Full Length Research Paper

Unsupervised data mining technology based on research of stroke medication rules and discovery of prescription

Xinjie Li[1,2], Mingqi Wang[3], Xujie Yu[3], Xiaoyan Li[3], Jingwei Liu[3], Zhenhua Jia[1,4*] and WEI Cong[5]

[1]Hebei Yiling Medicine Research Institute, Shijiazhuang 050035, China.
[2]Key Laboratory of collateral disease of Hebei Province, Shijiazhuang 050035, China.
[3]Beijing University of Chinese Medicine, Beijing 100029, China.
[4]Hebei Yiling Hospital Cardiovascular Division, Shijiazhuang 050091, China.
[5]Key Research Centre of State Administration of Traditional Chinese Medicine (Collateral Disease of Cardiovascular/TCM Collateral Disease Theory Key discipline of State Administration of TCM of China), Shijiazhuang 050035, China.

For collecting and sorting literature on treatment of ancient and modern stroke at the entry point, frequency and charts analysis methods were applied. Modern computer technology and statistics were used to do in-depth study of stroke literature data, explore stroke differentiation of symptoms and signs and regular treatment by ancient physicians. On this basis, the unsupervised data mining technology was applied to obtain each syndrome commonly used for drug pair and combination of core. In assembling empirically drug pair and combination of core, 3 new prescriptions were obtained for each syndrome.

Key words: Unsupervised data mining, stroke, medication rules, new prescription.

INTRODUCTION

Traditional Chinese medicine has a long history of stroke awareness, and has experienced thousands of years of transmutation for its etiology, pathogenesis and treatment medication. From the "outside wind theory" to the "internal wind theory"; from dispelling wind and releasing the exterior to nourish the blood and sinew, resolve phlegm, dispel wind, tonify qi, activate blood, calm the liver and extinguish wind; from houshiheisan, daxiaoxumingtang to dihuangyinzi, sanhuatang, buyanghuanwutang and zhenganxifengtang, almost every period, representative physicians had unique insights of stroke disease and accumulated very rich experience in its diagnosis and treatment.

The evolution of the ancient physicians treating stroke reflected the innovation and advancement of the theory and methods of stroke treatment in a particular historical period. The systematic analysis of the compatibility law of prescription is an important way to accurately grasp the pathogenesis of stroke evolution. Close attention has been paid to how to use modern scientific research methods, and how to carry out high efficient and relevant research on stroke.

Research purposes

In this study, we gathered literatures on treatment of ancient and modern stroke, and with the methods of modern computer technology combined with statistics; we established the disease prescription database. We used frequency, charts and other mathematical analysis methods to summarize and sum up the ancient physicians' knowledge of stroke and the general awareness of the law of the treatment and characteristics

---

*Corresponding author. E-mail: yijianxilai534@hotmail.com.

of the medicine on the basis of the analysis of stroke prescription drug law. We then applied the unsupervised data mining technology to excavate new prescriptions of stroke common syndromes and also used the experiment to verify the validity, and opened new avenues for the discovery of new stroke compound medicines. Through the study, we establish a compound Chinese medicine research and development mode about combination of data mining and the Chinese Medical Literature Analysis.

RESEARCH CONTENTS AND METHODS

The study on ancient and modern literature of stroke medication rules

Establishing the stroke literature database

We collected literature widely on treatment of ancient and modern stroke, sorted out in line with the literature included in the standard literature records of 855 cases. We normalized their symptoms, such as traditional Chinese medicine name, collated and entered data in Epidata 3.1 software and transformed them into the SPSS18.0 database with requirements of data mining and statistical analysis; and then proceeded to the next step of database analysis.

The study on each historical period of stroke medication rules

We used the computer statistical software, SPSS18.0 with frequency analysis method to do comparative analysis on the frequency of stroke drug use and percentage of drug genus. And thus, add up the high frequency of drug and high utilization of percentage of drug genus. From there, we then inferred the disease medication rules and dominant trend in the different historical periods.

The horizontal bar chart analysis method was used to compare lateral number of various drugs flavor and the percentage of all drugs in the same dynasty, analyze briefly stroke medication rules in the same period and compare vertical commonly used class of drugs of different dynasties to grasp stroke medication rules in history.

Unsupervised data mining technology on stroke for new discovery of prescription

From 855 prescriptions of total database of stroke literature, the followings were screened respectively: syndrome of qi deficiency and blood stasis, syndrome of obstruction of collateral due to wind-phlegm, syndrome of internal block of phlegm-heat, syndrome of yin deficiency liver, and four wind syndrome prescription. Each syndrome prescription was entered respectively into SPSS18.0 software and data were processed by the Institute of Automation, Chinese Academy of Sciences, Chen (2008) designed unsupervised pattern discovery analysis software. Analyzed Drug correlation was analyzed by improved mutual information and entropy clustering method of complex systems, respectively. The unsupervised data mining technology was applied to obtain each syndrome commonly used for drug pair and combination of core. In assembling empirically drug pair and combination of core, 3 new discovered prescriptions were obtained for each syndrome, and through research and discussion, Traditional Chinese Medicine experts identified new prescription, in line with the clinical treatment of stroke based on each syndrome compatible law.

RESULTS

The study on each historical period stroke medication rules

On the basis of the stroke literature database using frequency analysis method, the following were derived: For Tang and Song dynasties, 67 flavors with high-frequency of drugs were derived; among them, exterior-releasing medicine, tonifying and replenishing medicine, and wind-dampness-dispelling medicine were used most frequently. For Jin and Yuan dynasties, 52 flavors with high-frequency of drugs were derived; among them, tonifying and replenishing medicine, resolving phlegm drug, exterior-releasing medicine, liver-pacifying and wind-extinguishing medicine were used most frequently. For Ming and Qing dynasties, 70 flavors with high-frequency of drugs were derived; among them, tonifying and replenishing medicine, exterior-releasing medicine, resolving phlegm drug, liver-pacifying and wind-extinguishing medicine were used most frequently. Near the modern era, 47 flavors with high-frequency of drugs were derived; among them, tonifying and replenishing medicine, liver-pacifying and wind-extinguishing medicine, and resolving phlegm drug were used most frequently. In the total database, 78 flavors with high-frequency of drugs were derived; among them, tonifying and replenishing medicine, exterior-releasing medicine, resolving phlegm drug, liver-pacifying and wind-extinguishing medicine, wind-dampness-dispelling medicine, blood-activating and stasis-resolving medicine, warming the interior medicine, heat-clearing formula, qi-regulating medicine, orifice-opening medicine, etc., were used most frequently.

Horizontal bar chart analysis method was used to compare the dynasties' various types of drugs, flavor number and the percentage of all the medicines. Results were analyzed with the frequency analysis method. We also longitudinally compared the different dynasties' commonly used drug classes, and found out exterior-releasing medicines were used most frequently, followed by Tang and Song dynasties. On the contrary, in near the modern era, frequent use of wind-dampness-dispelling medicine led to its descending on time; tonifying and replenishing medicines were used frequently more than 10% of all drugs used in this period. The highest in Ming and Qing dynasties were liver-pacifying and wind-extinguishing medicines; frequency of resolving phlegm drug rose with time. Blood-activating and stasis-resolving medicines and orifice-opening medicines were used frequently with little changes in the treatment of stroke during each period.

Unsupervised data mining technology on stroke for new discovery of prescription [1-2]

From the total database of ancient and modern stroke

Table 1. 78 flavors high-frequency drugs of syndrome of qi deficiency and blood stasis.

S/N	Drug	Frequency	No	Drug	Frequency
1	Chuan Xiong	151	40	Bing Pian	24
2	Fang Feng	147	41	Ru Xiang	23
3	Dang Gui	133	42	Chi Shao	21
4	Ren Shen	115	43	Gui Zhi	20
5	Bai Shu	110	44	Yuan Zhi	20
6	Qiang Huo	80	45	Ling Yangjiao	20
7	Fu Ling	78	46	Ju Huang	19
8	Bai Shao	78	47	Cang Shu	19
9	Ma Huang	77	48	Mo Yao	19
10	Rou Gui	76	49	Wei Lingxian	18
11	Tian Ma	68	50	Bai Huashe	18
12	Du Huo	64	51	Wu Yao	18
13	Fu Zi	64	52	Du Zhong	18
14	Bai Zhi	56	53	He Shaowu	18
15	Niu Xi	54	54	Mai Dong	18
16	Huang Qi	52	55	Ge Gen	17
17	Huang Qin	50	56	Zhi Qiao	17
18	Chen Pi	47	57	Gao Ben	16
19	Xi Xin	42	58	Huang Xiang	16
20	Ban Xia	42	59	Niu Huang	16
21	Quan Xie	40	60	Xi Jiao	15
22	Jiang Can	38	61	Suan Zaoren	15
23	Tian Nanxing	37	62	Sheng Ma	14
24	Chuan Wu	36	63	Wu Qiaoshe	14
25	She Xiang	35	64	Chen Xiang	14
26	Mu Xiang	34	65	Xiang Fu	14
27	Gan Jiang	32	66	Hong Hua	14
28	Sheng Dihuang	31	67	Cao Xie	13
29	Shi Gao	30	68	Man Jingzi	12
30	Qin Jiu	30	69	Huang Lian	12
31	Fang Ji	30	70	Cao Wu	12
32	Ku Xingren	30	71	Di Long	12
33	Bai Fuzi	29	72	Shi Changpu	12
34	Jie Geng	27	73	Yi Yiren	11
35	Zhu Sha	27	74	Zhi Shi	11
36	Shu Dihuang	27	75	Da Zao	11
37	Jing Jie	26	76	Fu Shen	10
38	Bo He	26	77	Tao Ren	10
39	Sheng Jiang	25	78	Zhu Li	10

literature, the followings were screened respectively: syndrome of qi deficiency and blood stasis, syndrome of obstruction of collateral due to wind-phlegm, syndrome of internal block of phlegm-heat, syndrome of yin deficiency liver wind four syndrome prescription. By analyzing 277 cases drugs of syndrome of qi deficiency and blood stasis, obtaining 78 flavors high-frequency drugs (Table 1),applied respectively improved mutual information method, entropy clustering method of complex systems, obtaining 44 drug pair (Table 2)and 61 combination of core(Table 3),assembled empirically drug pair and combination of core, obtained 3 new prescription the syndrome(Table 4) (Shenlonghuoxuefang was inside, as depicted in Figure 1);By analyzing 213 cases drugs of syndrome of obstruction of collateral due to wind-phlegm, 53 flavors with high-frequency of drugs were obtained. By applying respectively the above method, 26 drug pairs, 19 combinations of core and 3 new prescriptions for the syndrome were obtained. By analyzing 128 cases of drugs for syndrome of internal block of phlegm-heat, 43

Table 2. 44 drug pairs of syndrome of qi deficiency and blood stasis.

S/N	Drug Pair		S/N	Drug Pair	
1	Chuan Xiong	Ma Huang	23	Huang Qi	Tian Nanxing
2	Fang Feng	Jing Jie	24	Huang Qi	Bo He
3	Dang Gui	Huang Qi	25	Huang Qin	Jie Geng
4	Dang Gui	Sheng Dihuang	26	Xi Xin	Xiang Fu
5	Ren Shen	Fu Zi	27	Jiang Can	She Xiang
6	Bai Shu	Xi Xin	28	Tian Nanxing	Suan Zaoren
7	Bai Shu	Yi Yiren	29	Chuan Wu	Jie Geng
8	Qiang Huo	Bo He	30	She Xiang	Man Jingzi
9	Bai Shao	Huang Qin	31	Mu Xiang	Du Zhong
10	Ma Huang	Huang Qi	32	Gan Jiang	Chi Shao
11	Ma Huang	Quan Xie	33	Sheng Di	Mai Dong
12	Rou Gui	Du Huo	34	Shi Gao	Ge Gen
13	Rou Gui	Fu Zi	35	Jie Geng	Bo He
14	Rou Gui	Gui Zhi	36	Zhu Sha	Bai Huashe
15	Rou Gui	Ge Gen	37	Jing Jie	Bo He
16	Tian Ma	Jing Jie	38	Sheng Jiang	Ge Gen
17	Du Huo	Xi Xin	39	Sheng Jiang	Zhu Li
18	Du Huo	Shi Gao	40	Ling Yangjiao	Xi Jiao
19	Fu Zi	Cao Wu	41	Ling Yangjiao	Fu Shen
20	Bai Zhi	Wu Yao	42	Niu Huang	Xi Jiao
21	Niu Xi	Xi Xin	43	Niu Huang	Wu Qiaoshe
22	Huang Qi	Xi Xin	44	Sheng Ma	Zhu Li

flavors with high-frequency of drugs were obtained. By applying respectively the above method, 26 drug pairs, 19 combinations of core and 3 new prescriptions for the syndrome were obtained. By analyzing 64 cases of drugs for syndrome of yin deficiency liver wind, 28 flavors with high-frequency of drugs were obtained. By applying respectively the above method, 20 drug pairs,19 combinations of core and 3 new prescriptions for the syndrome were obtained.

DISCUSSION

Application of mathematical statistics methods

Using TCM to carry out research on stroke has past two thousand years ago. the ancient physicians made further exploration and played with stroke etiology and pathogenesis of the law governing its prescription. Their expanded ideas and accumulated experience were conducive to further improve the prevention and treatment of this disease. However, these researches were scattered, disordered, low-level and very repetitive. But, with the rapid development and popularization of computer and software technology, the application of certain mathematical statistical methods can lead to a more comprehensive, systematic, and in-depth study of the incidence of the law and principles of treatment of

stroke. The frequency and bar chart analysis methods are simple, intuitive, practical, and more appropriate for this study.

In this study, we comparatively analyzed the frequency of stroke drug use, the percentage of drug genus though frequency analysis; and thus obtained high frequency of drug and high utilization of drug genus. We then inferred the laws of drug and the dominant trend of the disease in different historical periods. Horizontal bar chart analysis method was used to compare lateral number of various drugs flavor and the percentage of all drugs in the same dynasty. There were more user-friendly analyzed stroke medication rules in the same period, compared to the vertical commonly used class of drugs of different dynasties, making it more conducive to grasp stroke medication rules in history.

The application based on unsupervised data mining technology

Data mining technology of the present situation in the TCM Syndrome drug law

Along with the research of the Chinese literature system, TCM Syndrome drug law research also made some progress, but still did not find appropriate ways and means. Jianxin et al. (2007) by means of artificial

Table 3. 61 combinations of core for syndrome of qi deficiency and blood stasis.

S/N	Combinations of drugs			
1	Chuan Xiong	Fang Feng	Huang Qin	
2	Fang Feng	Qiang Huo	Bai Zhi	
3	Fang Feng	Rou Gui	Fang Ji	
4	Fang Feng	Huang Qin	Fang Ji	
5	Fang Feng	Shi Gao	Fang Ji	
6	Dang Gui	Bai Shu	Ban Xia	
7	Ren Shen	Hong Hua	Tao Ren	
8	Bai Shu	Fu Ling	Ban Xia	
9	Bai Shu	Fu Ling	Shu Dihuang	
10	Bai Shu	Shu Dihuang	Suan Zaoren	
11	Qiang Huo	Tian Ma	Tian Nanxing	
12	Qiang Huo	Niu Xi	Du Zhong	
13	Qiang Huo	Shu Dihuang	Du Zhong	
14	Fu Ling	Bai Shao	Sheng Dihuang	
15	Fu Ling	Chen Pi	Sheng Dihuang	
16	Fu Ling	Sheng Dihuang	Shu Dihuang	
17	Bai Shao	She Xiang	Zhu Sha	
18	Bai Shao	Zhu Sha	She Xiang	
19	Tian Ma	Jiang Can	Tian Nanxing	
20	Du Huo	Hong Hua	Tao Ren	
21	Bai Zhi	Bai Fuzi	Cao Wu	
22	Niu Xi	Sheng Dihuang	Suan Zaoren	
23	Niu Xi	Du Zhong	Bei Jie	
24	Niu Xi	Du Zhong	Yi Yiren	
25	Niu Xi	Suan Zaoren	Hong Hua	
26	Huang Qin	Fang Ji	Xing Ren	
27	Chen Pi	Ban Xia	Tian Nanxing	
28	Chen Pi	Wu Yao	Zhi Qiao	
29	Chen Pi	Wu Yao	Huang Lian	
30	Quan Xie	Jiang Can	Huo Xiang	
31	Chuan Wu	Jing Jie	Cao Wu	
32	She Xiang	Zhu Sha	Bing Pian	
33	Mu Xiang	Zhu Sha	Chen Xiang	
34	Mu Xiang	Cang Shu	Wei Ling Xian	
35	Mu Xiang	Wei Lingxiang	Man Jingzi	
36	Gan Jiang	Sheng Jiang	Da Zao	
37	Sheng Dihuang	Shu Dihuang	Suan Zaoren	
38	Jie Geng	Wu Yao	Zhi Qiao	
39	Jie Geng	Wu Yao	Huang Lian	
40	Zhu Sha	Bing Pian	Huo Xiang	
41	Zhu Sha	Bing Pian	Chen Xiang	
42	Shu Dihuang	Suan Zaoren	Hong Hua	
43	Jing Jie	He Shouwu	Cao Wu	
44	Bing Pian	Ru Xiang	Wei Lingxian	
45	Bing Pian	Huo Xiang	Xiang Fu	
46	Ru Xiang	Mo Yao	Di Long	
47	Ru Xiang	Xiang Fu	Di Long	
48	Chi Shao	Di Long	Tao Ren	
49	Yuan Zhi	Mai Dong	Shi Changpu	
50	Ju Hua	Gao Ben	Man Jingzi	
51	Dang Gui	Ban Xia	Huang Lian	Zhi Shi

Table 3. Contd.

52	Ma Huang	Rou Gui	Fang Ji	Xing Ren		
53	Bai Zhi	Cang Shu	Wei Lingxian	He Shouwu		
54	Bai Zhi	Cang Shu	He Shouwu	Wu Cao		
55	Quan Xie	Bai Fuzi	Bai Huashe	Gao Ben		
56	She Xiang	Bing Pian	Ru Xiang	Xiang Fu		
57	She Xiang	Bing Pian	Niu Huang	Xiang Fu		
58	Bing Pian	Ru Xiang	Chen Xiang	Xiang Fu		
59	Ren Shen	Ma Huang	Shi Gao	Fang Ji	Xing Ren	
60	Fu Ling	Chen Pi	Ban Xia	Huang Lian	Zhi Shi	
61	Tian Ma	Quan Xie	Jiang Can	Bai Fuzi	Bai Huashe	WuQiaoshe

Table 4. 3 new prescriptions for syndrome of qi deficiency and blood stasis.

S/N	3 New Prescriptions				
1	Ren Shen	Hong Hua	Tao Ren	Chi Shao	Di Long
2	Dang Gui	Huang Qi	Ru Xiang	Mo Yao	Di Long
3	Dang Gui	Huang Qi	Chi Shao	Di Long	Tao Ren

intelligence, data mining, clustering analysis techniques, initially revealed a unique insight of the Xin'an physicians on stroke pathogenesis, diagnosis and treatment; they dug out the idea of their clinical treatment, to provide a reference for the prevention and treatment of stroke. The cluster analysis had its limitations; it required a specific attribution of symptoms in the study of Chinese medicine symptom. There was no objective criterion for evaluating whether clustering results are good or bad in order to determine the usefulness of the results (Xiaoyu, 2010).

Guo (2011) conducted a statistical analysis on stroke hemiplegia medical case through the use of association rules of data mining technology. They summarized the characteristics of the type of medicine for each syndrome and commonly used drug pair. But association rules in the application process may exist such as the higher rules of support and confidence when there was no significance of application. This required the operation to ensure that there are high quality data for rule validation, as well as, a number of feedback correction (Tong et al., 2009).

Liang et al. (2009) used the covariance matrix of factor analysis method for 343 prescriptions, from Han to Yuan Dynasty. They summed up the 22 kinds of factors representing different drugs compositions. Regarding outside treatment of stroke in this period, before the Yuan Dynasty, an understanding of the etiology of stroke was obtained. There are still many arguments to the "outside wind". The majorities of physicians still disperse wind and eliminate the pathogenic factors as the main treatment, combined with resolving phlegm, freeing the collateral vessels and activating blood. The modern studies showed that it may be related to the study of Jianxin et al. (2011, 2012b), but factor analysis was only normalized in the same category to the more closely associated with several variables; each type of variables became a (public) factor, reflecting most of the information of the original variables to the few factors (Wang, 2010).

These were more or less the drawbacks in the aforementioned commonly used data mining techniques in TCM Syndrome drug law research. Another method can take effects in quality of medicine (Jianxin, 2010).

The advantages of unsupervised data mining techniques in TCM Syndrome drug law

Compared with cluster analysis, association rules analysis, factor analysis and the more commonly used data mining methods; unsupervised data mining technique is more suitable for the present needs of TCM Syndrome drug law. Many scholars, who applied the unsupervised data mining analysis method in their studies, made some valuable experience. An example is Huang et al. (2010) who researched on proprietary Chinese medicine prescription law of combating influenza and tuberculosis prescription drug law by using complex systems entropy clustering method of unsupervised data mining method. He obtained commonly used drugs, drug combinations and provided a reference and a basis for further prevention and control of influenza A (H1N1), tuberculosis and related drug screening.

The unsupervised data mining method applies not only dealing with linear data, but also to dealing with non-linear data. Because this method did not do a rigid linear segmentation with data, it was especially suitable for processing complex relationship data of existing multivariate and multi-level on TCM. Its advantages according to Tang et al. (2009) and Jianxin et al. (2007) were: 1) It was a very good and important method which portrayed correlation of complex system, and it did not require data consistency, data normalization, and required less calculation; 2) It objectively reflected the situation of the data, gathered out of class elements
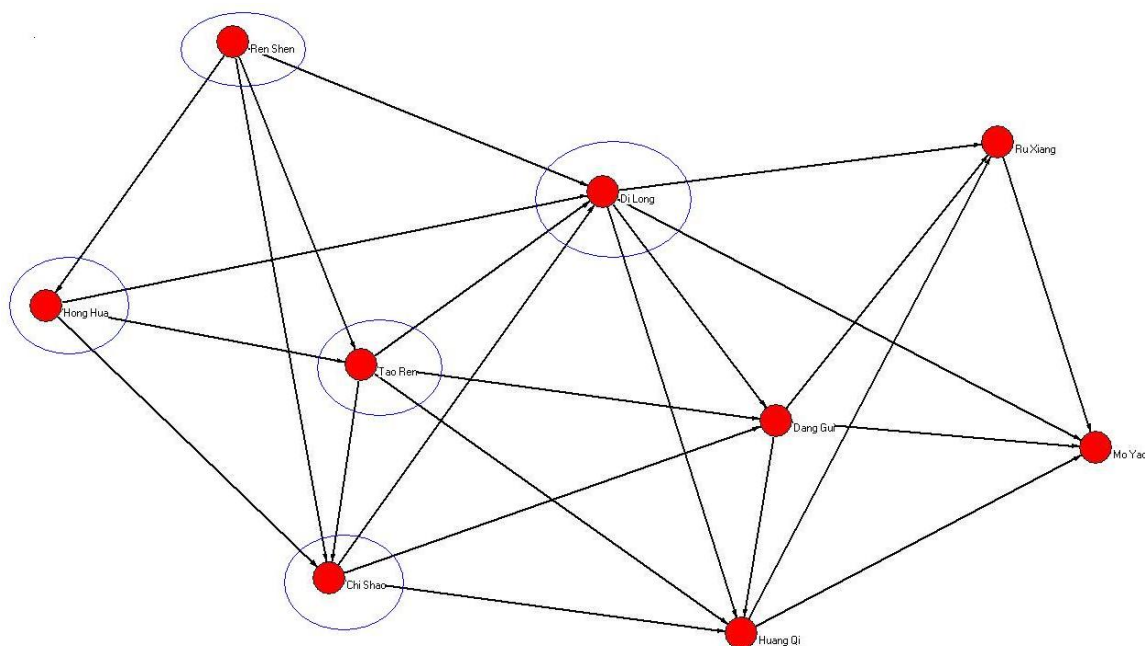
Figure 1. The three new prescriptions are discovered. Drugs of Shenlonghuoxuefang are in blue circle.

correlation, which was particularly large; 3) Convergence speed was fast for processing large amounts of data; 4) Could assign a drug to different class. This was consistent with the TCM theory, and paid special attention to the characteristics and properties of these drugs.

So, what is unsupervised data mining method? It was compared to the method of supervision. These two methods were the two strategies of machine learning methods. Supervision mining method achieved the classification and judgment of unknown samples by learning the training samples of known classes. In this case of no expert's knowledge of pre-participation, unsupervised mining method proceeded from the characteristics of the samples (variables), researched through some kind of algorithm for collecting the sample (variable) of similar characteristics together, and to achieve the purpose of distinguishing different characteristics of the samples (Qiu et al., 2004).

The improved mutual information and complexity of the system entropy clustering method with unsupervised data mining methods are more in line with the TCM theory. The latter, as an improved clustering method could not only lead to fast clustering, achieving a variable which appears in a different class, it also improved the shortcomings of traditional clustering methods "rigid" division in order to achieve a "soft" partition. The working principle proposed the concept of the variable "friends group" based on improved mutual information. It took a relative approach, selected a specific variable m, removed the value in the collection of m before the largest $Z1 \leq Z \leq (N-1)$ variables, and formed a congregation

of Z elements, denoted by D (m). Generally, Z was small compared to N, so this collection could be called "friends group" of the variable m. This is because each element of them is very related to m. Obtaining the "friends group" of each variable could lead to getting the clustering through the entropy clustering algorithm. For each variable, we all took their respective "friends group". If two variables were each in their respective friends and relatives, we think that these two variables were strongly correlated, for only strong related could be together. By analogy, the three variables could be together in a class if and only if any two variables were related. Z was finite, so this algorithm was certainly convergence. The class numbers were the algorithm automatically determined and a function of the number of variables N and the "friends group" of the number of Z (Tang et al., 2010).

This study, researched by using Epidata database platform, applied frequency statistics method and improved mutual information (Yang et al., 2005) entropy clustering method of complex systems (Chen, 2008) such as unsupervised data mining methods that primarily excavated the prescription drug law from the four levels implied in the database behind. The first level was a single herb, mainly by frequency method; the second level was the analysis of drug pair, mainly through improved mutual information method; it investigated commonly used drug pair for each syndrome and the treatment of stroke, and initially revealed the prescription drug law. One study showed that the way of taking drug may have effects (Jianxin et al., 2012a).

The third level involved rule of combination between

multiple drugs through entropy clustering method of complex systems; commonly combination of core obtained for each syndrome. The last level was that by assembling empirically drug pair and combination of core, new prescriptions were obtained for each syndrome and TCM experts by researching and discussing identified a new prescription clinical efficacy.

Conclusion

In the study, we have established stroke literature database, excavated the law of the treatment and characteristics of medicine of each historical period of stroke, and expanded its connotation, thus, providing a theoretical and clinical basis for diagnosis and treatment of stroke. Complex systems analysis methods were applied to stroke literature mining, by excavating the potential law of drugs combination, and discovering new prescription in line with the clinical treatment of stroke different syndrome.

ACKNOWLEDGEMENT

REFERENCES

Chen J (2008). Study on complex system modeling of syndrome in Traditional Chinese Medicine and its association with disease [D]. Chinese Academy of Sciences Graduate School p. 7.

Guo L (2011). Study on the Law of Stroke Hemiplegia's Syndromes [D].Hunan University of Traditional Chinese Medicine Doctoral Dissertation.

Huang L, Yang H, Tang S, Chen J (2010). Research on regulations of medicine usage in epidemic prevention based on complex system entropy method [J].The twelfth China Association for Science and Technology 22 sub-venues-"The forum Chinese medicine the status and role in the event of a major public health". Proceedings pp. 246-253.

Jianxin C, Guangcheng X, JING C (2007). An unsupervised pattern (syndrome in traditional Chinese medicine) discovery algorithm based on association delineated by revised mutual information in chronic renal failure data. J. Biol. Syst. 15(4):435–451.

Jianxin C, Huihui Z, Liangtao L, Chenglong Z, Fugang W, Qi S, Bing L, Wei W (2012a). Pharmacokinetics of oral administration of 2, 3, 5, 4'-tetrahydroxystilbene-2-O-β-d-glucoside from *Polygonum multiflorum* in beagle dogs. Afr. J. Pharm. Pharmacol. 6(14):1022–1025.

Jianxin C, Huihui Z, Xueling M, Xiao H, Liangtao L, Luya W, Jing H, Bing L, Wei W (2012b). The Effects of Jiang-Zhi-Ning and ItsMain Components on Cholesterol Metabolism. Evid-based. Compl. Alt. Volume 2012, Article ID 928234, p.10

Jianxin C, Huihui Z, Ying Y, Bing L, Jian N, Wei W (2011). Lipid-lowering and antioxidant activities of Jiang-Zhi-Ning in Traditional Chinese Medicine. J. Ethnopharmacol. 134:919–930.

Jianxin C, Xueling M, Huihui Z, Ying Y, Jing H, Shuzhen G, Bing L, Jianxin C, Zhenhua J, Xiangchun W, Guoqiang Y, Cong W,

Chenglong Z, Jianqiang Y, Yiling W (2010). Selecting biomarkers for primary hyperlipidemia and unstable angina in the context of neuro-endocrine-immune network by feature selection methods. J. Biol. Syst. 18(3):605–619.

Liang R, Huang J, Zhao W, Li J, Zhang F (2009). Research on the documents about the rules of the external treatment to Stroke before JIN and YUAN dynasty [J]. China J. Trad. Chin. Med. Pharm. 24(10):1332-1334.

Qiu D, Chen C (2004). Construction of Multi-classification Decision Tree Combining Unsupervised and Supervised Learning Strategy [J]. Mini-Micro Comput. Syst. 25(4):555-559.

Tang S, CHEN J, Wang B (2009). Designing New TCM Prescriptions Based on Complex System Entropy Cluster [J]. World Science and Technology – Modern. Trad. Chin. Med. 11(2):225-228.

Tang S, Chen J, Yang H, Huang L (2010). The New Research Field of Chinese Medicine: Syndrome-Entropy-Formula [J]. Chin. J. Basic Med. Trad. Chin. Med. 16(4):267-270.

Tong Y, Zhao Y, Yu J, Hu Y, Pan Y (2009). The application and research association rules in the TCM field [J]. Chin. J. Inform. Trad. Chin. Med. 16(7):95-96.

Wang B (2010). Study on syndrome differentiation and treatment rules of water-qi disease based on data analysis of ancient and morden medical records [D]. Heilongjiang University of Chinese Medicine Doctoral Dissertation pp. 34-35.

Xiaoyu SU (2010).The using Data Mining cluster analysis method in TCM clinical [J]. Pract. Clin. J. Integr. Trad. Chin. West. Med. 10(6):90-92.

Yang H, Zhao Y, Tang S, Huang L, Liu Y, Wang Y (2005). Study on the Relation of the Chinese herbs in the Prescription of Stroke Based on the Method of Entropy [J]. Chin. J. Basic Med. Trad. Chin. Med. 11(9):706.