ISSN 2706-6819 Vol.6, Issue 7: 1-6, DOI: 10.25236/IJFM.2024.060701

Research progress on evaluation tools for predicting recurrence risk of ischemic stroke in China

Jie Liu[1], Yuying Hao[2, a, *], Rui Song[2], Xin Yang[1], Li Cao[1]

[1]Shaanxi University of Chinese Medicine, Xianyang, 712046, China
[2]Department of Nursing, Affiliated Hospital of Shaanxi University of Chinese Medicine, Xianyang, 712000, China
[a]hyy8908@163com
[*]Corresponding author

Abstract: *By summarizing the overview, evaluation content, effect evaluation and limitations of existing models, this paper reviews some domestic models for ischemic stroke recurrence risk prediction, in order to provide reference for building a more simple, efficient and targeted ischemic stroke recurrence risk prediction model suitable for Chinese population, and at the same time, it can accurately identify patients with ischemic stroke recurrence in the early stage.*

Keywords: *Risk prediction; Ischemic stroke; Recrudescence; Review*

1. Introduction

Stroke, also known as cerebrovascular accident, is an acute focal injury of the central nervous system[1]. Stroke includes ischemic and hemorrhagic subtypes. Ischemic stroke is the most common[2]. In recent years, with the progress of medical technology, the mortality rate of stroke has been decreasing year by year, but it still has the characteristics of high morbidity, high recurrence rate and high disability rate[3, 4]. This not only brings serious disease burden to patients and their families, but also increases the social burden and the health cost invested by the state[5]. Among them, the nerve injury caused by the recurrence of ischemic stroke is more serious and more difficult to treat than the first stroke, and the mortality rate is higher[6]. It is worth noting that studies have shown that timely and effective secondary prevention can reduce the recurrence of ischemic stroke by about 80%[7]. The prediction results can obtain the related influencing factors and risk levels of stroke survivors' recurrence, which can identify high-risk groups early, prevent and reduce the risk of ischemic stroke recurrence, and provide reference for clinical nursing of stroke patients. Therefore, this study summarizes and summarizes the risk assessment tools for ischemic stroke recurrence constructed by Chinese scholars in recent years by consulting domestic and foreign literature.

2. Research status of risk prediction model of ischemic stroke recurrence

2.1 Scale

2.1.1 Essen scale (ERS)

The Eisen stroke risk score scale was first developed on the basis of CAPRIE study of clopidogrel and aspirin in the observation of the efficacy of high-risk ischemic events[8]. In 2013, Yang[9]reviewed the case data of 433 patients. On the basis of ERS, 8 variables were added to prepare a modified stroke risk score(SRS)scale, and a one-year follow-up was conducted. The study showed that the C statistics of ERS and SRS scores predicted within 30 days, 90 days, 180days, and 1year were 0. 617and 0. 638, respectively; 0. 619, 0. 583;0. 626, 0. 572; 0. 608, 0. 615, that is, ERS has a certain predictive value for the recurrence of ischemic stroke within 30 days and 1 year, but has a poor predictive value within 90 days and 180 days. SRS has a certain predictive value for the risk of early and middle recurrence after ischemic stroke. However, the population of the modified scale is only through case review. The clinical applicability of the scale needs to be verified, and prospective studies should be added in the future.

In 2018, Xia[10]developed an improved Essen scale. The study prospectively included 887 patients

ISSN 2706-6819 Vol.6, Issue 7: 1-6, DOI: 10.25236/IJFM.2024.060701

and followed up for 1 year by multivariate logistic regression analysis. The results showed that the AUC values of ERS and modified ERS were 0. 58and0. 70, respectively, suggesting that the modified ERS was superior to ERS in predicting the recurrence of ischemic stroke within 1 year in China. However, since the score is only used for some people with large vascular sclerosis and lacks external verification, the predictive effect and applicability of the score need to be further verified.

2.1.2 Recurrence risk perception scale for patients with ischemic stroke

In 2021, Chinese scholar Lin[11]and others finally formed the first draft of the scale through literature analysis, qualitative interview, expert correspondence, reliability and validity test and other steps, and selected 265cases of stroke patients in tertiary hospitals for reliability and validity test. The scale consists of two parts. The first part is the assessment of the possibility of recurrence, including three items;the second part consisted of 17 items and 3 dimensions. The Cronbach 'sαcoefficient of the total scale was 0. 850, indicating that the scale had good internal consistency. However, due to the small sample size and the limitations of the results, the scale should be verified and optimized by increasing the sample size to increase the universality of the population.

3. Traditional logistic regression model

3.1 Cox proportional hazards regression model

In 2013, an[12]reviewed the clinical data of 408 patients with primary ischemic stroke. Through the multivariate analysis of Cox proportional hazards regression model analysis, six influencing factors were finally include. The individual prognostic index of the model was PI=0. 025×age+0. 681×history of hypertension+0. 395×total cholesterol+0. 973×history+0. 283×(ESRS)+0. 636×disease progression. When the PI value was 2. 289 as the cutoff point, the area under the curve was 0. 827, the sensitivity was 0. 731, and the specificity was 0. 795. Therefore, when the prognostic index is greater than or equal to 2. 289, ischemic stroke has a greater possibility of recurrence. However, the recurrence rate of this model within 1 year and 2 years is inconsistent with the foreign recurrence rate, and no external verification is carried out, which may have certain limitations.

In 2017, in order to further verify the accuracy of the prediction model of ischemic stroke recurrence risk. An[13]divided the patients into 1058cases in the modeling group and 616 cases in the verification group by prospective study. Through Cox regression model analysis, five influencing factors were finally included in the main effect equation of Cox regression model. Based on this, the individual PI equation was PI=0. 265×age+0. 581×heart disease+0. 640×hypertension+0. 515×diabetes+0. 759×TC. The AUC was 0. 817, and the cut-off point was when the PI value was 1. 248, that is, when the PI value was≥1. 248, the patient was more likely to have stroke recurrence within 3 years. However, this model only conducted a 1-year cohort study, and the model fitting effect was not determined. The clinical applicability of this model needs to be further verified.

In 2022, Gao[14]used a multi-center, prospective research design to establish an early warning model of ischemic stroke recurrence based on TCM syndromes. A total of 1741 inpatients with first-episode ischemic stroke were collected. Through the multivariate Cox regression model, combined with the opinions of stroke experts, 16 recurrence influencing factors including TCM syndromes were finally formed. In this study, 60% of the enrolled cases were selected as the training set for model construction, and the remaining 40% were used as the validation data set. The AUC value of the modeling group was 0. 64±0. 02, and the AUC value of the validation model group was 0. 70±0. 03, indicating that the model had a moderate discrimination effect. In order to improve the practicability of the model, a risk score was developed on the basis of the model. Among them, for every 5 years of age, 1 point is added, and the scores corresponding to each influencing factor are calculated. According to the corresponding table of risk probability in different time periods, the ROC curve is divided into three groups of low, medium and high risk to predict the recurrence of stroke. The model integrates TCM syndromes into influencing factors and considers the early warning effect of TCM factors in the disease. However, this model ignores the possible influence of age on the recurrence of ischemic stroke, and its prediction results may be limited. In the future, the population range should be increased.

In 2022, Zhou[15]reviewed the clinical data of 538 patients with ischemic stroke, and used multivariate Cox proportional hazard regression to screen the influencing factors of ischemic stroke recurrence. On this basis, a predictive model equation was constructed and followed up for 5 years. The predictive equation was PI=2. 793×age+10. 444×hypertension+5. 557×diabetes+5. 078×heart disease+4. 874 ×cholesterol. However, due to the limited population included in the study, the

ISSN 2706-6819 Vol.6, Issue 7: 1-6, DOI: 10.25236/IJFM.2024.060701

population should be increased in the future. Prospective studies should be conducted to improve the accuracy of the model.

3.2 Logistic regression model

In order to accurately predict the influencing factors of ischemic stroke recurrence after aspirin treatment, Yu[16]reviewed the data of 1381 patients. Finally, 876 patients were included and followed up for 2 years. The model prediction equation Logit(p)=0. 4693×hyperglycemia+1. 0592× hypertension+ 0. 773×hyperlipidemia-0. 3804×mental status-0. 3661×medication compliance-1. 2531, and the model intercept was-1. 2531. The AUC value of the model is 0. 7529, the positive predictive rate of the model is 68. 41%, and the negative predictive rate is 85. 39%, suggesting that the model has good predictive performance, and after the goodness of fit test, it is suggested that the model has a good fitting effect on the data in general. And the study concluded that hypertension is the main factor in the recurrence of ischemic stroke, which is consistent with the results of Zhou[17]. The model has good predictive efficacy and provides a basis for secondary prevention of patients.

In 2022, Liu[18]through Meta analysis and quality evaluation, 7 risk factors for modeling were finally screened out. The model constructed by Logistic regression analysis was Logit(p)=α+0. 633× smoking+0. 622×heart disease history+0. 771×hypertension +0. 473×diabetes+0. 543×high fibrinogen+0. 8×low/high density lipoprotein+1. 16 statin medication compliance, where α is a constant term. For example, the study reported that the recurrence rate of stroke in China was 5. 48%[19]. At this time, the calculation method of the estimated constant term was 0. 0548/(1-0. 0548). In this model, the most relevant factor for recurrence is the medication compliance of statins, which is consistent with the research results of Huang[20], that is, statins can reduce the recurrence rate of ischemic stroke, so regular use of statins is worthy of attention. The model excludes the limitations of single-center research, and uses high-quality literature review to screen influencing factors, which increases the applicability of the model. The follow-up time of the included literature was different, and the model was not verified. The model may be biased. In the future, the reliability and accuracy of the model should be improved based on evidence-based literature review and patient data collection for model verification.

4. Prediction model of ischemic stroke recurrence risk based on machine learning

With the advancement of computer technology and intelligent medical care, machine learning has begun to be used for disease risk prediction. The artificial intelligence of the computer can be realized by the ability of the machine to learn from the inherent regularity information of the data and obtain new experience and knowledge. The disease risk prediction models constructed by BPNN, SVM algorithm, XGBoost algorithm and other machine learning algorithms are superior to traditional statistical methods[21, 22].

4.1 Column diagram

In 2020, Yuan[23]included 604 patients with primary ischemic stroke. Cox proportional hazard regression was used to construct the model. Finally, five influencing factors were included and a nomogram was developed. The top score of the nomogram is 0-100 points. The intersection of the vertical line and the top score of each influencing factor is the corresponding score of the influencing factor. The scores of all influencing factors are added to obtain the total score. The total score of the nomogram ranges from 0to280. When the score was>245, the risk of recurrence within 1 year was about>50%. The C-index of the model in the modeling group and the test group was 0. 72 and 0. 76, respectively, suggesting that the model has a certain predictive ability. However, this model is a single-center prospective case retrospective study, and there may be bias in the collection of patients' general social data and diagnostic results.

In 2021, Yu[24]reviewed the medical records of 2216 patients . Finally, 1244 patients were included and divided into training group (n=796) and verification group (n=448). The influencing factors were determined by multivariate logistic regression and the model was constructed. On this basis, a nomogram was developed. When the score>210 points, the probability of stroke recurrence>0. 5. However, the nomogram has certain limitations. First, the data are collected through case review, and the results may be biased. Secondly, the model validation population is all patients in the hospital, and external validation should be carried out in the future to ensure the population universality of the

ISSN 2706-6819 Vol.6, Issue 7: 1-6, DOI: 10.25236/IJFM.2024.060701

nomogram.

In 2022, Liu[25]included 821 patients with first-episode ischemic stroke. The histogram was established by R software, and the nomogram of ischemic stroke recurrence based on Cox proportional model was constructed. The total score range of the nomogram is 0-350 points. When the score is greater than 260, the probability of ischemic stroke recurrence within 3 years is greater than 50%;when the score>240, the probability of recurrence within 5 years is about>50%;the C-index of modeling and validation of the model in 3 years of recurrence were 0. 640 and 0. 603, respectively. The C-index of 5-year recurrence was 0. 671 and 0. 635, respectively, suggesting that the nomogram was better than 3 years in predicting 5year recurrence rate. However, this study is a small sample single-center study. The results of Cox analysis may be biased and multi-center verification should be carried out.

4.2 Risk prediction model of ischemic stroke based on artificial neural network

In 2013, Tan[26]constructed a risk prediction model of ischemic stroke recurrence based on neural network. A total of 589 patients were collected. The 16 factors screened by univariate logistic regression were used as input variables to construct three artificial neural models with different number of hidden layers. At the same time, the model is constructed by traditional methods. The results show that the prediction consistency of the single-layer artificial neural network is higher than that of the other two models. Therefore, the single-layer model will be compared with the traditional model in the future. The model was applied to the validation group. The AUC values of the artificial neural network-based ischemic stroke recurrence prediction model and the Logistic traditional model were 0. 787and0. 729, respectively. The sensitivity was 81% and 61. 9%, the specificity was 79. 3%and 72. 3%, and the accuracy was 84. 6% and 81. 7 %. It is suggested that the prediction ability of the model is medium, the model is more robust, and it is superior to the traditional regression model.

In 2022, Lu[27]used the same method to construct a risk prediction model of ischemic stroke recurrence based on blood pressure network. The model was applied to the validation set, showing that the AUC, sensitivity and specificity of the model based on blood pressure network and the traditional model were 0. 796and0. 735, respectively. 82. 1%, 63. 2%;80. 5%and75. 6%, respectively, suggesting that the predictive effect of the model is due to the traditional model.

4.3 Construction of ischemic stroke recurrence prediction model based on extreme gradient boosting (XGboost) algorithm

In 2018, Xu[28]reviewed the data of 6070 patients. Through data cleaning, missing value filling and assignment of categorical variables, a risk prediction model of recurrent readmission within 90 days of ischemic stroke based on Logistic regression and XGboost algorithm was constructed, and the AUC, specificity and sensitivity of the two were compared. The XGboost algorithm suggested that the top six influencing factors were hypertension, red blood cell distribution width, direct bilirubin, length of hospital stay, pulmonary infection, and alkaline phosphatase. ROC curve showed that the AUC based on XGboost and traditional model were 0. 792and0. 739, respectively, suggesting that the predictive effect was better than that of traditional model, which could be of significance for early recurrence of ischemic stroke. In 2021, Zheng[29] used the same method to construct model . Comparing the AUC, specificity, sensitivity and accuracy of the two, the XGboost algorithm can deeply mine the data to obtain the predicted value and the actual data error is minimized, so the prediction value of the model based on XGboost is higher. However, due to the limitations of the selected population, and only through retrospective studies, the applicability of the model to other populations needs to be further verified.

4.4 Prediction model of ischemic stroke recurrence risk based on deep learning

In 2022, Yuan[30]used 1960 electronic cases of ischemic stroke patients . After the removal of repeated values, the processing of missing values of personal information, and the numerical processing of non-numerical content, the importance of attribute reduction information entropy was repeatedly calculated, and finally five major categories of influencing factors were included. This study compared the efficacy of four machine learning models in predicting the recurrence of ischemic stroke. The results show that the accuracy rate and AUC value of the BiLSTM model are 84% and 0. 95, suggesting that the model has a good predictive effect, and the model is generally superior to other models. The model test based on BiLSTM can solve the problem of low accuracy of traditional models, which is consistent with the research results of Wu[31]. The scholar compared the traditional model. The

ISSN 2706-6819 Vol.6, Issue 7: 1-6, DOI: 10.25236/IJFM.2024.060701

results show that the model based on machine learning has a stronger ability to identify people with recurrent stroke within 1 year. At the same time, the BiLSTM model calculates the influencing factors with the strongest correlation with the recurrence of ischemic stroke through the attribute reduction information entropy algorithm, which can accurately identify the recurrence risk of patients in advance.

5. Comparison of risk model prediction

A total of 14 studies were included. In terms of research design, the models were mainly retrospective studies and prospective studies. Retrospective studies collect data conveniently, but there may be data missing, resulting in bias in the prediction effect of the model. Prospective studies can collect complete patient data, but there is a certain rate of loss of follow-up. In terms of model construction methods, three[23-25]studies constructed a nomogram, which could more intuitively reflect the level of recurrence risk;among them, 5studies[26-30]constructed models by computer and presented models by Internet technology, which can effectively evaluate the risk of recurrence of patients and provide convenience for the evaluation of medical staff. Therefore, in the future, the risk prediction model should be associated with the medical platform to provide more convenient risk prediction.

In terms of model validation, only one study supplemented external validation[13], and other models have not yet described whether external validation is performed. Therefore, in the study of model construction, the construction method of high-quality model should be followed, internal and external validation should be performed, and predictive performance and clinical applicability should be improved.

6. Summary and prospect

In summary, there are many domestic models for predicting the recurrence risk of ischemic stroke, but they all have certain limitations. Most of the models are single-center research designs. In future research, multi-center prospective research and external verification should be carried out to improve the applicability and accuracy of the model in the population. The risk prediction model of ischemic stroke recurrence based on machine learning can solve the problem of accuracy of traditional models. However, due to the complex use of the model and the lack of clinical applicability, in the future, it should be dynamically associated with the medical records or patient information of the hospital diagnosis and treatment platform, automatically analyze the prediction model, or create the Internet+nursing platform of the risk prediction model.

References

[1] Wang K, Shi Q, Sun C, et al. A machine learning model for visualization and dynamic clinical prediction of stroke recurrence in acute ischemic stroke patients: A real-world retrospective study. Front Neurosci. 2023, 17:1130831.

[2] Krishnamurthi RV, Ikeda T, Feigin VL. Global, Regional and Country-Specific Burden of Ischaemic Stroke, Intracerebral Haemorrhage and Subarachnoid Haemorrhage: A Systematic Analysis of the Global Burden of Disease Study 2017. Neuroepidemiology. 2020, 54(2):171-179.

[3] Mao C, Li M. Stroke Recurrence as a Challenge for Countries. JAMA Netw Open. 2022, 5(6): e2219698. Published 2022 Jun 1.

[4] Zhang Zhenxiang, Guo Yunfei, Lin Beilei, et al. Qualitative study on the characteristics and related factors of recurrence risk perception in stroke patients [J]. China General Practice, 2021, 24 (27): 3494-3500.

[5] Liu J, Yang Y, Yan K, et al. [Development and validation of nomograms for predicting stroke recurrence after firstepisode ischemic stroke]. Journal of southern medical univercity. 2022 Jan 20; 42(1): 130-136. Chinese.

[6] Lu T, Wang Y. Prediction Model Construction for Ischemic Stroke Recurrence with BP Network and Multivariate Logistic Regression and Effect of Individualized Health Education. Comput Math Methods Med. 2022, 2022:4284566.

[7] Hankey GJ. Secondary stroke prevention. Lancet Neurol. 2014, 13(2):178-194.

[8] CAPRIE Steering Committee. A randomised, blinded, trial of clopidogrel versus aspirin in patients at risk of ischaemic events (CAPRIE). CAPRIE Steering Committee. Lancet. 1996, 348(9038): 1329-1339.

[9] Weimar C, Goertler M, Röther J, et al. Predictive value of the Essen Stroke Risk Score and Ankle

ISSN 2706-6819 Vol.6, Issue 7: 1-6, DOI: 10.25236/IJFM.2024.060701

Brachial Index in acute ischaemic stroke patients from 85 German stroke units. J Neurol Neurosurg Psychiatry. 2008, 79(12):1339-1343.
[10] Yang ZY, Shi YH, Ren CC. Predictive value of modified Essen score for stroke recurrence [J]. Chinese Journal of Cerebrovascular Diseases (Electronic Edition), 2013, 7 (01):367-375.
[11] Lin BL, Zhang ZX, Guo YF, et al. Development and reliability and validity test of recurrence risk perception assessment scale for stroke patients [J]. Chinese Journal of Nursing, 2021, 56 (11): 1666-1671.
[12] An YC, Wang YX, Zhang J, et al. Construction of a predictive model for ischemic stroke recurrence [J]. Rehabilitation theory and practice in China, 2013, 19(03):210-213.
[13] An YC, Wang YX, Cheng Y. Construction of ischemic stroke recurrence model and its prospective predictive analysis [J]. Chinese Journal of Geriatric Cardiovascular and Cerebrovascular Disease, 2017, 19 (07):685-688.
[14] Gao Y, Xie YM, Wang ZF, et al. Construction and verification of early warning model of ischemic stroke recurrence based on TCM syndromes [J]. Chinese Journal of Traditional Chinese Medicine, 2022, 37(02):686-692.
[15] Zhou JJ, Wang CQ, Zhu SK, et al. Influencing factors of ischemic stroke recurrence and its prediction model construction: Based on 5-year follow-up data [J]. Practical Journal of Cardio-cerebro-pulmonary Vascular Disease, 2022, 30(02):13-17.
[16] Yu Y, Zhang X, Fu Y. Influencing factors and predictive models of stroke recurrence in patients with ischemic stroke treated with aspirin [J]. Chinese Journal of New Drugs and Clinical Sciences, 2019, 38(11):681-686.
[17] Zhuo Y, Wu J, Qu Y, et al. Clinical risk factors associated with recurrence of ischemic stroke within two years: A cohort study. Medicine (Baltimore). 2020, 99(26):e20830.
[18] Liu YH, Guo H, Gong SH, et al. Construction of a risk prediction model for recurrent acute ischemic stroke [J]. Chinese Journal of Geriatrics, 2022, 42(11): 2612-2618.
[19] Wang LD, Peng B, Zhang HQ, et al. Summary of ' China Stroke Prevention Report 2020'[J]. Chinese Journal of Cerebrovascular Disease, 2022, 19 (02): 136-144.
[20] Huang J, Jiang ZL, Shen L. Study on the relationship between small dense low density lipoprotein level and statin dose and recurrence of ischemic stroke [J]. Journal of Hebei Medical University, 2021, 42 (08):950-955.
[21] Fernández-Cadenas I, Mendióroz M, Giralt D, et al. GRECOS Project (Genotyping Recurrence Risk of Stroke): The Use of Genetics to Predict the Vascular Recurrence after Stroke. Stroke. 2017, 48(5):1147-1153.
[22] Zhang Z, Xu G, Cai B, et al. Genetic Variants in MicroRNAs Predict Recurrence of Ischemic Stroke. Mol Neurobiol. 2017, 54(4):2776-2780.
[23] Yuan K, Chen J, Xu P, et al. A Nomogram for Predicting Stroke Recurrence among Young Adults. Stroke. 2020, 51(6):1865-1867.
[24] Yu XF, Yin WW, Huang CJ, et al. Risk factors for relapse and nomogram for relapse probability prediction in patients with minor ischemic stroke. World J Clin Cases. 2021, 9 (31):9440-9451.
[25] Liu J, Yang YL, Yan Ke, et al. The nomogram can predict the recurrence of patients with first ischemic stroke [J]. Journal of Southern Medical University, 2022, 42 (01):130-136.
[26] Tan Y, Geng DQ, Huang SP. To establish a prediction model of ischemic stroke recurrence by artificial neural network [J]. China Health Statistics, 2013, 30 (05): 687-689.
[27] Lu T, Wang Y. Prediction Model Construction for Ischemic Stroke Recurrence with BP Network and Multivariate Logistic Regression and Effect of Individualized Health Education. Comput Math Methods Med. 2022, 2022:4284566.
[28] Xu Yuan, Ma Jianyong, Ge Yanqiu, Li Min, Ren Yuankai, Yi Yingping. Analysis of the predictive effect of XGboost model on the risk of recurrent readmission within 90 days after discharge of ischemic stroke [J]. Chinese Journal of Neuromedicine, 2018, 17 (8):813-818.
[29] Zheng WL. Construction of ischemic stroke recurrence prediction model based on XGboost algorithm [D]. Zunyi Medical University, 2021.
[30] Yuan S. Prediction of ischemic stroke recurrence risk based on deep learning [D]. Huazhong University of Science and Technology, 2021.
[31] Wu ML. Research on predicting the recurrence of ischemic stroke based on machine learning [D]. Shanxi Medical University, 2022.