Study on risk weights of different stroke risk factors

Yiyan Zhang, *Yi* Xin, *Hongyu* Kang, *Weiqi* Lv, and *Qin* Li[*]

Department of Biomedical Engineering, School of Life Science, Beijing Institute of Technology, 5 South Zhongguancun Street, Haidian District, Beijing 100081, China

Abstract. Stroke is a common disease that seriously endangers human health and life. According to the number of eight traditional risk factors of stroke to determine the stroke high-risk group, there are still some limitations. The study aims to explore the risk weights of different stroke risk factors. The research data used in this study was collected by the stroke prevention and control project. The study population included adults aged ⩾40 years residing in 16 provinces, municipalities and autonomous regions throughout China. 985, 325 subjects were included in the study, 15,835 stroke patients (1.65%) and 941,490 without stroke (98.35%). The Cumulative score model was applied to explore the stroke risk factors and their risk weights based on Logistic regression. From high to low, the risk of eight traditional factors is hypertension, family history of stroke, atrial fibrillation, diabetes mellitus, dyslipidemia, obvious weight, lack of physical activity and smoking. The weight coefficient of hypertension is about 7.7 times that of smoking.

1 Introduction

Stroke is a common and frequently-occurring disease that seriously endangers human health. In China, one person is diagnosed with stroke within an average of 12 seconds and one person dies of stroke every 21 seconds [1]. Stroke is the third leading cause of death in the United States and Britain, just behind heart disease and cancer. Stroke is also the first cause of adult disability, with 90% of survivors having varying degrees of functional loss [2]. World Health Organization predicted that if the mortality is not controlled, nearly 400 million people die from stroke per year in China by 2030 [3]. At present, the annual cost of treating cerebrovascular diseases in our country is over 10 billion, resulting in an indirect economic loss of nearly 20 billion each year [4]. Therefore, fully understanding the severity of stroke, increasing the effect of treatment and prevention of stroke, and reducing the morbidity and mortality of stroke has become an urgent task in public health.

Numerous studies have shown that stroke risk factors include hypertension, diabetes mellitus, heart disease, dyslipidemia, smoking, excessive drinking, aging and genetic factors [5]. Risk factors for stroke can be divided into non-intervention risk factors and interventional risk factors, and interventional risk factors can be divided into physiological risk factors and behavioral risk factors [6]. Several studies have shown that hypertension is

---

[*] Corresponding author: liqin@bit.edu.cn

the most important risk factor for ischemic stroke, and stroke is a major outcome of hypertension [7, 8]. Diabetes has been recognized as an independent risk factor for stroke and is one of the most important risk factors of cerebral hemorrhage in young people [9]. In recent years, some new risk factors in the pathogenesis of stroke have been found, such as hyperhomocysteinemia, individual genetic predisposition [10]. Despite many new risk factors have been found in recent years, hypertension, diabetes and atrial fibrillation remain the most common risk factors for acute ischemic stroke [11].

Guidelines for Classification and Treatment of Stroke in Anhui Province (2015 Edition) stipulate that the risk factors for stroke in the preliminary screening table of stroke, each item counts as 1 point, and 3 points or more is classified as high-risk population [12]. However, the risk of 8 risk factors for stroke is not equal, this calculation method still exist some limitations. To study the causes of stroke and risk weights of these factors, the more accurate stroke risk prediction model can be established.

The rest of this paper is organized as follows. Section 2 describes in detail the research data and method. Furthermore, Section 3 illustrates the model and displays the performance assessments of the model. Section 4 presents the discussion. Finally, Section 5 provides the conclusions.

2 Materials and methods

2.1 Participants

The data used in this study was obtained from a stroke screening and prevention project. The project conducted among 1,196,422 residing in 16 provinces, municipalities and autonomous regions throughout China. The subjects were selected using a multistage, probability proportionate-to-size, cluster-sampling technique. Data were collected using an interviewer-administered questionnaire. Trained and qualified investigators filled in the validated questionnaire following a face-to-face interview with each participant. In each screening site, all residents over the age of 40 (born before December 31, 1973) were enrolled as screening subjects. Residents living outside the screening site for more than six months were excluded. The sixth national census was used to determine the proportion of urban communities and suburban towns.

2.2 Research data items

Demographics include age, gender, nationality, marital status, employment status, education and current living district. Traditional eight risk factors are hypertension, atrial fibrillation, smoking, dyslipidemia, diabetes mellitus, lacking of physical training, obvious overweight and stroke family history. Ever diagnosis of stroke and eight risk factors was confirmed by doctors based on World Health Organization criteria.

2.3 Cumulative score model

Whether suffering from stroke as dependent variable (Yes=0, No=1), all 8 risk factors and relevant demographic factors as independent variables, multivariate logistic regression analysis was performed. The significance of Logistic regression coefficient β is that for each increase in the independent variable, the subject is at increased risk of stroke by a corresponding level. According to this principle, the study finally determine 10 times of each regression coefficient as a score of each independent variable. The regression

coefficient is retained to three decimal places. The higher the cumulative score of a subject, the greater the risk of stroke.

SPSS (Statistical Product and Service Solutions) Version 19.0 was applied to conduct statistical analysis. A two-tailed test was used to judge the statistical significance with a test level of 0.05.

2.4 Assessment method of prediction model

In this paper, the sensitivity, specificity and Youden index as the evaluation index of the prediction. Youden index equals the sum of sensitivity and specificity minus one. Critical threshold for dividing high-risk group was based on the Youden index. The receiver operating characteristic curve (ROC curve) and the area under the curve (AUC) were used to evaluate the established prediction model and its practical value.

3 Results

3.1 Stroke morbidity

Based on the population base of the sixth census, the morbidity of stroke in 16 provinces and municipalities was 1.35%. Among them, the stroke morbidity of urban population is 1.20%, the morbidity of rural population is 1.49%. The morbidity of rural population is slightly higher than that of urban areas. Considered from region, the highest prevalence 2.78% was found in Hebei Province, followed by Heilongjiang Province (2.13%), Henan Province (1.82%), Shandong Province (1.73%) and Beijing (1.70%). Overall, the morbidity of the South China is lower than the North China.

3.2 Stroke risk prediction model

Multivariate non-conditional logistic regression analysis was used to select variables as risk factors. The multivariate non-conditional logistic regression result showed that 15 factors had a $P> 0.05$, which was not statistically significant. As can be seen from table 1, 30 items were included in the cumulative score model, the score of each variable was 10 times of corresponding regression coefficient.

Table1. Cumulative score model for stroke risk prediction.

Factors	β	Sig.	Exp (β)	95% C.I.		Score
				Lower limit	Upper limit	
Hypertension	1.515	0.000	4.550	4.383	4.722	15
Arterial fibrillation	0.782	0.000	2.185	2.074	2.302	8
Smoking	0.197	0.000	1.218	1.167	1.271	2
Dyslipidemia	0.481	0.000	1.617	1.564	1.672	5
Diabetes mellitus	0.677	0.000	1.969	1.887	2.054	7
Lack of physical	0.289	0.000	1.335	1.289	1.384	3
Obvious overweight	0.403	0.000	1.496	1.442	1.553	4
Stroke family history	1.168	0.000	3.215	3.087	3.347	12
45-50 years old	-2.125	0.000	0.119	0.103	0.139	-21
50-54 years old	-1.693	0.000	0.184	0.166	0.204	-17

55-59 years old	-1.166	0.000	0.312	0.287	0.339	-12
60-64 years old	-0.737	0.000	0.479	0.445	0.515	-7
65-69 years old	-0.428	0.000	0.652	0.610	0.697	-4
70-74 years old [a]	-0.179	0.000	0.836	0.783	0.893	-2
Male vs female	-0.443	0.000	0.642	0.619	0.666	-4
The Zang nationality	-0.479	0.035	0.620	0.397	0.968	-5
Uyghur nationality	-0.880	0.239	0.415	0.096	1.794	-9
Other nationalities	0.269	0.533	1.309	0.562	3.050	3
Married	0.424	0.001	1.528	1.179	1.981	4
First marriage	0.593	0.000	1.810	1.422	2.304	6
Divorce	0.611	0.000	1.843	1.433	2.371	6
Unaccounted marital	0.726	0.000	2.066	1.437	2.970	7
Professional	0.290	0.000	1.336	1.215	1.469	3
Enterprise manager	0.125	0.008	1.134	1.034	1.243	1
Peasant	0.111	0.000	1.118	1.050	1.190	1
Student	0.420	0.000	1.522	1.434	1.615	4
Unemployed	0.224	0.003	1.251	1.077	1.452	2
Other	0.261	0.000	1.299	1.185	1.423	3
Junior high school	0.700	0.025	2.015	1.093	3.712	7
Technical secondary	0.622	0.046	1.863	1.011	3.433	6
Constant	-4.968	0.000	0.007			

[a] For 70-74 years old group, on account of the small number of population, there was no statistical significance.

The score range calculated using the cumulative score model is (-30, 64). First, the score was divided into ten levels. Then, the actual morbidity of stroke in each level is calculated. Finally, the fit function was used to match the score to the actual morbidity.
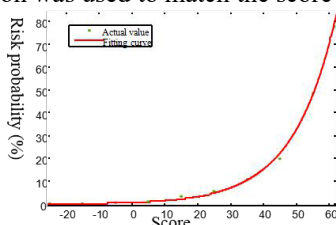

Fig. 1. Predictive curve of stroke risk.

As shown in figure 1, the stroke risk presents an increasing trend as the score increases. The fitting function is:

$$P(x) = 0.6832 e^{0.07705x}$$

(1)

P represents the stroke risk and x represents the score calculated by the cumulative score model.

3.3 Assessment of stroke risk prediction model

The threshold should be determined by considering the sensitivity and specificity of the prediction results. Therefore, this paper used the Youden index to determine the best cut-off point.

Table 2. Assessment of cumulative score model (Test set).

Threshold	Sensitivity	Specificity	Youden index
6	0.818	0.694	0.511
7	0.799	0.721	0.519
8	0.783	0.746	0.529
9	0.761	0.763	0.523
10	0.742	0.784	0.525
11	0.721	0.797	0.518

The assessment results in table 2 showed that when 8 was the threshold of the score, the Youden index reaches the highest. At this point, the sensitivity and specificity was 78.3% and 74.6%, respectively. The AUC of the cumulative score model was 0.817.

4 Discussion

Our results show that hypertension has become the first predisposing factor of stroke, much higher than other factors. Hypertension is one of the important reasons that cause and promote cerebral arteriosclerosis. Long-term blood pressure rising can lead to the formation and aggravation of atherosclerosis, and then stroke. If the patient's blood pressure can be reduced to a reasonable level, the incidence of stroke can be reduced by at least 38% -35% [13].

In the present study, atrial fibrillation was the third highest risk factor for stroke, after the family history of stroke. Professor Zhu Jun pointed out that heart with atrial fibrillation would beat rapidly and irregularly, blood cannot be pumped out completely and stasis in the atria, resulting in thrombosis. After the blood clots fall, they will flow to the brain along with the blood and block brain blood vessel, and ultimately lead to ischemic stroke [14].

Diabetes mellitus is the fourth stroke risk factor. Insulin changes, changes in plasma lipoproteins and glucose metabolism caused by diabetes will lead to atherosclerosis and thrombosis. Therefore, diabetes mellitus becomes a vital important factor for stroke. Piernik-Yoder found that one in three acute stroke patients have diabetes mellitus. On the age of onset, diabetic stroke has the characteristics of young age, and more significant impact on women [15].

The INTERSTROKE study findings showed that five risk factors accounted for more than 80% of the global risk of all stroke (ischaemic and intracerebral haemorrhagic): hypertension, current smoking, abdominal obesity, diet, and physical activity [16]. In our study, hypertension also was the greatest risk factor for stroke, but smoking, lacking of physical training and obvious overweight contributed less risk weights. In addition, our results found that the prevalence of stroke showed a regional difference.

This is the first large-scale population-based study on stroke prevalence in China. These data may not represent the country as a whole because of regional differences in socioeconomic status and risk behaviors. Future studies will include participants from urban and rural communities from different parts of the country to improve external validity.

5 Conclusion

This study demonstrates that assessing the risk level of stroke based on risk weights of different risk factors is meaningful and practicable. Prevention and intervention of high risk factors for stroke will greatly reduce the incidence of stroke. Especially in northern China, stroke prevention should be strengthened for those with a family history of stroke,

hypertension, atrial fibrillation and diabetes. Regular physical activity and reducing smoking, these changes of lifestyle will be beneficial to reduce the risk of stroke.

References

1. Sun X Q, Liu X M, Liu Z Y, Li H, Wang M J 2013 *J.* Population status and stroke survey (Incidence analysis of stroke screening population). *Medical Laboratory Science and Clinics* **24** 59-66.

2. Higashida R, Alberts M J, Alexander D N 2013 *J.* Interactions Within Stroke Systems of Care A Policy Statement From the American Heart Association/American Stroke Association. *Stroke* **44** 2961-84.

3. Wang L D 2011. China Stroke Declaration.

4. Liu G F, Li X, Wang Y L, Du W L, Zhou Y, Li Z X, Zhao X Q, Wang C X, Gao X L, Guo M N, Xie X Q, Hu M, Wang Y J 2012 *J.* Evaluation of Medical Service Performance in Cerebrovascular Disease Inpatients in Beijing. *Chinese Journal of Stroke* **7** 606-8.

5. Marshall J, Wang Y, McKevitt C, Rudd A G 2013 *J.* Trends in Risk Factor Prevalence and Management Before First Stroke Data From the South London Stroke Register 1995–2011. *Stroke* **4** 3298-304.

6. He H L, Wang P X 2012 *J.* Research Progress of Risk Factors and Preventive Measures for Stroke. *Chinese Journal of Practical Nursing* **4** 194-6.

7. Han Y J 2011 *J.* Research Progress on Risk Factors of Ischemic Stroke. *China Medicine and Pharmacy* **1** 57-9.

8. Remme W J, Swedberg K 2001 *J.* Task force for the diagnosis and treatment of chronic heart failure, European Society of Cardiology. Guidelines for the diagnosis and treatment of chronic heart failure. *Eur Hear* **22** 1527-60.

9. Oldstein L B, Adams R, Becker K 2001 *J.* Primary prevention of chemic stroke: A statement for healthcare professionals from the Stroke Council of the Americoll Heart Asseciation. *Circulation* **103** 163-82.

10. Lgu E, Bungard T J, Tsuyuki R T 2004 *J.* Stroke prophylaxis in institutionalized elderly patients with atrial fibrillation. *J Am Geriatr Soc* **52** 428-38.

11. Li Y J, Wang M X 2016 *J.* Updated data on risk factors for subtypes of acute ischemic stroke. *Chin J Clinicians (Electronic Edition)* **10** 722-7.

12. Wang Y, Wang C Q, Tian Y H, Wang K, Wang J Y, Su X, Yang W M, Chen J, Ni C M 2016 *J.* Stroke Grading Guideline of Anhui Province. *Anhui Medical Journal* **4** 367-82.

13. Faraco G, Iadecola C 2013 *J.* Hypertension A Harbinger of Stroke and Dementia. *Am Heart Assoc* **62** 810-7.

14. Li Y N 2012 *J.* Atrial fibrillation is a risk factor for stroke. *Qiuyi Wenyao* **12** 46-49.

15. Piernik-Yoder B, Ketchum N 2013 *J.* Rehabilitation Outcomes of Stroke Patients With and Without Diabetes. *Arch Phys Med Rehab* **8** 1508-12.

16. O'Donnell M J, Xavier D, Liu L S, Zhang H Y, Chin S L, Rao-Melacini P, Rangarajan S, Islam S, Pais P, McQueen M J, Mondo C, Damasceno A, Lopez-Jaramillo P, Hankey G J, Dans A L, Yusoff K, Truelsen T, Diener H-C, Sacco R L, Ryglewicz D, Czlonkowska A, Weimar C, Wang X Y, Yusuf S 2010 *J.* Risk factors for ischaemic and intracerebral haemorrhagic stroke in 22 countries (the INTERSTROKE study): a case-control study. *Lancet* **376** 112-23.