

Verification and comparison of three prediction models of ischemic stroke in young adults based on the back propagation neural networks

Yuyang Chen, BS^a, Yingqi Mao, MM^a, Xiaoyun Pan, BS^a, Weifeng Jin, PhD^b, Tao Qiu, PhD, MD^{a,*}

Abstract

This work aims to explore risk factors for ischemic stroke in young adults and analyze the Traditional Vascular Risk Factors Model based on age, hypertension, diabetes, smoking history, and drinking history. Further, the Lipid Metabolism Model was analyzed based on lipoprotein a [LP (a)], high-density lipoprotein (HDL), low-density lipoprotein (LDL), apolipoprotein AI (apo AI), apolipoprotein B (apo B), and the Early Renal Injury Model based on urinary microalbuminuria/creatinine ratio (UACR). Besides, we estimated glomerular filtration rate (eGFR), cystatin C (Cys-C), homocysteine (Hcy), β 2 microglobulin (β 2m), and validated their predictive efficacy and clinical value for the development of ischemic stroke in young adults.

We selected and retrospectively analyzed the clinical data of 565 young inpatients admitted to Zhejiang Provincial Hospital of Chinese Medicine between 2010 and 2020, 187 of whom were young stroke patients. A single-factor analysis was used to analyze the risk factors for stroke in young people and developed a traditional vascular risk factors model, a lipid metabolism model, and an early kidney injury model based on backpropagation (BP) neural networks technology to predict early stroke occurrence. Moreover, the prediction performance by the area under the receiver operating characteristics (ROC) curve (AUC) was assessed to further understand the risk factors for stroke in young people and apply their predictive role in the clinical setting.

Single-factor analysis showed that ischemic stroke in young adults was associated with hypertension, diabetes, smoking history, drinking history, LP(a), HDL, LDL, apo AI, apo B, eGFR, Cys-C, and β 2m ($P < .05$). The BP neural networks technique was used to plot the ROC curves for the Traditional Vascular Risk Factors Model, the Lipid Metabolism Model, and the Early Kidney Injury Model in enrolled patients, and calculated AUC values of 0.7915, 0.8387, and 0.9803, respectively.

The early kidney injury model precisely predicted the risk of ischemic stroke in young adults and exhibited a certain clinical value as a reference for morbidity assessment. Whereas the prediction performance of the Traditional Vascular Risk Factors Model and the Lipid Metabolism Model were inferior to that of the early kidney injury model.

Abbreviations: β 2m = β 2 microglobulin, apo AI = apolipoprotein AI, apo B = apolipoprotein B, AUC = area under the receiver operating characteristics curve, BP = backpropagation, Cys-C = cystatin C, eGFR = estimated glomerular filtration rate, Hcy = homocysteine, HDL = high-density lipoprotein, LDL = low-density lipoprotein, LP (a) = lipoprotein a, ROC = receiver operating characteristics, UACR = urinary microalbuminuria/creatinine ratio.

Keywords: neural networks, risk factors, statistical model, stroke

Editor: Mustafa Büyükmumcu.

YC and YM contributed equally to this work.

This work has been supported by the National Natural Science Foundation of China (Grant No. 81774230), the Natural Science Foundation of Zhejiang Province (Grant No. LY16H270002), and the Administration of Traditional Chinese Medicine of Zhejiang Province, China (Grant No. 2019ZB036).

The authors have no conflicts of interests to disclose.

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

^aThe First School of Medicine, ^bThe College of Pharmaceutical Science, Zhejiang Chinese Medical University, Hangzhou, PR China.

* Correspondence: Tao Qiu, The First School of Medicine, Zhejiang Chinese Medical University, 548th Binwen Road, Hangzhou 310053, PR China (e-mail: qiutao2002002@163.com).

Copyright © 2021 the Author(s). Published by Wolters Kluwer Health, Inc.

This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial License 4.0 (CCBY-NC), where it is permissible to download, share, remix, transform, and buildup the work provided it is properly cited. The work cannot be used commercially without permission from the journal.

How to cite this article: Chen Y, Mao Y, Pan X, Jin W, Qiu T. Verification and comparison of three prediction models of ischemic stroke in young adults based on the back propagation neural networks. *Medicine* 2021;100:11(e25081).

Received: 12 November 2020 / Received in final form: 14 February 2021 / Accepted: 16 February 2021

<http://dx.doi.org/10.1097/MD.00000000000025081>

1. Introduction

Ischemic stroke in young adults is an increasingly prevalent type of stroke in clinical practice, characterized by a high rate of disability and a long life expectancy. Globally, there is an incidence of 2 million young stroke patients annually,^[1] accounting for 5% to 15% of all strokes.^[2] A large cohort study originating in the Netherlands revealed that the incidence of ischemic stroke in young adults increased by almost 50% between 1998 and 2010.^[3] Although this group represents only a small fraction of the stroke population, the psychosocial and economic burden of any chronic or devastating disease in this particular group has a significant impact and adversely affect the development of society, specifically in the developing world. In terms of lifelong dependency and disability, the consequences of ischemic stroke in young adults are enormous. Evidence has shown that only 40% of young stroke victims can resume their former occupation, while about a third can never return to work. At the same time, the relative rarity of ischemic stroke in young adults appear to delay diagnosis and often brings hurdles in the search for its cause.^[4] Notably, ischemic stroke in young adults is not caused by a single factor, but rather by a combination of factors. As such, it is important to analyze the risk factors for ischemic stroke in young adults to help predict and prevent the onset of ischemic stroke in young adults. BP neural networks, a growing data mining algorithm, can mine implicit information among data and optimize model performance by adjusting the weight distribution among variables. We explored risk factors for ischemic stroke in young adults using BP neural networks analysis to develop a traditional vascular risk factor model, a lipid metabolism model, and an early kidney injury model. Also, we assessed predictive efficacy by area under the receiver operating characteristics curve (AUC).

2. Methods

2.1. Study subjects

A retrospective observational design was used for this study. A total of 565 young inpatients (18 ≤ age ≤ 45 years) were enrolled from Zhejiang Provincial Hospital of Chinese Medicine between January 2010 and March 2020. Out of these, 312 were males while were 253 females, with a mean age of 36 years. Besides, 187 were young stroke patients while 378 were non-young stroke patients. The mean age of the 2 groups of patients in each physical and chemical index was comparable by t-test with no significant differences. The inclusion criteria were:

1. 18 ≤ age ≤ 45 years;
2. met the diagnostic criteria for acute ischemic stroke;
3. presence of new infarction confirmed by cranial magnetic resonance imaging (MRI);
4. time from onset to hospital admission ≤ 7 days.

On the other hand, the exclusion criteria included:

1. incomplete medical history or physical and chemical parameters;
2. coexistence of chronic inflammatory reactions, malignancies;
3. pregnant or lactating women;
4. participation in other clinical trials.

The study was approved by the Ethics Review Committee of Zhejiang Provincial Hospital of Chinese Medicine (Ethics approval number: 2020-KL-117-02).

2.2. Data collection and definition

Information on young inpatients was obtained by reviewing medical records, among them: general information; clinical symptoms; and physical and chemical information collections: blood parameters such as blood routine, biochemistry, coagulation, lipids, and early kidney injury (The physical and chemical parameters of blood were collected and tested by the Laboratory of Zhejiang Provincial Hospital of Chinese Medicine). Data were collected from all admissions and patients were included in the stroke group when they were diagnosed with acute ischemic stroke based on their discharge diagnosis, otherwise they were included in the non-stroke group. The estimated glomerular filtration rate (eGFR) was calculated using the Chronic Kidney Disease Epidemiology Collaboration formula.^[5]

2.3. Statistical analysis

Microsoft Excel (Microsoft 365, 2020) was used for data entry, and after proofreading, the data was imported into the Excel database. Statistical analysis was performed using R programming language (version 3.5.0). The counting data were tested by a Chi-Squared test. The enumeration data were compared by χ^2 tests, while measurement data were compared by the independent sample t-test between 2 groups. Non-parametric tests were performed for non-normal distribution data. A BP neural network model was used for prediction, and the closer the R2 to 1, the better the fit. The value of the BP neural networks model for predicting the occurrence of ischemic stroke in young adults was evaluated using the ROC curve. The AUC comparisons were made using non-parametric tests. *P* values < .05 were considered significant.

3. Results

3.1. The risk factors for ischemic stroke in young adults

A total of 565 young inpatients were enrolled in the study, including 187 young stroke patients and 378 non-young stroke patients. Single-factor analysis showed that ischemic stroke in young adults was related to hypertension, diabetes, smoking history, drinking history, lipoprotein a [LP(a)], high-density lipoprotein (HDL), low-density lipoprotein (LDL), apolipoprotein AI (apo AI), apolipoprotein B (apo B), eGFR, cystatin C (Cys-C), and β_2m (*P* < .05), but not with urinary microalbuminuria/creatinine ratio (UACR) (*P* = .058) and Hcy (*P* = .116) (Figs. 1–3).

3.2. BP neural networks model building

The BP neural networks model were entered with the occurrence of ischemic stroke in young adults as the dependent variable (assignment: yes = 1, no = 0) and the variables set by each model as independent variables [assignment of categorical variables: yes = 1, no = 0; standardized formulas were applied to the enumeration data, and all input layer covariates were normalized to the (0, 1) interval]. The traditional vascular risk factor model represents a BP neural networks model with an implied layer of 3, which predicts the data with an accuracy of 63.27%. The Lipid Metabolism Model depicts a BP neural networks model with an implied layer of 3, which predicts the data with an accuracy of 76.17%; the early kidney injury model represents a BP neural networks model with an implied layer of 5, which predicts the data with an accuracy of 82.80% (Table 1).

Traditional Vascular Risk Factors Model

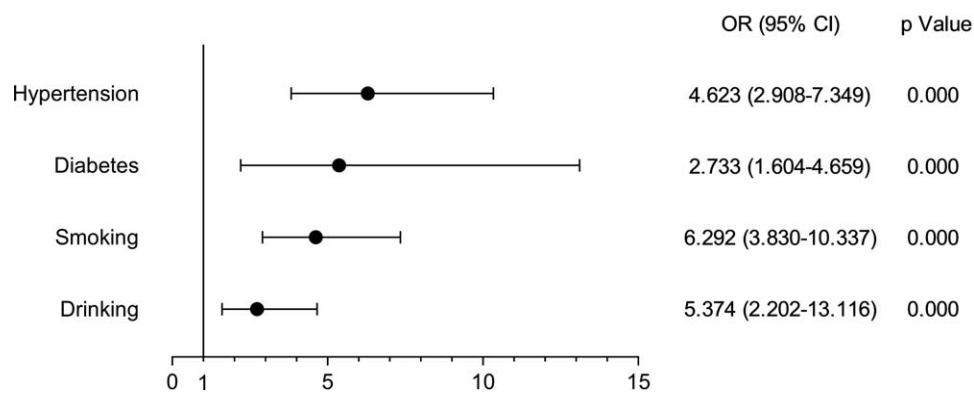


Figure 1. Correlation between risk factors and stroke in young adults in traditional vascular risk factor model.

3.3. Comparison of the predictive values of three models

The AUC values for the traditional vascular risk factor model, the lipid metabolism model, and the early kidney injury model were 0.7915, 0.8387, and 0.9803, and the differences were statistically significant for all three models (Table 2, Fig. 4).

4. Discussion

Ischemic stroke in young adults is becoming increasingly prevalent in clinical practice. There are at least 1.5 million new young stroke patients across the globe each year. Its incidence is annually increasing hence having a huge impact on the productivity and life expectancy of patients. Of note, this trend poses a huge threat to social and economic stability. Nevertheless, ischemic stroke in young adults is not caused by a single factor, but rather a combination of factors that multiply the risk.^[6] Therefore, it is essential to analyze the risk factors associated with ischemic stroke in young adults geared towards predicting and preventing this type of stroke.

BP neural networks,^[7] also known as backpropagation neural networks, is a multilayered feedforward neural networks in the artificial neural networks, which is the most representative and extensive. BP neural networks mine implicit information within a dataset to optimize model performance by adjusting the weight distribution of the variables. In this paper, we explored the risk factors for ischemic stroke in young adults and utilized BP neural networks analysis to develop the traditional vascular risk factor

model, the lipid metabolism model, and the early kidney injury model. Additionally, the predictive efficacy in areas under the ROC curve (AUC) was assessed. Our data revealed that ischemic stroke in young adults was associated with hypertension, diabetes, smoking history, drinking history, LP(a), HDL, LDL, apo AI, apo B, eGFR, Cys-C, and β 2m, but not with UACR and Hcy. The accuracy of the traditional vascular risk factor model, the lipid metabolism model, and the early renal injury model was 63.27%, 76.17%, and 82.80%, with AUC values of 0.7915, 0.8387, and 0.9803 respectively. Notably, the early kidney injury model exhibited higher fidelity, better agreement with actual observations, and more reliable predictions.

The importance of traditional vascular risk factors for ischemic stroke in young adults was evident. The increased prevalence of traditional risk factors including hypertension and diabetes mellitus in the young population has been explored in numerous studies^[8] and coincides with an increase in hospitalization rates for young stroke patients. Also, the prevalence of the traditional vascular risk factors for ischemic stroke in young adults is almost double that of the general population.^[9] The present study suggests that among the traditional vascular risk factors, hypertension and diabetes were the primary risk factors for ischemic stroke in young adults. This was in line with the findings of Shahid et al.^[10] Hypertension can be directly or indirectly implicated in the formation of atherosclerotic plaque through vascular endothelial injury, secretion of vasoactive substances,

Lipid Metabolism Model

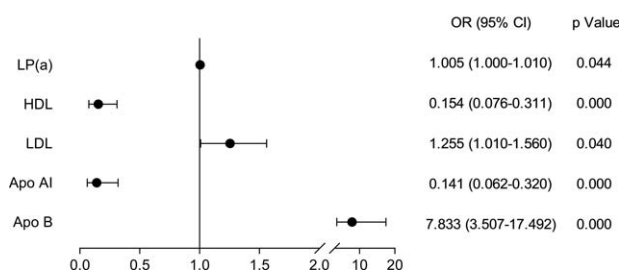


Figure 2. Correlation between risk factors and stroke in young adults in lipid metabolism model.

Early Renal Injury Model

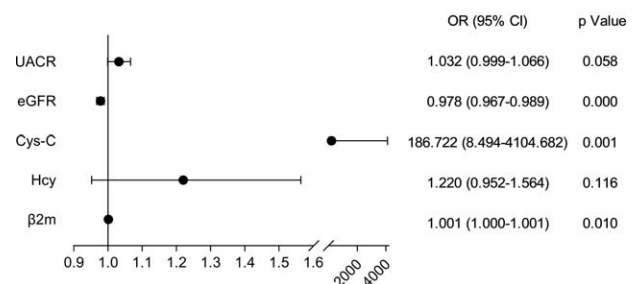


Figure 3. Correlation between risk factors and stroke in young adults in early kidney injury model.

Table 1
Prediction of the training set samples with the BP neural networks model.

Predicted value	Observed value		Total
	Stroke (+)	Stroke (-)	
Traditional Vascular Risk Factors Model (+)	106	41	147
Traditional Vascular Risk Factors Model (-)	79	315	394
Lipid Metabolism Model (+)	19	1	20
Lipid Metabolism Model (-)	12	22	34
Early Renal Injury Model (+)	23	1	24
Early Renal Injury Model (-)	0	20	20

and inflammatory reactions, which leads to ischemic stroke in young adults. Diabetes mellitus triggers the formation of advanced glycation end-products (AGEs), insulin resistance, and abnormal polyol metabolism, which in turn causes microvascular and macrovascular pathologies. Chronic hyperglycemia causes disorders of glucose and lipid metabolism, and increases blood viscosity, thereby aggravating vascular endothelial cell damage and accelerating the development of atherosclerosis. In total, 50% to 70% of people diagnosed with diabetes eventually become disabled or die from atherosclerosis of the heart, cerebral vessels, or peripheral arteries. Furthermore, our prediction model based on traditional risk factors showed positive predictive power for ischemic stroke in young adults. However, this is often at the disadvantage of misreporting or concealing patient history, hence, this work aimed to further identify more precise and better predictive models.

Increasingly recent epidemiological evidence supported the role of abnormal lipid metabolism in increasing the risk of ischemic stroke.^[11] Apo B promotes the phagocytosis of oxidized LDL by monocytes or macrophages, thereby promoting the formation of foam cells. Therefore, Apo B promotes atherosclerosis hence promoting thrombosis and increasing the risk of ischemic stroke.^[12] In line with the known literature,^[13,14] we also found that Apo B was a greater risk factor for ischemic stroke in young adults. Whereas HDL and Apo AI played a protective role in the development of ischemic stroke in young adults. Studies have shown that HDL plays a vital role in stopping the development and progression of atherosclerotic plaque formation by primarily protecting the vascular endothelium, preventing oxidation, and fighting inflammation. HDL reduces the formation of atherosclerotic plaque by transporting cholesterol in a reverse direction,^[15] which in turn reduces the incidence of stroke. Similarly, Apo AI, present in HDL, maintains the HDL structure, is produced by the small intestine and liver, recognizes receptors on tissue cells, promotes lipid transport, regulates enzyme activity, and binds lipoproteins to cell membrane

receptors. Moreover, it activates the transport of lecithin cholesterol acyltransferase (LCAT) as well as break down cholesterol.^[16] In addition, Apo AI inhibits LDL oxidation, regulates inflammatory response, promotes cholesterol outflow from the arterial blood vessel wall, and fights atherosclerosis. The lipid metabolism model we developed also demonstrated better accuracy compared to traditional vascular risk factor models.

Notably, Cys-C showed a dramatic effect on ischemic stroke in young adults. Cys-C is a protein that passes freely within the glomerulus and is not reabsorbed by tubular epithelial cells, and is a reliable indicator of glomerular filtration capacity.^[17] Studies have shown that high serum concentrations of Cys-C are implicated in arterial stenosis and atherosclerotic plaque formation^[18] and are associated with an increased risk of stroke.^[19] Nevertheless, Hcy showed a protective effect against ischemic stroke in young adults in our analysis. This contradicted numerous known studies that have been conducted on the effects of Hcy on stroke risk. We believe this was since the non-stroke group of patients were also inpatients in Zhejiang Provincial Hospital of Chinese Medicine and were relatively unrepresentative causing errors in Hcy levels. Here, we found that the early kidney injury model was more accurate compared to the traditional vascular risk factor model and the lipid metabolism model in predicting ischemic stroke in young adults. Previous works^[20] demonstrated that the vasculature of the brain and kidney has similar hemodynamics and that the mechanisms of vascular regulation in both have some similarities. Both organs have low vascular resistance, thereby allowing continuous perfusion of blood at high volumes, and both are target organs for the pathological process of atherosclerosis, which are both highly similar high-risk factors for endothelial injury.^[21] Early renal damage is characterized by disorders and abnormalities of glomerular endothelial cell function. On the other hand, cerebrovascular damage is caused by endothelial cell dysfunction, ischemic atherosclerosis, hypoperfusion, and disruption of the blood-brain barrier, both of which are characterized by lesions in small arteries. Small blood vessels are susceptible to lipid hyaline degeneration and endothelial dysfunction under the long-term influence of certain high-risk factors, and these pathological changes underlie cerebrovascular pathology and renal vascular injury. When the kidneys are damaged, disturbances in lipid metabolism and chronic inflammatory processes accelerate the formation of atherosclerosis thus increasing the risk of ischemic stroke in young adults.

This was a single-centered, retrospective, and observational study with a small sample size and patients restricted to Zhejiang Provincial Hospital of Chinese Medicine. These findings require further multi-centered, large-sample, prospective studies to establish whether eGFR, Cys-C, and β 2m as targets of intervention improve the prediction and prevention of ischemic

Table 2
Predictive value of the BP neural networks model for stroke in young adults.

Models	Correct rate (%)	Sensitivity (%)	Specificity (%)	Youden index (%)	Positive Likelihood Ratio (%)	Negative Likelihood Ratio (%)	AUC (95% CI)	P Value
Traditional Vascular Risk Factors Model	63.27	57.30	88.48	45.78	497.51	48.26	0.7915 (0.7498–0.8333)	<.0001
Lipid Metabolism Model	76.17	61.29	95.65	56.94	1409.68	40.47	0.8387 (0.7367–0.9407)	<.0001
Early Renal Injury Model	82.80	100.00	95.24	95.24	2100.00	0.00	0.9803 (0.9398–1.000)	<.0001

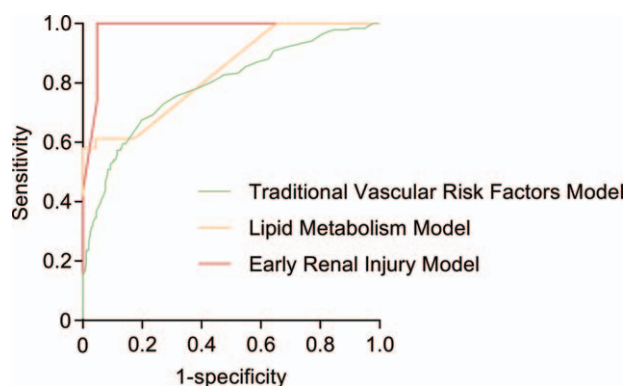


Figure 4. ROC curves of the BP neural networks model for stroke in young adults.

stroke in young adults. The study has the following limitations. First, the results of blood and urine specimens were only the results of specimens at a single time point after admission. Secondly, the treatment administered to patients by physicians might cause acute kidney injury, which might then influence model building and the findings. Thirdly, 92 patients were excluded from the analysis due to missing information. Besides, despite controlling several potential confounders, the possibility of residual uncontrolled confounding was a concern given the design of the study. Also, the performance of the BP neural networks model has generally been recognized in the medical field. However, its selection of input variables, parameter estimation of weighting coefficients and hypothesis testing, as well as medical interpretation of predictor variables needs improvement. We propose animal studies should be conducted to investigate the role and mechanism of early renal injury in cerebrovascular disease. Moreover, whether reducing early renal injury prevents ischemic stroke in young adults remains to be exposed.

5. Conclusion

In conclusion, our findings reveal that the early kidney injury model accurately predicts the risk of stroke in young stroke patients, and has a certain clinical value as a reference for the assessment of ischemic stroke in young adults. Moreover, its predictive power is more robust compared to the traditional vascular risk factor model and the lipid metabolism model. In clinical practice, the renal status of young patients at high risk of stroke should be timely noted to make timely adjustments for treatment options and better address the risk factors and causes of ischemic stroke in young adults. This is essential for the prevention of future strokes.

Author contributions

Conceptualization: Yuyang Chen, Yingqi Mao.

Data curation: Yuyang Chen, Yingqi Mao.

Formal analysis: Yuyang Chen, Yingqi Mao.

Investigation: Yuyang Chen, Yingqi Mao, Xiaoyun Pan.

Methodology: Yuyang Chen, Yingqi Mao, Xiaoyun Pan, Weifeng Jin.

Project administration: Yuyang Chen, Yingqi Mao, Xiaoyun Pan, Weifeng Jin.

Resources: Xiaoyun Pan, Weifeng Jin.

Software: Xiaoyun Pan, Weifeng Jin.

Supervision: Weifeng Jin.

Validation: Weifeng Jin.

Visualization: Weifeng Jin.

Writing – original draft: Yuyang Chen, Yingqi Mao.

Writing – review & editing: Tao Qiu.

References

- [1] Ekker MS, Boot EM, Singhal AB, et al. Epidemiology, aetiology, and management of ischaemic stroke in young adults. *Lancet Neurol* 2018;17:790–801.
- [2] Yesilot N, Putaala J, Bahar SZ, et al. Ethnic and geographical differences in ischaemic stroke among young adults. *Curr Vasc Pharmacol* 2017;15:416–29.
- [3] Ekker MS, Verhoeven JI, Vaartjes I, et al. Stroke incidence in young adults according to age, subtype, sex, and time trends. *Neurology* 2019;92:e2444–54.
- [4] Lo WD, Kumar R. Arterial ischemic stroke in children and young adults. *Contin Minneap Minn* 2017;231:158–80.
- [5] Levey AS, Stevens LA. Estimating GFR using the CKD Epidemiology Collaboration (CKD-EPI) creatinine equation: more accurate GFR estimates, lower CKD prevalence estimates, and better risk predictions. *Am J Kidney Dis Off J Natl Kidney Found* 2010;55:622–7.
- [6] Schöberl F, Ringleb PA, Wakili R, et al. Juvenile stroke. *Dtsch Arzteblatt Int* 2017;114:527–34.
- [7] Cleophas TJ, Cleophas TF. Artificial intelligence for diagnostic purposes: principles, procedures and limitations. *Clin Chem Lab Med* 2010;48:159–65.
- [8] Hathidara MY, Saini V, Malik AM. Stroke in the young: a Global Update. *Curr Neurol Neurosci Rep* 2019;19:91.
- [9] George MG, Tong X, Bowman BA. Prevalence of cardiovascular risk factors and strokes in younger adults. *JAMA Neurol* 2017;74:695–703.
- [10] Shahid R. Risk factors and subtypes of ischemic stroke in young patients: an observational study from a teaching hospital in Saudi Arabia. *Funct Neurol* 2019;34:79–84.
- [11] Barkas F, Milionis H. Treating dyslipidemia for the primary and secondary prevention of stroke. *Semin Neurol* 2017;37:286–93.
- [12] Robinson JG, Williams KJ, Gidding S, et al. Eradicating the burden of atherosclerotic cardiovascular disease by lowering apolipoprotein B lipoproteins earlier in life. *J Am Heart Assoc* 2018;7:e009778.
- [13] Sun Y, Hou X-H, Wang D-D, et al. Apolipoprotein B/AI ratio as an independent risk factor for intracranial atherosclerotic stenosis. *Aging* 2019;11:6851–62.
- [14] Ma Y-H, Leng X-Y, Dong Y, et al. Risk factors for intracranial atherosclerosis: a systematic review and meta-analysis. *Atherosclerosis* 2019;281:71–7.
- [15] Favari E, Thomas MJ, Sorci-Thomas MG. High-density lipoprotein functionality as a new pharmacological target on cardiovascular disease: unifying mechanism that explains high-density lipoprotein protection toward the progression of atherosclerosis. *J Cardiovasc Pharmacol* 2018;71:325–31.
- [16] Chen W, Wu Y, Lu Q, et al. Endogenous ApoA-I expression in macrophages: a potential target for protection against atherosclerosis. *Clin Chim Acta Int J Clin Chem* 2020;505:55–9.
- [17] Zhang B, Yang N, Lin S-P, et al. Suitable concentrations of uric acid can reduce cell death in models of OGD and cerebral ischemia-reperfusion injury. *Cell Mol Neurobiol* 2017;37:931–9.
- [18] Ren J, Dong X, Nao J. Serum cystatin C is associated with carotid atherosclerosis in patients with acute ischemic stroke. *Neurol Sci Off J Ital Neurol Soc Ital Soc Clin Neurophysiol* 2020;41:2793–800.
- [19] Wang Y, Zhang Y, Ma Q, et al. Determination of clinical cut-off values for serum cystatin C levels to predict ischemic stroke risk. *J Stroke Cerebrovasc Dis Off J Natl Stroke Assoc* 2019;28:104345.
- [20] Toyoda K. The cerebro-renal interaction in stroke neurology. *Neurology* 2012;78:1898–9.
- [21] Ninomiya T. Risk of stroke in kidney disease. *Contrib Nephrol* 2013;179:58–66.