes of Health Stroke Scale at admission; OTT time, onset-to-treatment time; pre-stroke mRS, pre-stroke modified Rankin Scale.

To test the performance of the START nomogram for predicting the probability of unfavorable outcome at 3-month, we calculated area under the Receiver Operating Characteristics (AUC-ROC) curve to assess discrimination of the model, which was also referred to as a "C-statistic" for binary outcomes.[27] The C-statistic is equivalent to the probability that the measure or predicted risk of unfavorable outcome is higher in patients who had unfavorable outcome than those who had favorable outcome.[28] A C-statistic ranging from .70 to .80 indicates an adequate power of discrimination, a range of .80-.90 is considered excellent.[29] The calibration of the models can be assessed by calibration plots, which predicts probabilities against the actual observed risk. The 45-degree line represents the perfect predictions, in which predicted outcome perfectly corresponds with actual outcome. Calibration is tested with the Hosmer-Lemeshow (HL) test. A nonsignificant HL test indicates good calibration. All analyses were performed with Stata software version 13 (Stata Corp, College Station, TX).

Result

A total of 3419 patients registered in the Nanjing First Hospital Stroke Registry cohort, baseline characteristics of the selected 306 patients are shown in Table 1. Among the 306 eligible patients, 97 patients had unfavorable outcome, 32% patients were female, and the mean age of the patients were 68.2 ± 12.3 years. The mean NIHSS score on admission was 7.9 ± 5.7. Among them, history of hypertension, diabetes mellitus, coronary artery disease, atrial fibrillation, and previous stroke were 67.3%, 24.2%, 18%, 20.9%, 23.9%, respectively.

The START nomogram showed good discrimination. The model was validated in our data with AUC-ROC value of .766 (95%CI: .7013-.8304, $P < .0001$) (Fig 1). The HL test showed that the calibration of START nomogram was good ($P = .1261$). Calibration plots are shown in Figure 2. Calibration curves for estimating 3-month unfavorable outcome indicated there was no apparent departure from perfect fit, with good correspondences between predicted outcome and actual outcome.

Discussion

The present study externally validated the START nomogram in Chinese AIS patients. It suggested that the START nomogram displayed good discrimination and calibration. This study is the first external validation of START nomogram.

To the best of our knowledge, the START model was recently first prognostic model by using nomogram for individualized prediction of probability of 3-month unfavorable outcome after IVT for stroke patients. Comparing to other clinical prediction tools (risk groupings, classification and regression tree analyses, probability tables, artificial neural networks), nomogram have the highest accuracy and the best discriminating characteristics. Nomogram is a graphical statistical instrument that incorporates several continuous variables by using traditional statistical methods. Nomogram uses the continuous scales to calculate the continuous probability of a particular outcome.[16] Therefore, nomograms provide superior individualized risk estimations that facilitate modern medical decision-making.

Among previous studies, only DRAGON[8] score and ASTRAL[9] score were designed to predict 3-month

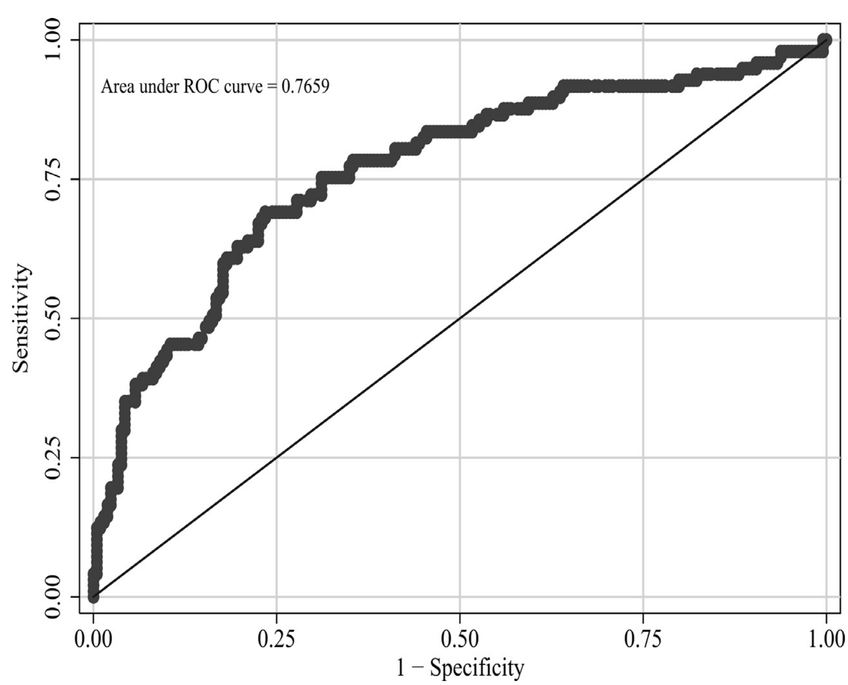

Figure. 1. *Receiver operating characteristic (ROC) curve of the START nomogram for 3-month unfavorable outcome (mRS 3-6) in 306 Chinese stroke patients.*

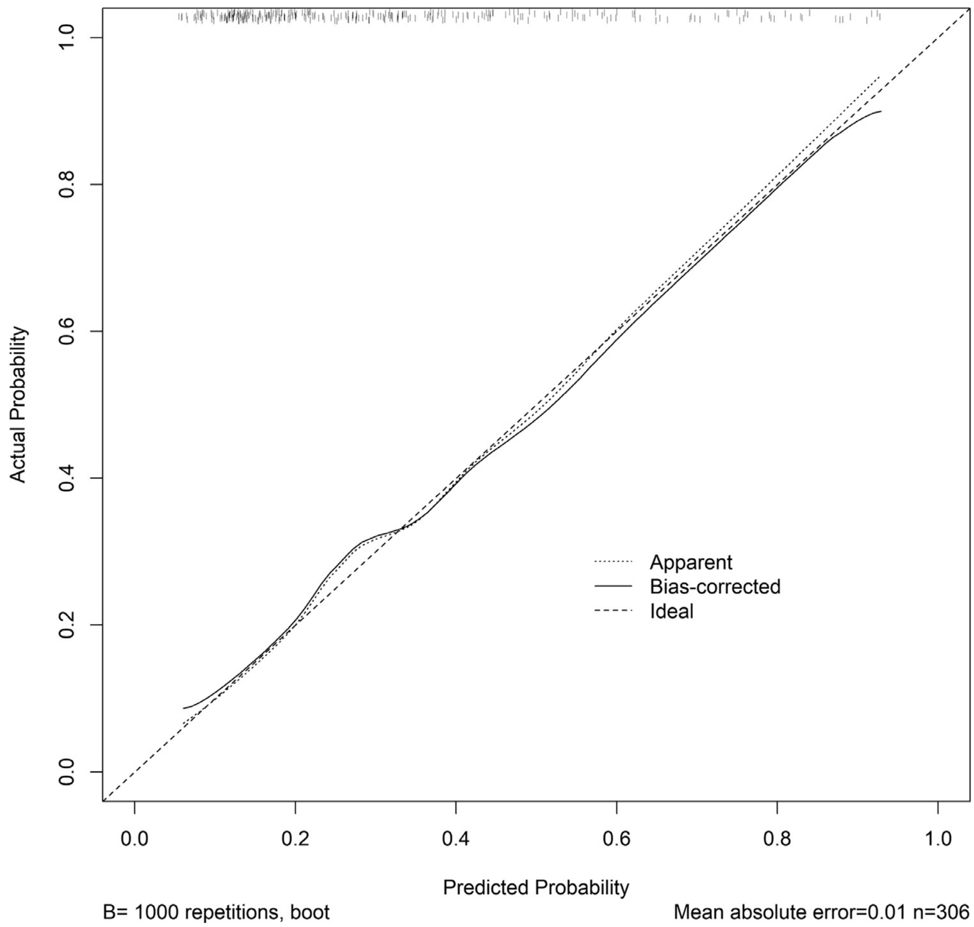

Figure. 2. *Calibration plot of START nomogram for 3-month unfavorable outcome (mRS 3-6). Dashed line is reference line where an ideal nomogram would lie. Dotted line is the performance of nomogram, while the solid line corrects for any bias in nomogram.*

unfavorable functional outcome after IVT for AIS. Compared with them, START nomogram show better performance for predicting the 3-month unfavorable outcome in AIS patients after IVT.[15] The first reason was likely that START nomogram contained the recognized factors for predicting 3-month unfavorable functional outcome such as age,[30,31] pre-mRS score,[32] NIHSS score,[33] onset-to-treatment time.[34,35] In addition, by using the continuous variables, START nomogram calculate the continuous probability of 3-month unfavorable outcome that can avoid information loss.[5] Moreover, START nomogram can be easily calculated in clinical practice, because only 4 readily obtainable clinical variables at admission construct the model. Nurses of stroke unit can easily obtain all variables. Hence, START nomogram is a good prognostic model for individualized prediction of the probability of unfavorable outcome after IVT for stroke.

However, the START nomogram has not been validated externally. A newly developed prognostic model needs to be validated with patient data not used in the development process and preferably selected from an appropriate (similar) patient population in a different center (external validation).[36,37] Therefore, its use in clinical practice is limited, external validation for the START nomogram is needed.

Our study has several limitations. First, this is a single SC study that could not represent nationwide areas and could have selection bias. Second, this is a retrospective analysis which bases on an ongoing database. The amount of missing data are large because this is the limitation of such study design. Third, the sample size is small, however, the variables of eligible patients are complete and correct.

In conclusion, our study demonstrate that the START nomogram has a good performance for predicting the 3-month unfavorable outcome after IVT for Chinese AIS patients. START nomogram is a reliable and simple clinical instrument to predict unfavorable outcome after acute stroke. Prospective studies are expected to validate the START nomogram in large sample size of Chinese populations and other population.

Conflicting Interests

The authors declare that there is no conflict of interest.

References

1. Benjamin EJ, Virani SS, Callaway CW, et al. Heart disease and stroke statistics-2018 update: a report from the American Heart Association. Circulation 2018;137:e67-e492.

2. Randolph SA. Ischemic stroke. Workplace Health Saf 2016;64:444.
3. Powers WJ, Rabinstein AA, Ackerson T, et al. 2018 guidelines for the early management of patients with acute ischemic stroke: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. Stroke 2018;49:e46-e110.
4. Jampathong N, Laopaiboon M, Rattanakanokchai S, et al. Prognostic models for complete recovery in ischemic stroke: a systematic review and meta-analysis. BMC Neurol 2018;18:26.
5. Moons KG, Altman DG, Reitsma JB, 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.
6. Saposnik G, Guzik AK, Reeves M, et al. Stroke Prognostication using Age and NIH Stroke Scale: SPAN-100. Neurology 2013;80:21-28.
7. Flint AC, Faigeles BS, Cullen SP, et al. THRIVE score predicts ischemic stroke outcomes and thrombolytic hemorrhage risk in VISTA. Stroke 2013;44:3365-3369.
8. Strbian D, Meretoja A, Ahlhelm FJ, et al. Predicting outcome of IV thrombolysis-treated ischemic stroke patients: the DRAGON score. Neurology 2012;78:427-432.
9. Ntaios G, Faouzi M, Ferrari J, et al. An integer-based score to predict functional outcome in acute ischemic stroke: the ASTRAL score. Neurology 2012;78:1916-1922.
10. Saposnik G, Fang J, Kapral MK, et al. The iScore predicts effectiveness of thrombolytic therapy for acute ischemic stroke. Stroke 2012;43:1315-1322.
11. Kent DM, Selker HP, Ruthazer R, et al. The stroke-thrombolytic predictive instrument: a predictive instrument for intravenous thrombolysis in acute ischemic stroke. Stroke 2006;37:2957-2962.
12. Cooray C, Mazya M, Bottai M, et al. External validation of the ASTRAL and DRAGON scores for prediction of functional outcome in stroke. Stroke 2016;47:1493-1499.
13. Strbian D, Seiffge DJ, Breuer L, et al. Validation of the DRAGON score in 12 stroke centers in anterior and posterior circulation. Stroke 2013;44:2718-2721.
14. Liu G, Ntaios G, Zheng H, et al. External validation of the ASTRAL score to predict 3- and 12-month functional outcome in the China National Stroke Registry. Stroke 2013;44:1443-1445.
15. Cappellari M, Turcato G, Forlivesi S, et al. The START nomogram for individualized prediction of the probability of unfavorable outcome after intravenous thrombolysis for stroke. Int J Stroke 2018;13:700-706.
16. Shariat SF, Karakiewicz PI, Suardi N, et al. Comparison of nomograms with other methods for predicting outcomes in prostate cancer: a critical analysis of the literature. Clin Cancer Res 2008;14:4400-4407.
17. Balachandran VP, Gonen M, Smith JJ, et al. Nomograms in oncology: more than meets the eye. Lancet Oncol 2015;16:e173-e180.
18. Jehi L, Yardi R, Chagin K, et al. Development and validation of nomograms to provide individualised predictions of seizure outcomes after epilepsy surgery: a retrospective analysis. Lancet Neurol 2015;14:283-290.
19. Callegaro D, Miceli R, Bonvalot S, et al. Development and external validation of two nomograms to predict overall survival and occurrence of distant metastases in adults after surgical resection of localised soft-tissue sarcomas of the extremities: a retrospective analysis. Lancet Oncol 2016;17:671-680.
20. Hijazi Z, Oldgren J, Lindback J, et al. The novel biomarker-based ABC (age, biomarkers, clinical history)-bleeding risk score for patients with atrial fibrillation: a derivation and validation study. Lancet 2016;387:2302-2311.
21. Wang Y, Li J, Xia Y, et al. Prognostic nomogram for intrahepatic cholangiocarcinoma after partial hepatectomy. J Clin Oncol 2013;31:1188-1195.
22. Huang YQ, Liang CH, He L, et al. Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal cancer. J Clin Oncol 2016;34:2157-2164.
23. Tevis SE, Weber SM, Kent KC, et al. Nomogram to predict postoperative readmission in patients who undergo general surgery. JAMA Surg 2015;150:505-510.
24. Shariat SF, Capitanio U, Jeldres C, et al. Can nomograms be superior to other prediction tools? BJU Int 2009;103:492-495. discussion 495-497.
25. Chun FK, Karakiewicz PI, Briganti A, et al. A critical appraisal of logistic regression-based nomograms, artificial neural networks, classification and regression-tree models, look-up tables and risk-group stratification models for prostate cancer. BJU Int 2007;99:794-800.
26. Alba AC, Agoritsas T, Walsh M, et al. Discrimination and calibration of clinical prediction models: users' guides to the medical literature. JAMA 2017;318:1377-1384.
27. Uno H, Cai T, Pencina MJ, et al. On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data. Stat Med 2011;30:1105-1117.
28. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. Circulation 2007;115:928-935.
29. Lee YH, Bang H, Kim DJ. How to establish clinical prediction models. Endocrinol Metab (Seoul) 2016;31:38-44.
30. Saposnik G, Cote R, Phillips S, et al. Stroke outcome in those over 80: a multicenter cohort study across Canada. Stroke 2008;39:2310-2317.
31. Ford GA, Ahmed N, Azevedo E, et al. Intravenous alteplase for stroke in those older than 80 years old. Stroke 2010;41:2568-2574.
32. Karlinski M, Kobayashi A, Czlonkowska A, et al. Role of preexisting disability in patients treated with intravenous thrombolysis for ischemic stroke. Stroke 2014;45:770-775.
33. Weimar C, Konig IR, Kraywinkel K, et al. 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;35:158-162.
34. Emberson J, Lees KR, Lyden P, et al. Effect of treatment delay, age, and stroke severity on the effects of intravenous thrombolysis with alteplase for acute ischaemic stroke: a meta-analysis of individual patient data from randomised trials. Lancet 2014;384:1929-1935.
35. Moser DK, Kimble LP, Alberts MJ, et al. Reducing delay in seeking treatment by patients with acute coronary syndrome and stroke: a scientific statement from the American Heart Association Council on Cardiovascular Nursing and Stroke Council. J Cardiovasc Nurs 2007;22:326-343.
36. Steyerberg EW, Moons KG, van der Windt DA, et al. Prognosis Research Strategy (PROGRESS) 3: prognostic model research. PLoS Med 2013;10:e1001381.
37. Toll DB, Janssen KJ, Vergouwe Y, et al. Validation, updating and impact of clinical prediction rules: a review. J Clin Epidemiol 2008;61:1085-1094.