Journal of Traditional Chinese Medical Sciences (2014) 1, 9–19

HOSTED BY
ELSEVIER

Available online at www.sciencedirect.com
ScienceDirect
journal homepage: http://www.elsevier.com/locate/jtcms

A new way to analyze the traditional Chinese medicine syndrome: heat toxin syndrome in cerebral infarction

Zhichen Zhang [a], Shaozhen Ji [b], Xuejie Yu [b], Xianglan Jin [b], Liping Zhang [b], Rongjuan Guo [a], Hong Zheng [a], Mingqi Wang [b], Yunling Zhang [a,*]

[a] *Dongfang Hospital, Beijing University of Chinese Medicine, Beijing, China*
[b] *Beijing University of Chinese Medicine, Beijing, China*

Received 15 March 2014; accepted 15 June 2014

Available online 18 December 2014

KEYWORDS
Acute cerebral infarction;
Heat toxin syndrome;
Turning point;
Data mining

Abstract *Objective:* To establish a diagnostic system for heat toxin syndrome of acute cerebral infarction. Based on this toxin syndrome diagnostic system, the general principles of heat toxin development will be uncovered, and the critical turning point at which the heat toxin syndrome occurs will also be explored.
Methods: In this study, a total of 271 hypertension patients with cerebral infarction within 72 h were recruited from the Affiliated Dongfang Hospital of the Beijing University of Chinese Medicine, the Affiliated Dongzhimen Hospital of Beijing University of Chinese Medicine, the Affiliated Renmin Hospital of Peking University, the Second Affiliated Hospital of Tianjin University of Traditional Chinese Medicine, the Affiliated Hospital of Shandong University of Traditional Chinese Medicine, the Affiliated Hospital of Changchun University of Traditional Chinese Medicine, the Affiliated Hospital of Hebei University of Traditional Chinese Medicine and China Meitan General Hospital from August, 2008, to December, 2009. The patients' Chinese medical information was recorded on days 1, 3, 5, 7, and 14 during their hospitalizations. The medical records were recorded according to traditional Chinese medicine (TCM) theory and included the serum marker levels at the beginning and at the end of the trial. The time line was also analyzed.
Results: The level of Hs-CRP, PAG, NSE, OX-LDL, and MMP-9 were abnormal and, were higher in CI patients compared to hypertension patients. In the study of the heat toxin diagnosis system, according to the entropy clustering results, 30 combinations of the medical information can be sorted into the traditional syndromes, but 13 combinations cannot be sorted. To obtain more precise symptoms related to the heat toxins, a logistic regression equation was set up with the

* Corresponding author. Tel.: +86 10 67689634; fax: +86 10 67691949.
E-mail address: yunlingzhang2004@163.com (Y. Zhang).
Peer review under responsibility of Beijing University of Chinese Medicine.

http://dx.doi.org/10.1016/j.jtcms.2014.11.008

variables from the unsorted medical information; the dependent variables were fever and BP fluctuation. Weighted variables were obtained. MLP analysis demonstrated that the diagnosis model was stable and precise. The accuracy reached 83.82%. The ROC test showed that seven points of the diagnosis system was the best cutting point, with a sensitivity of 0.857 and a specificity of 0.955. Progressing stroke was related to heat toxin syndrome. When the turning point appeared, the combination of symptoms, such as coma, aphasia, gummy eyes, and halitosis, predicted the deterioration or recovery of CI. The heat toxin syndrome existed in every subtype of CI; however, the observed heat toxin levels were highest in PACI and lowest in LACI. Meanwhile, blood and sputum stasis syndromes transformed into heat toxicity were one source of heat toxin syndrome.

Conclusion: Heat toxin syndrome, as well as qi/blood/sputum stasis, co-existed in the CI patients, and the transformation frequently appeared during the process. Three to five days after the onset of CI was the turning point, at which time several combinations of medical indicators make it possible to predict the development of CI.

Introduction

Acute cerebral infarction, also known as ischemic stroke or stroke in traditional Chinese medicine, is the sudden loss of blood circulation in the brain, causing a corresponding sudden loss of neurological function. Stroke is a huge threat to human health, which can cause permanent nerve damage. Stroke can also cause disability or death immediate diagnosis and treatment.

Hypertension is an independent risk factor for acute cerebral infarction. Hypertension marks the start of chronic damage to the cerebrovascular wall, and laboratory markers of the syndrome that manifest during an acute cerebral infarction are observed latent at this stage.

There are no effective prevention and treatment methods for ischemic stroke at present. Although administering thrombolytics is an effective method in acute cerebral infarction patients, that efficacy is limited to a narrow window of 3 h after the first stroke symptoms. Due to this limitation, most patients miss the treatment period, and it is estimated that the proportion of stroke patients who benefit from thrombolysis is less than 5%. Therefore, early warning of ischemic stroke symptoms is essential. Stroke is a major challenge for both traditional Chinese and Western medicine. Western medicine places emphasis on urgent needs in clinical research and the exact meaning of biochemical markers to evaluate the condition of the patient, for the early detection and the early treatment of stroke. Traditional Chinese medicine, in contrast, emphasizes understanding the syndrome and therapies are based on the dynamic state of a disease. Defining the type of syndrome before and after stroke is an essential prerequisite for the prescription of Chinese medicine and for providing a fundamental theory.

In both traditional Chinese and Western medicine, the analysis and diagnosis of the disease reflect an understanding of its root causes. Many modern diseases are complex, and multiple factors must be considered. In the Chinese medicine theory, external and internal pathogenic processes, give rise to pathological factors, such as wind, fire, phlegm and blood stasis. Single or complex syndrome differentiation and treatment through traditional Chinese medicine has achieved certain effects, but further improvement of the curative effect is difficult, and the efficacy achieved in clinical trials cannot be repeated. The academician Yongyan Wang, who had previous clinical experience, while reviewing the disease process, put forth the Chinese medicine theory of the "poison damaged brain"[1] in stroke patients, which obtained better curative effects in clinical practice.

Internal poison damage in the collaterals is the result of a dynamic relationship between the etiology of the disease and the pathogenesis of the disease and is also an important reason for the disease occurrence. Poison damage involves the interaction of a variety of pathogenic factors and the pathogenic transformation of multiple components. To examine internal poison damage in the collaterals, under the guidance of the whole theory, modern life science technology has become the best way to detect its causes. A large quantity of data has been collected through randomized clinical trials and analyzed by "system integration" using the modern computer and information science theory in conjunction with experimental trials to analyze and recognize different presentations of this dynamic process. Thus, such techniques as data mining and the decision tree method have assisted in the diagnosis of TCM syndrome types.[2–4] These approaches help to confirm the specific state of internal poison damage in the collaterals accurately and to execute the planned clinical treatment and will also become the crossover point between the traditional Chinese medicine and Western medicine and between traditional and modern research.

The turning point is the change of run trend or run rate in the process of disease development. In the mathematical field, this term refers to the connection point of the convex curve and the concave curve. When a point on the function of the image that makes the second derivative of the function zero, and the third derivative is not zero, this point is the turning point of the function. In everyday life, the term "turning point" is used to illustrate the decline or fall of certain circumstances after they have risen for a given length of time. Introducing mathematical theory to the study of traditional Chinese medicine theory allows a

good interpretation of the characteristics of the evolution of the etiology and pathogenesis of disease.

The diversity of the etiology of stroke, the differences in patients' responses, and their different behaviors in different times and spaces form the complex etiology system and complex clinical symptom system. Over time, each of the two systems (the complex etiology system or complex clinical symptom system) has a corresponding role to perform, but both can be reflected in the evolution of the pathogenesis. Studies of the complex symptom system are expected to reveal the law of evolution of parts of its pathogenesis. However, study of complex symptom systems requires: 1) establishing a "complexity" way of thinking and the methods and techniques to develop and address complex data; and 2) epidemiological investigation of stroke based on clinical epidemiological methods. Thus, the use of dimensionality reduction methods is critical for dealing with the complexity systems to explore the law of pathogenesis evolution of stroke disease, and only the maturity of modern data mining technologies allows this type of research. How the syndrome elements change at different times is a clue to research incentives and changing rules, and we will be able to define the quantitative to qualitative progression of a syndrome as it develops. This process is an important turning point between the onset and the different possible outcomes and has important strategic significance for the diagnosis and treatment of the disease.

The current study has improved the understanding of the syndrome elements and, its biological indicators and established an identification system and the turning point of acute cerebral infarction.

Materials and methods

Diagnostic criteria

Diagnostic criteria of Hypertension: From the 2005 Chinese guidelines for the management of hypertension, hypertension is defined by systolic blood pressure $\geq$140 mmHg and/or diastolic blood pressure $\geq$90 mmHg, except in the case of secondary hypertension. Hypertensives are further divided into low-, mid-, and high-risk groups.

Hypertension low-risk group: hypertension without high cholesterol and (or) high blood sugar; mid- and high-risk groups: hypertension with high cholesterol and (or) high blood sugar.

Diagnostic criteria of TCM Stroke: the State Administration of Traditional Chinese Medicine, the emergent encephalopathy coordination of scientific research group developed Industry Standard Diagnosis and Treatment Evaluation Criteria of Stroke.[5]

Diagnostic criteria of Western medicine for acute cerebral infarction: Refer to the diagnostic criteria of acute cerebral infarction from Various Types of Cerebrovascular Disease Diagnosis[6] revised by the 4th National Cerebrovascular Disease Conference of the Chinese Medical Association in 1995. Atherosclerosis cerebral infarction, lacunar infarction and asymptomatic cerebral infarction were selected.

Main points of the diagnosis of atherosclerotic thrombotic cerebral infarction: (1) onset often occurs in a quiet state; (2) most have no obvious headache or vomiting; (3) a slow disease progression, and mostly gradual or stepwise onset that is related to cerebral atherosclerosis (this pattern can also be observed in arteritis and blood disease); (4) consciousness or mild conscious disturbances one to two days after the onset; (5) symptoms and signs of the internal carotid artery system and/or the vertebral basilar artery system are apparent; (6) the event is diagnosed with CT or MRI; and (7) the cerebrospinal fluid from a lumbar puncture is often free of blood.

Main points of the diagnosis of Lacunar Infarction: (1) it is caused by hypertension arteriosclerosis with acute or subacute onset; (2) most patients have no conscious disturbances; (3) the event is diagnosed with CT or MRI; and (4) the clinical manifestations are not severe. The common manifestations are pure sensory stroke, pure motor hemiparesis, ataxic hemiparesis syndrome, dysarthria/clumsy hand syndrome or sensorimotor stroke.

Main points of the diagnosis of asymptomatic cerebral infarction: (1) no brain or retinal symptoms of vascular disease are found; (2) it can only confirmed by imaging; and (3) the clinical diagnosis is given as the situation dictates.

Inclusion and exclusion criteria

The patients included in this study met the following criteria: (1) their cases were in accordance with the diagnostic criteria above; (2) their CI onset occurred within 72 h; (3) head CT or MRI confirmed no new intracranial infarct or hemorrhage; (4) their ages were between 35 and 75 years old, irrespective of gender; and (5) their informed consent was obtained before they participated in this study. The study was approved by the local ethics committee.

The patients excluded from the study were those who: (1) were diagnosed with transient ischemic attack (TIA); (2) were found to have cerebral hemorrhage or subarachnoid hemorrhage; (3) were women in pregnancy or breastfeeding; (4) had a combination of severe diseases of the liver, the kidney, or the hematopoietic system and cancer, or of immune system diseases and bone and joint disease; (5) had mental deficiency or mental illness that prevented their cooperation with the examination; (6) had other diseases in acute phases; (7) were participating in clinical trials of drug research; (8) had CI onset more than 72 h previously; (9) had apoplexy caused by a brain tumor, cerebral trauma or hematologic disease, as confirmed by auxiliary examination; (10) had cerebral embolism caused by rheumatic heart disease, coronary heart disease, or other heart disease with atrial fibrillation; (11) had mental deficiency or mental illness; (12) had incomplete research data.

Clinical data

A total of 271 cases of patients with hypertension complicated with acute cerebral infarction from the Departments of the Affiliated Dongfang Hospital of Beijing University of Chinese Medicine, the Affiliated Dongzhimen Hospital of Beijing University of Chinese Medicine, the Affiliated Renmin Hospital of Peking University, the Second Affiliated Hospital of Tianjin University of Traditional Chinese Medicine, the Affiliated Hospital of Shandong University of Traditional Chinese Medicine, the Affiliated Hospital of

Changchun University of Traditional Chinese Medicine, the Affiliated Hospital of Hebei University of Traditional Chinese Medicine and China Meitan General Hospital from August, 2008, to December, 2009 were recruited in this study. All patients were examined by CT or MRI. Consent was received from all of the participants.

Observation methods of TCM symptoms

The diagnoses of syndromes were made by experienced doctors of TCM. A unified medical questionnaire of symptoms was formulated from the literature research and expert investigation. The symptoms were observed by neurologists, who had been specially trained for this program, through observation, listening, interrogation, and pulse-taking along with the medical history of patients at 1, 3, 5, 7, and 14 days from the onset of CI.

Syndrome differentiation

Because hypertension is the most common risk factor for cerebrovascular disease, and there is no clear TCM syndrome element scale, the study used part of Standards of Syndrome-Differentiated Diagnosis in "Apoplexy"[5] to define the syndrome elements. Heat syndrome and blood-stasis syndrome were assessed by the Standards of Syndrome-Differentiated Diagnosis in "Apoplexy". The patients with scores $\geq$7 in both the heat syndrome and blood stasis syndrome characteristics were assigned to the heat-blood stasis syndrome group. The participants with scores $\geq$7 in heat syndrome characteristics were assigned to the heat syndrome group. The patients with scores $\geq$7 in the blood stasis syndrome characteristics went to the blood stasis syndrome group. The patients with scores $<$7 in both the heat syndrome and blood stasis syndrome characteristics were assigned to the non-heat-blood stasis syndrome group. The basic data and the frequency of occurrence were analyzed.

The level of serum markers in blood

The serum markers were tested as routine by the clinical laboratory at the Affiliated Dongfang Hospital of Beijing University of Chinese Medicine and at the Affiliated Dongzhimen Hospital of Beijing University of Chinese Medicine. The concentration of OX—LDL was tested by the enzyme-linked immunosorbent (ELISA), using FLx800 (BioTek, America) Fluorescence Microplate Reader.

The clinical intervention programs for acute cerebral infarction

The patients who met the standard of hypertension complicated with acute cerebral infarction were randomly divided into kudiezi injection group and xuesaitong-injection group with a ratio of 1:1. Each group member received kudiezi injection or xuesaitong injection by intravenous infusion every day for 14 days. The other basic treatments for each group were the same. The kudiezi injection (Chinese Drug Approval Number Z20025449, Shuangding Pharmaceutical Co. Shenyang, Ltd.) was given as a 250 ml diluted injection of 20 ml kudiezi in 5% glucose or 0.9% sodium chloride. The xue-sai-tong injection (Chinese Drug Approval Number Z32020419, Xuzhou Pharmaceutical Co. Xuzhou, Ltd.) was given as a 250 ml diluted injection of 400 mg xue-sai-tong in 5% glucose or 0.9% sodium chloride. Through the above-mentioned methods, syndrome identification, serum marker testing, and clinical interventions, the whole process of low risk, middle risk and high risk to acute cerebral infarction were simulated. Through comparing the results of the three studies, we picked out the key information about poison damage of the brain in acute cerebral infarction.

Diagnostic model of the heat toxin syndrome

Assessment of the heat toxin syndrome

Because the intensity of Chinese medical diagnostic information was high and it is widely distributed, we first classified the Chinese medical diagnostic information in the analysis. Compared with traditional cluster analysis, the entropy clustering method in the study of Chinese medical diagnostic information of acute cerebral infarction could not only be applied to linear data but also to non-linear data, which are based on the complex system of entropy. Because the entropy partition would not have rigid linear segmentation, it was especially suitable for complex multivariate and multi-level data. The combination of the four diagnostic methods' information was the study object. The characteristics of toxin damage were mined to find a reasonable and representative combination of the information gleaned through the four diagnostic methods.

The definition of mutual information based on information entropy can measure any statistical correlation between variables, and it has been widely applied to various fields. On the basis of mutual information, the proposal of the concept of a correlation coefficient can overcome the shortcoming that the correlation coefficient is symmetrical and cannot be directly compared.

The combination of TCM symptoms

Because the damage caused by toxins was diffuse and it was difficult to classify one specific Chinese medical diagnosis, the research group selected indicators closely related to vital signs (blood pressure, body temperature, heart rate and respiration) and classified them.

The establishment of the Chinese medical diagnostic information

The neural network is a new and rare method for epidemiology research. Any problem related to pattern recognition or pattern clustering, especially when it cannot be solved by existing methods or the effect is not good, can be analyzed by neural networking. The researcher simply selects the input and output, and the rest can be performed by the neural network. It can not only avoid the subjectivity and thinking mode of the brain's decision-making but also it can eliminate fatigue of the brain decision-process. Compared with previous methods, its advantage is obvious and its potential could be developed by the majority of epidemiologists. The TCM indicator is ambiguous, so we need a tool to solve such complex problems. One important

advance in scientific research is the new theory and methods to solve complex systems problems. The comprehensive integrated study system, from qualitative to quantitative, is undoubtedly an important way to study large-scale, complex systems. The combination of the machine system, the expert system and the knowledge system constitutes a highly intelligent man—machine combination. The neural network is an important methodological tool to study complex systems.

These four diagnostic methods cover the symptoms and signs. They reflect the characteristics of heat excess, which are similar to heat toxin, however, further proof was needed to confirm the relevance and overlap of these paradigms.

We constructed a six hidden layer MLP network to solve the problem of the heat toxin syndrome diagnosis model. The input variables of the network were information derived from the four diagnostic methods and vital signs with thermal potential. The input factors were: a sudden change of blood pressure (variation exceeded 30% of the original blood pressure), and the thermal potential puzzled (fever puzzled). The input covariates were: fainting, aphasia, shortness of breath, quick pulse, halitosis, gum in the eyes, dark face, yellow sputum. The MLP network analysis methods and the learning process included: 10 units of the input layer, 6 hidden layers, and two output layer units. In 271 cases of acute cerebral infarction (factors and covariates complete cases), 70% of the cases were used for the learning sample, while 30% of the cases were used for the test sample. All of the cases were randomly arranged. The selection of the parameters: the learning cycle = 2000, learning speed = 1.0 decreased to 0.1, the momentum factor = 0.5 down to 0.1, and a random number generator was used to set small random numbers to each neuron weight coefficient.

Establishment of the dialectical system of heat toxin syndrome of acute cerebral infarction

Through the research above, four diagnostic methods for the identification of heat toxin syndrome of acute cerebral infarction and the importance of syndrome factors were established, but the specific significance of the importance scores of the four diagnostic methods needs further study. Therefore, this part of the study increased the sample size further, extending the syndrome factors back to the model-like cases to verify and confirm the diagnosis scores by the ROC (receive operating characteristic) curve.

The study on turning point in the pathogenic process of heat toxin syndrome of acute cerebral infarction

To study the dynamic evolution of the macro characterization and the micro indicators of acute cerebral infarction pathogenesis during days 1—14 after the onset of stroke, the heat toxin scores of patients were compared before and after treatment, and pictures were drawn to portray the trend according to the time of the order of development. We then analyzed the relationship between the outcome and the development of the disease process, and looked for the key turning point in the pathogenesis progression of heat toxin syndrome of acute cerebral infarction, starting from different outcomes of patients with heat toxin syndrome.

Statistical methods

Statistical analyses were performed with the SPSS software package (version 17.0). The counting variables were expressed as means ± SD. The measurement variables were described by frequency and percentage; $\alpha = 0.05$ for inspection standards, $p < 0.05$ was considered a significant difference, and $p < 0.01$ was considered a very significant difference.

Results

General data of enrolled cases

Through the comparison above, the average age of the patients with acute cerebral infarction was significantly higher than that of the patients with hypertension. In terms of the gender, the study included more males than females. For patients with acute cerebral infarction included in this study, the average onset time was 27.58 h, which corresponds to the early stage. The systolic blood pressure of patients with acute cerebral infarction was significantly higher than that of patients with hypertension, which shows that patients with acute cerebral infarction usually had poorly controlled hypertension. It also could be the reason that the blood pressure elevated due to acute cerebral infarction. The hypertension history of the patients acute cerebral infarction was significantly longer than that of the patients with hypertension, which showed that the vascular damage was sustained in conditions of hypertension, and the damage gradually increased over time (Table 1).

Results of the classification basis

In the study of the comparison between hypertension and acute cerebral infarction, the patients with acute cerebral infarction had Chinese medical diagnosis combination of

Table 1 General information of patients with acute cerebral infarction.

Demographic and clinical data		
	Heat toxin syndrome	Non-heat toxin syndrome
Cases	87	184
Age (y)	65.94 ± 4.62	63.7 ± 14.62
Gender (male/female)	48/39	110/74
Body temperature (°C)	37.63 ± 3.28	36.89 ± 1.31
Onset time (hours)	26.43 ± 2.11	27.58 ± 2.03
Systolic blood pressure (mmHg)	158.94 ± 15	153.31 ± 22.47
Diastolic blood pressure (mmHg)	89.58 ± 3.65	88.8 ± 14.38
Resting heart rate (beats/min)	83.09 ± 9.73	75.51 ± 10.31
Body mass index (BMI)	25.54 ± 11.49	25.12 ± 12.12
History of hypertension (y)	3.3 ± 1.25	7.10 ± 2.29

Table 2 Basic parameters of the MLP network model.

Factor	Covariate	The number of units	The covariates' rescaling method	Hidden layers	The number of units in the hidden layer	Activation function	The dependent variable of the output layer	The number of units	Activation function	Error function
Sudden change of blood pressure (variation exceeded 30% of the original blood pressure) SMTZ14	Fainting ZFZ6	12	Standardization	1	6	Hyperbolic tangent	Heat toxin yes or no	2	Softmax	Cross-entropy
The thermal potential puzzled (fever puzzled) SMTZ3	Aphasia ZFZ12									
	Shortness of breath SMTZ5									
	Quick pulse SMTZ21									
	Halitosis JCZ16									
	Gum in eyes JCZ15									
	Dark face JCZ21									
	Yellow sputum JCZ29									

strong heat symptoms. There was a certain correlation between continuous fever and heat symptoms in vital signs. We analyzed the relationship between continuous fever and other vital signs. The results are as follows (Supplemental materials Table 2).

We dynamically compared the patients' information which correspond to both the thermal potential puzzled and sudden change of blood pressure, we can find that within 14 days after the onset of acute cerebral infarction, they are closely related and their correlation is stable (Supplemental materials Fig. 1).

According to the body temperature, with blood pressure linkage as the basis of classification, a LOGISTIC regression equation can be created. We performed entropy clustering combinations of four diagnostic methods' results, combined with the relationship of body temperature and blood pressure, and establishing the regression equation. The results are as follows (Supplemental materials Table 3 and Table 4).

After the inspection of the regression equation above and the standardized transformation of the partial regression coefficient, a combination of the information from four diagnostic methods that reflects thermal potential puzzled and a sudden change of blood pressure was generated. The following study will verify the rationale of this combination of four diagnostic methods' information with thermal potential and toxic insult features from a multifaceted approach that includes aspects of both Chinese and Western medicine.

Results of validation of critical four diagnostic methods information identification of heat toxin syndrome of acute cerebral infarction

Basic information of MLP network model (Table 2)

Composition of training samples and test samples

We used a six-hidden-layer MLP network via input variables, such as factors and covariates in the context of network learning, and the results are shown in the following table (Table 3).

Learning and prediction accuracy of the MLP network model

The 271 samples of cases involved in the study were divided up, with 70% used for the study, and 30% for the test. All the

Table 3 MLP model training and proportion of validation sample.

Information of cases		
	The number of samples (*n*)	Percentage (%)
Training sample set	184	67.9
Validation sample set	87	32.1
Number of valid samples	271	100
Exclude the number of samples	0	
Total	271	

Table 4 Results of the MLP prediction accuracy.

	Observed results	Forecast results		Correct percentage
		No	Yes	
Training sample set	No	114	35	76.51%
	Yes	32	25	43.86%
Validation sample set	No	57	11	83.82%
	Yes	22	29	43.14%

Note: total number of samples is 271.

cases included in the study were effective samples. The heat toxin syndrome diagnosis model simulated by the MLP network has a high accuracy (training sample set: 76.51%; test sample set: 83.82%). The results are shown in the table above (Table 4).

MLP network diagram as follows (Supplemental Materials: Fig. 2).

Probability of predictive simulation as the figure (Supplemental materials Fig. 3).

The abscissa indicates whether heat toxin syndrome is established or not, from left to right is NO and YES. The ordinate represents the prediction accuracy, it is generally believed that the boxplot portion greater than 0.5 or more is the correct part of a model, so it can be observed that the part of the model of prediction accuracy slightly higher than 0.5 prompts a good prediction function of the model.

Information gain in MLP network

For an established neural network, each neural node contains important information; the neural network is to judge the presence or absence of heat toxin syndrome. The form of the established model is similar to the form of "if it contains certain information, heat toxin syndrome set up; otherwise, heat toxin syndrome does not set up". Thus, the question is which property (variable, derived from any of the four diagnostic methods' information) is suitable or not to act as the concept these nodes need to borrow from information theory. We use a statistic, "Information Gain", to measure the ability of a property to distinguish the data samples. The greater the information gain is for the property of each node the more concise the network; for example, a network can be understood as, if the thermal potential puzzled and then heat toxin syndrome set up, then discuss the points of quick pulse, shortness of breath. It is very valuable to use thermal potential puzzled as the node of the network at this moment. The total number of the four diagnostic methods' information and vital signs involved in the MLP network built by the institute is 10, which are the factors and covariates in the MLP network. The information gain results are graphed (cumulative gain chart) as follows (Supplemental materials Fig. 4).

The cumulative gain chart aimed at the percentage of the total number of observations, shows the percentage of the total number of observations in the specified category "gain". For example, on the "yes" curve, the first point is located at (10%, 30%), which indicates that if all the observations are reordered in accordance with the prediction virtual probability "yes", the top 10% will contain approximately 30% of the observations (heat toxin syndrome patients) which actually have a category "yes". Likewise, the top 20% will contain approximately 50% of the heat toxin syndrome patients; the top 30% will contain 70% of the heat toxin syndrome patients (Table 5).

Table 5 The importance rank of dependent variables and covariates in the MLP network.

Variable name	Importance	Normalized importance (%)
Sudden change of blood pressure	0.148	93.90
Thermal potential puzzled	0.158	96.00
Fainting	125	79.30
Aphasia	0.034	21.80
Gum in eyes	0.132	83.30
Halitosis	0.117	74.00
Dark face	0.053	33.30
Yellow sputum	0.069	43.40
Shortness of breath	0.035	22.00
Quick pulse	0.078	49.40

The diagonal is the baseline curve. If 10% of the observations from the ratings data set were randomly selected, the "gain" of observations with the category "yes" will be approximately 10%. The farther the curve is above the baseline, the greater the gain is.

Therefore, what content do we need to gain? It depends on the error cost of the factors and covariates. The next step is to select a classification rule to reduce error costs, to generate a model with the best combination of sensitivity and clarity. Then, assigning the importance of factors and independent variables is the key to the success of this heat toxin syndrome system in practical application (Supplemental materials Fig. 5).

The results above indicate that the lift chart was derived from the cumulative gain chart; the value on the y-axis corresponds to the cumulative gain ratio of the curves to the baseline. Therefore, the category "yes" increase of 10% is 30%/10% = 3.0.

The importance rank of dependent variables and covariates in the MLP network

The importance of correctly ranking the independent variables within the network model is that by manipulating different values of the independent variables, it is possible to predict the resulting change in the values of the dependent variables. The standardized importance is the importance value divided by the highest importance value and is expressed as a percentage. Sudden change of blood pressure (93.90%), thermal potential puzzled (96.00%), fainting (79.30%), gum in the eyes (83.30%), and halitosis (74%) are highly important to the system, as revealed in the experimental results above (Supplemental materials Fig. 6).

Results of establishment of the dialectical system of heat toxin syndrome of acute cerebral infarction

The heat toxin threshold was determined by the ROC curve.

The ROC curve refers to the subjects' working characteristic curve, the composite indicator that reflects the sensitivity and specificity of the continuous variables and reveals the relationship between the sensitivity and the specificity of the mapping method. By setting a plurality of different threshold values through the continuous variable, a series of sensitivities and specificities were calculated. Then, the curve was plotted with the sensitivity as the ordinate, and (1 − specificity) as the abscissa. The larger the area under the curve, the higher the diagnostic accuracy was. The point closest to the upper left of the plot is the higher threshold for both the sensitivity and specificity on the ROC curve. The ROC curves plotted have the following variables as parameters (Supplemental materials Table 9).

The heat toxin threshold was determined by the ROC curve.

$$\text{Sensitivity} = \text{true positive rate (TPF)} = a/(a + c)$$

$$\text{Specificity} = 1 - \text{the false-positive rate (the FPF)} = 1 - b/(b + d)$$

$$\text{Youden index} = \text{sensitivity} + \text{specificity} - 1$$

Draw ROC curve applied the SPSS 17.0 as follows (Supplemental materials Table 10 and Fig. 7).

The results showed that when heat toxin syndrome scored seven, the model's combination of sensitivity and specificity was the best. Based on the analysis above, the research group set up heat toxin syndrome identification system, in tabular form as follows (Table 6).

The recognition system above lacks the tongue score. Because of the difficulty of the examination of patients with acute cerebral infarction, in this study, the data of the tongue is temporarily absent (Supplemental materials Fig. 8).

The analysis of Hs-CRP showed that on Day 3, with the combination of dark face/yellow sputum/gum in the eyes/shortness of breath, the Hs-CRP was at a high level, which prompted a more severe inflammatory response, however preclude the possibility exists that infection caused the inflammatory response. The Hs-CRP level of this group of patients had a greater reduction after treatment, with the same tendency as the NIHSS result (Supplemental materials Fig. 9).

The MMP-9 level, combined with fainting/yellow sputum/halitosis, had a large variation on Day 3. In contrast, the combination of dark face/gum in the eyes/yellow sputum exhibited no obvious difference in MMP-9 levels before and after treatment at Day 7 (Supplemental materials Fig. 10).

Aspect of OX-LDL comparison. On Day 3 the fainting/yellow sputum/halitosis group exhibited a large difference in OX-LDL levels before and after treatment, however, on Day 7 the combination of dark face/gum in the eyes/yellow sputum had a small effect on OX-LDL. Interestingly, the individual diagnostic methods revealed that during early period after onset (Days 3–5), in the fainting/yellow sputum/halitosis/shortness of breath group, the inflammatory cytokines had a more substantial response to the treatment. However, when the information from the four diagnostic methods were combined above, there was no significant change in the microeconomic indicators before and after treatment appeared the late onset (Days 7–14).

Relationship between heat toxin syndrome of acute cerebral infarction and progressive stroke

Progressive Stroke is a clinical subtype of stroke; some ischemic cerebrovascular disease patients are still in a condition of deterioration after systematic treatment, called progressive stroke (PS). The European Stroke Association defined progressive stroke in 2004, according to the Scandinavian Stroke Scale scoring system. Within 3 days after the onset of stroke, the patient's level of consciousness or the motor score of the arm, leg, or eye declines by ≥2 points, and/or the language score declines by ≥3 points after a continuous nerve function evaluation, or the patient dies in 72 h. Because PS etiology and pathogenesis are complex and the symptoms change quickly, its clinical diagnosis and treatment are quite tricky. This variation of stroke outcome has similarities with heat toxin syndrome of acute cerebral infarction, based on levels of consciousness in NIHSS and sports score definitions.

The NIHSS score cardinality of progressive stroke and non-progressive stroke has no significant difference. In the

Table 6 Heat toxin syndrome of the acute cerebral infarction identification system.

Pathogenic characteristics	Macro characterization					Microeconomic indicators
Abrupt onset	◆ Vital signs	Sudden change of blood pressure		Thermal potential puzzled		Hs-CRP
Rapidly	◆ Core symptoms	Fainting	5 points	Gum in eyes	1 point	MMP-9
changing		Aphasia	2 points	Halitosis	1 point	OX-LDL
Group changes		Yellow sputum	1 point	Quick breath	2 points	OX-LDL
Dirty cloud		Dark face	4 points			
of diffuse	◆ Tongue and pulse	Quick pulse	4 points			
Very violent						
Diagnostic criteria	1 The pathogenic characteristics + core symptom score greater than 7 points + 2 item of vital signs 2 The pathogenic characteristics + core symptom score greater than 7 points + more than 1 microscopic indicator					

observation period, as the disease changes, both stroke types display a different degree of increase (Supplemental materials Table 12 and Fig. 11).

The results above reveal that within three to seven days after the stroke onset, the NIHSS score had small amplitude fluctuations, in both the progressive stroke and non-progressive stroke patients.

Relationship between the frequency of four diagnostic methods of heat toxin syndrome of acute cerebral infarction and progressive stroke

The results above show that the fainting/aphasia/gum in the eyes/halitosis group has obvious distribution differences between progressive stroke and non-progressive stroke. The four diagnostic methods combined reveal a higher frequency of progressive stroke in these patients, and also reveal dramatic changes in their disease condition, indicating that the prognosis is poor (Supplemental materials Table 13 and Fig. 12).

Relationship between heat toxin syndrome of acute cerebral infarction and infarction type and severity

The OCSP classification method: The clinical classification of ischemic stroke has great significance in the individualized treatment, prognostic assessment and secondary prevention of stroke. The Bamford typing method in the Oxfordshire community stroke project (OCSP) is a commonly used clinical stroke typing method. Through the OCSP classification of patients with cerebral infarction, we explored the relationship between OCSP clinical classifications and term prognosis of heat toxin syndrome of acute cerebral infarction. The following is a brief introduction of the OCSP classification method.

Total anterior circulation infarct (TACI): activity disorder of the higher nerves (i.e., aphasia, miscalculation and spatial disorientation), homonymous hemianopia in at least two of the three parts (face, upper and lower limbs) and movement disorder and/or sensory deficits.

Partial anterior circulation infarct (PACI): at least two of the three signs of TACI, or only the higher nervous activity dysfunction, or a sensory-motor defect range that is more limited than in TACI.

Lacunar infarction (LACI): shows lacunar syndrome, pure motor stroke, pure sensory stroke, ataxic hemiparesis or sensorimotor stroke, involving the face and arm or arm and leg.

Posterior circulation infarction (POCI): ipsilateral cranial nerve paralysis and contralateral sensorimotor disorder, bilateral sensorimotor disorder, obstruction of synergistic activity of the eyes and cerebellar dysfunction without ipsilateral long tract signs or ipsilateral visual field defects.

Dramatic changes of heat toxin syndrome of acute cerebral infarction are reflected not only in terms of macro-characterization changes but infarct location and are also closely associated with changes in the syndrome. Heat toxin syndrome of acute cerebral infarction is more common in PACI but rarely observed in LACI. The analysis of the relationship between the specific infarction type and heat toxin syndrome is as follows (Supplemental materials Table 14).

The results above indicate that a dynamic evolution exists in the acute cerebral infarction syndromes. The proportions of heat toxin syndromes in different infarction types were different. Three to five days from onset, the number of heat toxin syndrome patients increased significantly; this trend was most prominent in the patients of posterior circulation infarction (Supplemental materials Table 15).

In addition to the dynamic evolution of heat toxin syndrome appearance during the observation period, the other traditional syndromes also displayed a dynamic evolution. As the time from the stroke onset increases, the proportion of each syndrome is different in each different infarction type (Supplemental materials Fig. 13).

Stasis syndrome occurs throughout the posterior circulation infarction, however, the proportion significantly increased at 5 days after the stroke onset. The proportion of heat toxin syndrome, heat syndrome and other syndromes with empirical characteristics is higher during Days 3—5, and decreased significantly after Day 5 (Supplemental materials Fig. 14).

Stasis syndrome also seen in the greatest proportion in former circulation infarctions, however because the significantly increased proportion of heat toxin syndrome and heat syndrome, during Days 3—5, the proportions of stasis syndrome and sputum syndrome decreased notably. After Day 5, the proportions of stasis syndrome and sputum syndrome increased visibly (Supplemental materials Fig. 15).

In lacunar infarctions, heat toxin syndrome accounts for a low proportion while stasis syndrome holds the highest proportion throughout the study period, however wind syndrome has a higher proportion in this type of infarction than in any other infarction types.

Acute cerebral infarction patients are divided into the three categories above according to the OCSP classification, and the proportions of wind, heat, sputum, stasis, weakness and heat toxin syndromes in each type of infarction and the dynamic trends were compared. Heat toxin syndrome has a certain proportion in each infarction type, and coexists with other traditional syndromes, with a similar trend with heat syndrome, sputum syndrome, etc. These results suggest that heat toxin syndrome has several points of overlap with heat syndrome and sputum syndrome. This crossover of heat toxin syndrome and other syndromes is not only reflected in the same time trend; the NIHSS score variation of heat toxin syndrome also has similarities and differences with other syndromes (Supplemental materials Table 16 and Fig. 16).

The results show that the baseline of the heat toxin syndrome NIHSS scores was higher, the trend was relatively stable within 14 days after the stroke onset, there was no significant rise or decline, and this was different from the variation observed in the other traditional syndromes (Supplemental materials Table 17 and Fig. 17).

The results show that the heat toxin syndrome proportion was lower in patients of Zhong Jing Luo, but still had a certain proportion during day 1—5. Heat toxin syndrome had a higher proportion in patients of Zhong Zang Fu, and the proportion continued to rise during day 1—5, and the proportion declined after five days.

Discussion

Based on the study of four diagnostic methods and a biological indicator group about hypertension and acute

cerebral infarction, there was an obvious difference in the distribution of wind, phlegm, blood stasis, qi depression and blood deficiency in hypertensive disease and acute cerebral infarction patients. Wind and qi depression made up a higher proportion of the hypertension patients, but in acute cerebral infarction patients, the elements of fire, phlegm, and blood stasis increased notably. There was single pattern or two united patterns in hypertension, while there were multiple united patterns in acute cerebral infarction. Compared with the acute cerebral infarction group, the hypertension had a less combined type, with the blood stasis and qi deficiency types more prevalent. The combination of patterns was more complicated in the acute cerebral infarction group, with the combination of phlegm stasis and fire most prevalent. Further study showed that the acute cerebral infarction patients with thermal potential, based on information from the four diagnostic methods, displayed an obvious rise in Hs-CRP, OX-LDL, and MMP-9 compared to the patients without heat potential, and presented a united change in their weight class and biological indicator groups.

The standardization of the syndrome of stroke is important for basic research and clinical treatment. With the development of pathogenesis definitions through traditional Chinese medicine, in addition to the traditional standards of "wind, fire, phlegm, blood stasis and deficiency", most stroke disease syndromes about poison evil miss the standards, therefore research into the concept of the poison damaged brain is very important.

Because hypertension and acute cerebral infarction are complicated systemic diseases, by combining information from the four diagnostic methods and defining the biological indicator linkage characteristics, we were able to integrate data mining and clinical epidemiology. On the one hand, clinical epidemiological research idea and design methods had been applied to make the clinical test observation; On the other hand, data mining method had been used for reflecting the nonlinear relationship between the syndrome and symptoms.

With the help of a variety of data mining methods, the contribution of four diagnostic methods and their comprehensive analysis have created a new breakthrough for study of a huge complicated system in traditional Chinese medicine. In addition to the traditional linear analysis methods (i.e. logistic regression, spearman correlation), the current study introduced the principal component analysis method, which cuts into the issue from the relationship of multiple numerical variables, using the method of "reducing dimensions", a multivariate statistical method reducing several variables into a few irrelevant variables. No supervised and supervised data mining methods had been combined to form the syndrome diagnostic model of acute cerebral infarction with a heat toxin pattern, which met the requirements of the syndrome concept fuzziness of qualitative research and objective quantitative research in syndrome diagnostic standard research. Different weights were assigned to the four diagnostic methods in the syndrome diagnosis, which has important significance to the study of the dynamic evolution of the syndrome process of acute cerebral infarction with heat toxin pattern. There was a good judgment accuracy (83.82%) in the inspection stage, and the four diagnostic methods combined with excellent sensitivity and specificity, as identified by ROC curve. Specifically, when the data from the four different diagnostic methods was aggregated to score more than 7 points, it can form the heat toxin pattern.

With the heat and phlegm syndrome characteristics, such as fainting, increased eye secretion, yellow sputum, shortness of breath, and the characteristics of blood stasis, such as tarnished face, the heat toxin pattern is not independent, usually crossing, overlapping and included with other patterns. These results show that there is a transformation relationship among the different patterns, while the former diagnosis only emphasized the beginning statement of the syndrome characteristics and rules of specific four diagnostic information, without attention to the dynamic evolution of the syndrome.

The former results involving blood pressure[7,8] and temperature[9] and the related four diagnostic methods for the heat toxin pattern exhibited important significance in acute cerebral infarction and its prognostics. The results show that the principal components rapidly changed after admission from the 3rd to the 5th day, meaning that the blood pressure and body temperature change dramatically; therefore, it could be concluded that the 3rd-5th day is the turning point for acute cerebral infarction with heat toxin pattern. According to the NIHSS score, the heat toxin pattern score, the disease time, the score changes before and after treatment, and the change in inflammation factors (Hs-CRP, MMP-9/OX-LDL), it can be concluded that the condition has a drastic change, and this turning point is critical for two completely opposite changes in the disease progression, deterioration and restoration.

Conclusions

This study had simulated the cerebral ischemia damage process in three levels, hypertension low risk group, hypertension high risk group and an acute cerebral infarction group. This is a qi-deficiency, viscera function disorder and qi inverse randomly process, and the pericardium, diet, mood and fatigue are the causes. Wet phlegm, blood stasis, wind, fire and thermal depression are pathological products; grasping this essence of pathogenesis, on the basis of clinical reality, with the aid of the data mining method, the author studied the formation and development of toxin damage to the brain, summarizing the pattern of acute cerebral infarction with heat toxin and defined the production, transformation and the turning point.

Acute cerebral infarction with heat toxin pattern has four basic stages: aggregation, attack, reconciliation and change. Aggregation: toxin is from the accumulation of other pathogenic factors. Attack: attack of toxin marks the break of old balance. Reconciliation: the body has recovered the balance of Yin and Yang according to its own adjustment. Change: in this stage phlegm, blood stasis and deficiency will occur with the readjustment of body.

In the view of the whole evolution process of acute cerebral infarction with heat toxin pattern, after the attack of toxin, it changes rapidly, which reflects pathogenic rigors violence, and at the 3rd-5th day (turning point), the toxin break the defense, causing a deficiency of body, or changes

to other pathogenic factors. In summary, the process of toxin conversion is rapid and intense.

In conclusion, the results of this study indicate the following:

(1) Qi/blood/sputum/stasis existed in the whole process of hypertension and CI development and their dynamic evaluation;
(2) Heat toxin syndrome, as well as qi/blood/sputum/stasis, co-existed in the CI patients, and transformation frequently occurred during the process;
(3) Days 3—5 after the onset of CI, which was the turning point, the combination of information from several medical diagnostic methods with this model could predict the disease progression after CI.

Acknowledgments

This work was supported by the National Basic Research Program of China (973 Program) under Grants No. 2012CB518406 and 2006CB504805 and the National Science Foundation of China (Grant No. 81173463).

Appendix A. Supplementary data

Supplementary data related to this article can be found online at http://dx.doi.org/10.1016/j.jtcms.2014.11.008.

References

1. Wang Yongyan, Zhang Zhibin. Reconsideration several problems of perfect syndrome method system. *Tianjin Tradit Chin Med*. 2007;24:1—4.
2. Liu Qiang, Chen Jianxin, Chen Jing. Through the dynamic correlation coefficient cluster clinical syndrome elements extraction method trial. *Chin Traditional Chin Med Information J*. 2007;14:100—102.
3. Chen Huajun. Towards based architecture for traditional Chinese medicine. In: *Proceedings of the International Workshop*. 2003:77—89.
4. Geng Deping, Gao Ying, Tang Tao. Digital TCM diagnosis system based on data mining. *China JMIT*. 2003;S1:15—20.
5. Encephalopathy emergency group in state administration of traditional chinese medicine. Stroke disease diagnosis and curative effect evaluation standard. *J Beijing Univ Traditional Chin Med*. 1996;19:55—56.
6. Tsivgoulis G, Spengos K, Zakopoulos N, et al. Twenty four hour pulse pressure predicts long term recurrence in acute stroke patients. *J Neurol Neurosurg Psychiatry*. 2005;76:1360—1365.
7. Jain S, Namboodri KKN, Kumari S, et al. Loss of circadian rhythm of blood pressure following acute stroke. *BMC Neurol*. 2004;4:1.
8. The fourth national conference of cerebrovascular disease. All kinds of cerebrovascular disease diagnosis points. *Chin J Pract Intern Med*. 1997;17:312.
9. Wong AA, Read SJ. Early changes in physiological variables after stroke. *Ann Indian Acad Neurol*. 2008;11:207—209.