

## **An Effective Traditional Chinese Medicine Syndrome Model based on Proportional Odds Model and Similarity**

**Tong-sheng CHEN**

*Institute of Artificial Intelligence, School of Information Science and Technology, Xiamen University  
Xiamen, 361005/ Fujian, China  
E-mail: tongschen@yahoo.com.cn*

**Shao-zi LI \***

*Department of Cognitive Science, Xiamen University  
Xiamen, 361005/ Fujian, China  
\*Corresponding Author: szlig@xmu.edu.cn*

**Chang-le ZHOU**

*Institute of Artificial Intelligence, School of Information Science and Technology, Xiamen University  
Xiamen, 361005/ Fujian, China  
E-mail: dozzero@xmu.edu.cn*

Received: 02-04-2009

Accepted: 08-12-2009

### **Abstract**

This paper discusses proportional odds model and similarity on the construction of Traditional Chinese Medicine (TCM) chronic gastritis syndrome model. The syndromes explain the true relationships among a set of symptoms, which can be characterized by marginal probabilities and conditional probabilities. The former represents the distribution of levels of the syndrome across the symptoms. The latter describes the probabilities of an individual in a certain syndrome being at a particular level of the symptom, so it enables us to characterize the nature of the syndromes and the essence of the symptoms. A graph model for TCM syndromes gives intuitive mathematical insight into its use and computation. Similarity among syndromes explains that the model we construct is altogether fitting and proper.

*Keywords:* Syndrome model, Proportional odds model, Similarity, Score test.

### **1. Introduction**

Syndrome differentiation of the chronic gastritis is an important part in the syndrome theory of TCM, which comes from the thinking of the human brain with analyzing rules of numerous cases, based on the learning from the ancient Chinese medical documents and accumulating individual experiences, guided by the theory of TCM. So it is sure that the syndrome theory of TCM is a priori, objective, and full of quantity and standardization.

The chronic gastritis is a common disease of the digestive system, with gastric inflammation being its notable features. This disease seriously influences the working and living of people. As social competition being fierce day by day, the rhythm of life is being accelerated constantly, the chronic gastritis patients are greatly increased. Compare to Western medicine, Chinese medicine have many advantages in the treatment of chronic gastritis.<sup>1-2</sup> Therefore, artificial intelligence has been widely used successfully in the medical field, using of artificial intelligence technology

in the field of traditional Chinese medicine research on chronic gastritis has a certain practical significance.

This paper consists of six sections: the first section is an introduction explaining the study's motivation and goals; the second section Design of TCM syndrome model; the third section explains the symptom selections; the fourth section explains syndrome similarity measures; the fifth section consists of empirical results and analysis; and the sixth section presents our conclusions and future work.

## 2. Design of TCM Syndrome Model

This section is divided into two parts which include syndrome theory of TCM, and syndrome model.

### 2.1. Syndrome theory of TCM

Zhang zhongjing (East Han dynasty) valued the function of the syndrome theory. They were rules of treatments and standards of syndrome differentiation formulated by the ancient physicians. The principle that “treatment based on syndrome differentiation” showed that both treatment and syndrome differentiation were used to diagnostic patients, though with different periods of time.

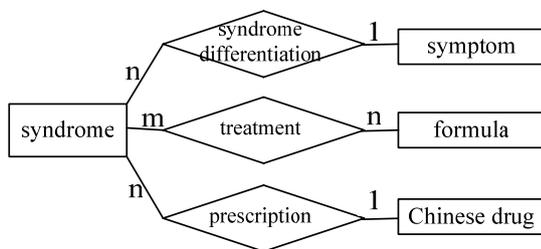


Fig. 1. An overview diagram for TCM essence.

On TCM's recommendation, the syndrome differentiation theory describes syndrome-symptom relation, the treatment theory describes syndrome-formula relation, and the prescription theory links syndrome to drug.

Now look at Figure 1, there are four entities (i.e., syndrome, symptom, formula and drug) and three relations (i.e., syndrome-symptom relation, syndrome-formula relation and syndrome-drug relation), for TCM essence. Here, syndrome can have numerous symptoms, a symptom that can be included by many syndromes. A syndrome must include at least one symptom, but the

symptom need not necessarily syndrome. Syndrome-symptom relation means a syndrome differentiates for a specific symptom (For brevity, attributes of entities and relationships are not shown in the diagram).

There are three syndrome models, each focusing on a different perspective of the relation and with its own theory. From the time of Zhang zhongjing down to today, syndrome-symptom model was the main stream of TCM thought for as long as 2000 years. Such model, referred to as syndrome model hereafter, will be considered as restrictions in syndrome-symptom relation.

Recently, many TCM knowledge were developed and achieved the acceptable performance, such as standardization of syndrome differentiation (Chen et al. 2005)<sup>3</sup>, syndrome classification (Li et al. 2006)<sup>4</sup>, syndrome-formula relation (Yu et al. 2005)<sup>5</sup>, syndrome-medicine model (Lu et al. 2008)<sup>6</sup>, and syndrome evaluation of the curative effect (Peng et al. 2009)<sup>7</sup>, goals to establish objective and quantitative diagnosis standards for syndrome differentiation. (Zhang et al. 2007)<sup>8</sup>. We survey the previous related studies as follows.

Zhang et al. (2007) used latent tree models to analyze the data set, and found natural clusters in the data set that correspond well to TCM syndrome types. It provides statistical validation to TCM syndrome types and suggests the possibility of establishing objective and quantitative diagnosis standards for syndrome differentiation. Yu et al. (2005) proposed MRAGEP in accordance with the syndrome-formula relationship of TCM. It can get an amazing function of good precision. The experiments prove the effectiveness of MRAGEP, and its accuracy is 62.5%. Lu et al. (2008) established a set of perfect Chinese medicine syndrome animal model and assessment method.

### 2.2. Syndrome model

The concept behind syndrome models is to combine several main symptoms to prevent over-fitting from the little sample. Any combination of two or more symptoms can be termed as feature selection. Over a half century, the syndrome differentiation has become a universally understood technique within the TCM field. However, this model is limited by the objective and quantitative standards. Some scholars build on the foundation of syndrome differentiation to construct a

syndrome model. Nevertheless, this approach assumes that single symptom is under the domination of the specific syndrome, and uses bayesian networks to analysis and found natural clusters in the dataset. Those assumptions are not consistent with reality (Zhang 2008) and as a result models created using this method rarely yield significantly positive performance. Chen et al. (2006)<sup>9</sup> presented entropy partition method and its ideal requirement for partition result, then apply this method to TCM dataset without understanding any expectation of the objectives. Wang et al. (2006)<sup>10</sup> pointed syndrome is composed of some symptoms, and each symptom demonstrates different values in different syndrome. Shi et al. (2007)<sup>11</sup> held the main advantage of neural network is that it may realize uncertain inference of syndrome differentiation in TCM, without request experts to provide all possible combinations for certainty degrees of symptoms and syndromes. Therefore, those aforementioned papers stress single symptom can not be under the domination of the specific syndrome.

TCM syndrome model constructing has several important decision phases such as symptom selections, syndrome similarity measures, and empirical analysis. Symptom selections phase involves distributing the different symptoms within the model and determining their weights. The classifier is often behind the idea of separating symptoms from the case collections to the point that coined to describe a certain syndrome. Syndrome similarity measures are rules of model appropriateness to regulate and adjust the relationship among syndromes. SimRank is one of the best known methods. In symptom selections phase, there are three steps, including cases collection, syndrome feature selection based on the classifier, and performance testing. In syndrome similarity measures phase, there are four steps, including compute syndrome-to-syndrome similarity, compute symptom-to-symptom similarity, compute syndrome-to-symptom weight and confirm model appropriateness. In empirical analysis phase, there are two steps, including association analysis, and analysis based on proportional odds model with deviance residuals testing.

### 3. Symptom Selections

This section is divided into three parts which include cases collection, syndrome feature selection based on the classifier, and performance testing.

#### 3.1. Case collection

In our experiment, we collect 709 cases from the digestion outpatient department of the Affiliated Shuguang Hospital of Shanghai University of Traditional Chinese Medicine during Feb., 2006 and Dec, 2006.

According to the "Diagnostic Criteria for the diagnosis of chronic gastritis combining traditional and western medicine" of the Integrated Traditional and Western Medicine Digest Special Committee, the chronic gastritis is divided into five syndromes: incoordination between liver and stomach (INCRD) (n=240); dampness-heat of spleen and stomach (DAMPH) (n=77); deficiency of spleen and stomach (DEFSS) (n=151); blood stasis in stomach (BLDSS) (n=84); and yin deficiency of stomach (YDEFS) (n=157).

In accordance with the general rules of the working of things, although chronic gastritis are divided into five syndromes, every syndrome is all correlated with 74 kinds of symptoms, i.e., the single symptom can't be under the domination of the specific syndrome. Methods of predict and classify used are related to datasets. The symptom is all dichotomous classified in the chronic gastritis case gathered. The specific symptom was coded 1 if this symptom was determined to have been present, and 0 if not. Every case belongs to single syndrome.

#### 3.2. Syndrome feature selection

This paper does not deal with data missing and situation with latent parameters. Except that the data missing while the case is gathered. In addition, this study thinks that there is no latent parameter that may influence the symptom.

The study learned from the dichotomous classification using logistic regression. After making analyzing the symptom, multi-class chronic gastritis classification and dependence of the symptom produced, the dependence of every syndrome and main symptom is arranged as Table 1 in descending form, for example "distending pain of stomach" has the strongest

dependence for INCRD, secondly it is pink tongue, stringy and slippery pulse, thin-white fur.

Table 1. Main symptoms of chronic gastritis.

syndrome	Symptoms
INCRD	distending pain of stomach, pink tongue, stringy and slippery pulse, thin-white fur, belching, gastric upset, acid regurgitation, purplish tongue, dull pain of stomach, little fur, yellow and greasy fur, deep and weak pulse
DAMPH	yellow and greasy fur, slippery pulse, burning pain of stomach, red tongue, thin-white fur, pink tongue, yellowish urine, bitter taste in mouth, heartburn, little fur, deep and weak pulse, acid regurgitation
DEFSS	pale tongue, deep and weak pulse, dull pain of stomach, teeth-print tongue, cold limbs, white and greasy fur, likeness of being warmed and pressed, weakness, red tongue, loose stool, vomiting of water, dry mouth
BLDSS	purplish tongue, stabbing pain of stomach, hemafecia, thready and unsmooth pulse, ecchymosis on tongue, dim complexion, pink tongue, distending pain of stomach, red tongue, acid regurgitation, loose stool, gastric distention
YDEFS	little fur, hunger pain, stringy and thready pulse, red tongue, constipation, fissured tongue, white and greasy fur, loose stool, pale tongue, teeth-print tongue, aggravated in the night, insomnia

In order to achieve predict optimization, this study optimize features with dynamic programming, for preventing over-fitting from too many features. After dealing with, each syndrome has different optimized number of feature, shows as Table 2, for example the optimized number of feature is 6 for INCRD, i.e., it should include 6 feature parameters in the predict: distending pain of stomach, pink tongue, stringy and slippery pulse, thin-white fur, belching and gastric upset.

Once the maximum likelihood estimates  $b_0$  and  $b_1$  are found, we substitute these values into the response function in

$$E\{Y_i\} = \pi_i = \frac{\exp(\beta_0 + \beta_1 X_i)}{1 + \exp(\beta_0 + \beta_1 X_i)} \quad (1)$$

to obtain the fitted response function. We shall use  $\hat{\pi}_i$  to denote the fitted value for the  $i$ th case:

$$\hat{\pi}_i = \frac{\exp(b_0 + b_1 X_i)}{1 + \exp(b_0 + b_1 X_i)} \quad (2)$$

The fitted logistic response function is as follows:

$$\hat{\pi} = \frac{\exp(b_0 + b_1 X)}{1 + \exp(b_0 + b_1 X)} \quad (3)$$

If we utilize the logit transformation, we can express the fitted response function in (3) as follows:

$$\hat{\pi}' = b_0 + b_1 X \quad (4)$$

i.e.  $\hat{\pi}' = b_0 + b_1 * \text{distending pain of stomach} + b_2 * \text{pink tongue} + b_3 * \text{stringy and slippery pulse} + b_4 * \text{thin-white fur} + b_5 * \text{belching} + b_6 * \text{gastric upset}$

We shall rely on standard statistical software programs specifically designed for logistic regression to obtain the maximum likelihood estimates coefficient.

Since maximum likelihood estimates  $b_0 = -3.712$ ,  $b_1 = 1.441$ ,  $b_2 = 1.704$ ,  $b_3 = 1.463$ ,  $b_4 = 0.930$ ,  $b_5 = 0.906$  and  $b_6 = 1.21$ , the fitted logistic response function is as follows:

$$\hat{\pi} = \frac{\exp(-3.712 + 1.414X_1 + 1.704X_2 + 1.463X_3 + 0.930X_4 + 0.906X_5 + 1.21X_6)}{1 + \exp(-3.712 + 1.414X_1 + 1.704X_2 + 1.463X_3 + 0.930X_4 + 0.906X_5 + 1.21X_6)} \quad (5)$$

If we utilize the logit transformation, we can express the fitted response function in (5) as follows:

$$\hat{\pi}' = -3.712 + 1.414X_1 + 1.704X_2 + 1.463X_3 + 0.930X_4 + 0.906X_5 + 1.21X_6 \quad (6)$$

The multiple logistic regression model for DAMPS is as follows:

$$\hat{\pi}' = -2.052 + 2.395X_1 + 1.240X_2 + 0.617X_3 + 0.351X_4 + 0.014X_5$$

(7), i.e.,

$$\hat{\pi}' = -2.052 + 2.395 * \text{yellow and greasy fur} + 1.240 * \text{slippery pulse} + 0.617 * \text{burning pain of stomach} + 0.351 * \text{red tongue} + 0.014 * \text{thin-white fur}.$$

The model for DEFSS is as follows:

$$\hat{\pi}' = -4.722 + 1.931X_1 + 2.223X_2 + 1.9X_3 + 1.831X_4 + 1.645X_5 + 0.870X_6 + 1.153X_7 + 1.295X_8$$

(8), i.e.,

$$\hat{\pi}' = -4.722 + 1.931 * \text{pale tongue} + 2.223 * \text{deep and weak pulse} + 1.9 * \text{dull pain of stomach} + 1.831 * \text{teeth-print tongue} + 1.645 * \text{cold limbs} + 0.870 * \text{white and greasy fur} + 1.153 * \text{likeness of being warmed and pressed} + 1.295 * \text{weakness}.$$

The model for BLDSS is as follows:

$\hat{\pi}' = -3.153 + 2.309X_1 + 2.179X_2 + 1.937X_3 + 1.166X_4$   
 (9), i.e.,  $\hat{\pi}' = -3.153 + 2.309 * \text{purplish tongue} + 2.179 * \text{stabbing pain of stomach} + 1.937 * \text{hemafecia} + 1.166 * \text{thready and unsmooth pulse}$ .

The model for YDEFS is

$\hat{\pi}' = -3.354 + 2.271X_1 + 1.817X_2 + 1.222X_3 + 0.888X_4 + 1.062X_5 + 1.119X_6 - 0.228X_7$   
 (10), i.e.,  $\hat{\pi}' = -3.354 + 2.271 * \text{little fur} + 1.817 * \text{hunger pain} + 1.222 * \text{stringy and thready pulse} + 0.888 * \text{red tongue} + 1.062 * \text{constipation} + 1.119 * \text{fissured tongue} - 0.228 * \text{white and greasy fur}$ .

Table 3 shows the varying predicting accuracy. For example the syndrome INCRD is 84.17%, DEFSS is 85.33%.

Table 2. Symptoms of different chronic gastritis selected by optimized models.

syndrome	#of feature	Symptoms
INCRD	6	distending pain of stomach(Y1), pink tongue(Y41), stringy and slippery pulse (Y53), thin-white fur(Y47), belching (Y16), gastric upset (Y17)
DAMPH	5	yellow and greasy fur(Y49), slippery pulse(Y55), burning pain of stomach(Y5), red tongue(Y42), thin-white fur(Y47)
DEFSS	8	pale tongue(Y40), deep and weak pulse(Y54), dull pain of stomach (Y3), teeth-print tongue (Y45), cold limbs(Y26), white and greasy fur (Y48), likeness of being warmed and pressed (Y8), weakness(Y28)
BLDSS	4	purplish tongue(Y43), stabbing pain of stomach (Y4), hemafecia(Y38), thready and unsmooth pulse(Y51)
YDEFS	7	little fur(Y50), hunger pain(Y2), stringy and thready pulse(Y52), red tongue(Y42), constipation (Y36), fissured tongue (Y44), white and greasy fur(Y48)

### 3.3. Performance testing

Friedman-Goldszmid<sup>12</sup>, HGC<sup>13</sup>, Cheng<sup>14-15</sup> chooses 8 symptom features of the gastritis syndrome as the feature parameter of the Bayesian classifier. Those 8 symptoms have reflected some symptoms of five chronic gastritis syndromes. See Table 3 in categorized accuracy. Three kinds of average accuracy of performing the algorithm are about 60%. In BLDSS and YDEFS, three kinds perform only 35.71% of supreme rate of accuracy of algorithms. But in INCRD and DEFSS, the categorized accuracy of HGC is above 80%, the reason is that there are more cases of these two syndromes and HGC have certain advantages on the little sample. This study showed that average accuracy 81.12% is better than before and enhances 26.94% compared to HGC.

Table 3. Predicts accuracy by different algorithms.

Syndrome	F-G	HGC	Cheng	Logistic
INCRD	67.29%	81.31%	61.68%	84.17%
DAMPH	42.22%	48.89%	44.44%	61.54%
DEFSS	69.33%	81.33%	88%	85.33%
BLDSS	21.43%	35.71%	35.71%	80.95%
YDEFS	31.91%	27.66%	25.53%	82.28%
Average accuracy	54.30%	63.91%	57.62%	81.12%

## 4. Syndrome Similarity Measures

We examine the constituents of the syndrome taking symptoms as input factors, adopting gini index and dynamic programming to select optimal numbers of symptom, and then employ logistic regression in conjunction with maximum likelihood estimates to predict accuracy. Now we shall focus center on how the syndrome interacts with its symptoms.

### 4.1. Similarity

In recent years, similarity techniques have been used in many different fields, including recommender system (Marínez et al. 2008)<sup>16</sup>, dendrogram (Upadhyaya et al. 2008)<sup>17</sup>, and medical analysis (Kong et al. 2008)<sup>18</sup>, and have demonstrated excellent performance in all of these areas. SimRank (Jeh and Widom 2002)<sup>19</sup> is based on a simple and intuitive graph-theoretic model, applicable in any domain with object-to-object relationships, that

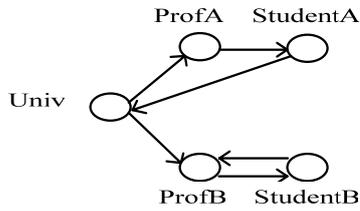


Fig. 2. A small Web graph G.

measures similarity of the structural context in which objects occur, based on their relationships with other objects.

We compute a measure that says "two objects are similar if they are related to similar objects." This general similarity measure, called SimRank, is based on a simple and intuitive graph-theoretic model. A similarity measure can be used to cluster objects, such as for collaborative filtering in a recommender system, in which "similar" users and items are grouped based on the users' preferences.

More precisely, objects  $a$  and  $b$  are similar if they are related to objects  $c$  and  $d$ , respectively, and  $c$  and  $d$  are themselves similar. The base case is that objects are similar to themselves.

As an example, consider the tiny Web graph  $G$  shown in Figure 2, representing the Web pages of two professors ProfA and ProfB, their students StudentA and StudentB, and the home page of their university Univ. Edges between nodes represent hyperlinks from one page to another. From the fact that both are referenced by Univ, we may infer that ProfA and ProfB are similar. We generalize this idea by observing that once we have concluded similarity between ProfA and ProfB, and considering that ProfA and ProfB reference StudentA and StudentB respectively, we can also conclude that StudentA and StudentB are similar. Continuing forth, we can infer some similarity between Univ and ProfB, ProfA and StudentB, etc.

Let us denote the similarity between objects  $a$  and  $b$  by  $s(a, b) \in [0, 1]$ . We write a recursive equation for  $s(a, b)$ . If  $a = b$  then  $s(a, b)$  is defined to be 1. Otherwise,

$$s(a, b) = \frac{C}{|I(a) \parallel I(b)|} \sum_{i=1}^{|I(a)|} \sum_{j=1}^{|I(b)|} s(I_i(a), I_j(b)) \quad (11)$$

where  $C$  is a constant between 0 and 1. One SimRank equation of the form (11) is written for each (ordered) pair of objects  $a$  and  $b$ , resulting in a set of  $n^2$

SimRank equations for a graph of size  $n$ . Eq. (11) says that to compute  $s(a, b)$ , we iterate over all in-neighbor pairs  $(I_i(a), I_j(b))$  of  $(a, b)$ , and sum up the similarity  $s(I_i(a), I_j(b))$  of these pairs. Then we divide by the total number of in-neighbor pairs,  $|I(a) \parallel I(b)|$ , to normalize. That is, the similarity between  $a$  and  $b$  is the average similarity between in-neighbors of  $a$  and in-neighbors of  $b$ . As discussed earlier, the similarity between an object and itself is defined to be 1.

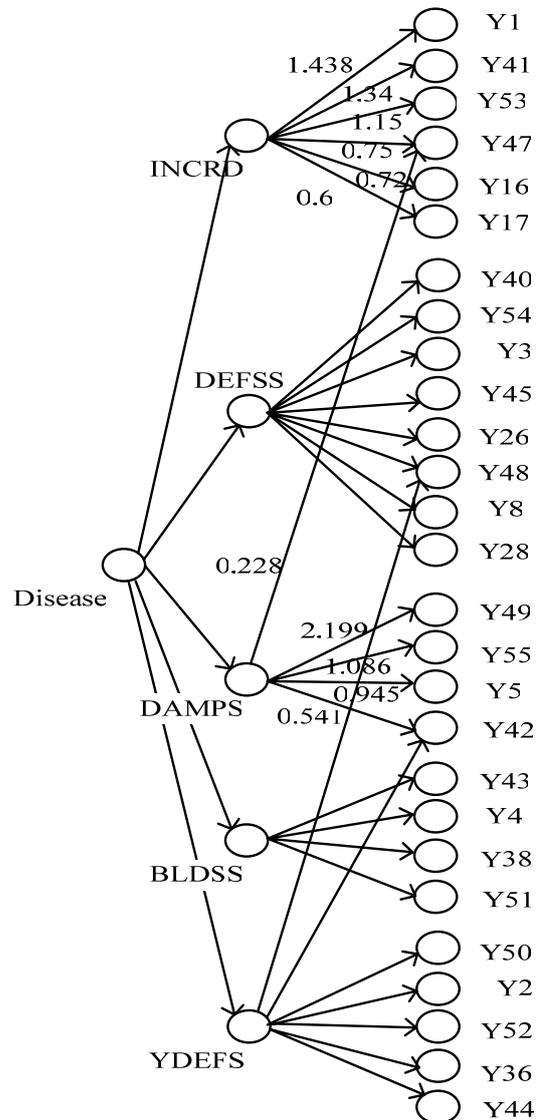


Fig. 3. Novel syndrome model graph.

**4.2. Syndrome-to-syndrome similarity**

The thought of model appropriateness should find some expression in graph (Figure 3) that conforms to the syndromes. Similarity is a general approach that exploits the syndrome-to-syndrome relationships found in TCM syndrome model.

In Table 4 (C=1.0) we concluded similarity between syndromes. For instance, INCRD and DAMPS is 0.680, INCRD and DEFSS is 0.125, etc. The range from 0 to 0.680 of similarity among syndromes explains that the model we construct is altogether fitting and proper.

Table 4. Portion of syndrome similarity scores.

Similarity scores
s(INCRD,DAMPS) = 0.680
s(INCRD,DEFSS) = 0.125
s(INCRD,BLDSS) = 0
s(INCRD,YDEFS) = 0.321
s(DAMPS,DEFSS) = 0.297
s(DAMPS,BLDSS) = 0
s(DAMPS,YDEFS) = 0.588
s(DEFSS,YDEFS) = 0.618
...

Table 5. Portion of symptom similarity scores.

s(1,49) = 0.68	s(49,48) = 0.44	s(55,2) = 0.59
s(41,49) = 0.68	s(49,8) = 0.30	s(5,40) = 0.30
s(53,49) = 0.68	s(49,43) = 0	s(5,48) = 0.44
s(47,49) = 0.84	s(49,50) = 0.59	s(5,43) = 0
s(47,55) = 0.84	s(55,5) = 1	s(5,50) = 0.59
s(16,49) = 0.68	s(55,42) = 0.79	s(42,40) = 0.46
s(17,49) = 0.68	s(55,40) = 0.30	s(42,48) = 0.62
s(49,55) = 1	s(55,48) = 0.44	s(42,43) = 0
s(49, 5) = 1	s(55,43) = 0	s(42,50) = 0.79
s(49,42) = 0.79	s(55,38) = 0	s(42,2) = 0.79
s(49,40) = 0.30	s(55,50) = 0.59	...

**4.3. Symptom-to-symptom similarity**

In Table 5 (using C=1.0) we concluded similarity between symptoms. For instance, symptom 1 and symptom 49 is 0.68 (i.e., s(1,49) = 0.68), symptom 42

and symptom 48 is 0.62, etc. Table 2 indicates the meaning of each symptom code.

**4.4. Syndrome-to-symptom weight**

Since the similarity says that we consider only the percentage of times that objects are occurred together, not the absolute number of times. This measure constraint has necessitated us not to use it for syndrome-to-symptom similarity measure. Instead, we employ gini index. The feature selection measure provides a contribution ranking for each feature. In Table 6 we concluded normalized weight between syndrome and symptom. For instance, weight between INCRD and symptom Y1 is 1.438 (i.e., w(INCRD, 1) = 1.438), INCRD and symptom Y41 is 1.34, etc.

Table 6. Portion of syndrome-to-symptom weight.

w(INCRD, 1) = 1.438	w(DAMPS,47) = 0.228
w(INCRD,41) = 1.34	w(DAMPS,49) = 2.199
w(INCRD,53) = 1.15	w(DAMPS,55) = 1.086
w(INCRD,47) = 0.75	w(DAMPS, 5) = 0.945
w(INCRD,16) = 0.72	w(DAMPS,42) = 0.541
w(INCRD,17) = 0.6	...

**4.5. Confirm model appropriateness**

Once the syndrome-to-syndrome similarity, symptom-to-symptom similarity and syndrome-to-symptom weight has been obtained, the usual next steps are to examine the appropriateness of the model and, if the fit is approved by some physicians. We shall go to empirical analysis phase.

**5. Experimental Results and Analysis**

Association measures is a graphical procedure for representing associations in a table of frequencies. We will concentrate on a two-way contingency table. If the contingency table has I rows and J columns, the plot produced by association measures contains a set of points corresponding to the rows. The positions of the points reflect associations. Row points that are close together indicate rows that have similar profiles (conditional distributions) across the columns.

If we divide the frequencies in each row (syndrome) by the corresponding row total, we obtain a profile of severities of symptom. The profiles for the different syndrome (rows) are shown in a line graph in Figure 4.

In general, the profiles are different: however, the profiles for marginal and DAMPH are similar. The lines in the figure appear to be quite different from one another. This suggests that the various levels of symptom are not distributed over the syndromes in the same way.

Let us consider an  $I \times J$  contingency table with relative frequencies  $p_{ij}$ . The measures of dependence indicates the deviation from independence of each cell (i, j) in the contingency table.

A two-way contingency table has a row variable with index  $i=1, \dots, r$  and a column variable with index  $j=1, \dots, c$ . The contents of the cells in the table are generally frequencies  $f_{ij}$ , or relative frequencies or proportions  $p_{ij}$ . The differences between the observed proportions can be modeled with a part due to the model of independence between rows and columns and a part due to the dependence between rows and columns. The model of independence between row and column variables postulates that the  $p_{ij}$  can be modeled by the product of the marginal proportions  $p_{i.}$  and  $p_{.j}$ , that is  $p_{i.} p_{.j}$ .

Therefore, we decide if  $A_i$  and  $B_j$  are independent on if there is a difference between  $p_{ij}/p_{i.}$  and  $p_{.j}$ . For each i, plot  $p_{ij}/p_{i.}$  and  $p_{.j}$ , row points that are apart from each other indicate the row categories are unrelated to the column categories.<sup>20</sup>

For clarity a total of 227 cases comply with the conditions. In our 227 collections: INCRD (n=107), DEFSS (n=75), DAMPH (n=45).

**5.1. Syndrome and belch severity profile**

Table 7 contains the frequencies of J=4 different levels of belch severity (0-3: normal, mild, moderate, severe) found at I=3 syndromes.

Table 7. Frequencies of levels of belch severity.

Syndrome	Severity				Total
	0	1	2	3	
INCRD	31	53	14	9	107
DEFSS	53	12	7	3	75
DAMPH	22	13	9	1	45
Total	106	78	30	13	227

Seldom have severity "3", so we amalgamate (i.e., add up) "2" and "3" to "2" (above moderate). If we divide the frequencies in each row (syndrome) by the

corresponding row total, we obtain Table 8 which shows 3x3 contingency table with relative frequencies.

Table 8. Syndrome x belch severity: Row proportions.

Belch syndrome	Severity		
	0	1	2
INCRD	0.290	0.495	0.215
DEFSS	0.707	0.160	0.133
DAMPH	0.489	0.289	0.222
marginal prob.	0.467	0.344	0.189

Utilizing this data to plot  $p_{ij}/p_{i.}$  and  $p_{.j}$ , we obtain a profile of severities of symptom. The profiles for the different syndrome (rows) are shown in a line graph in Figure 4.

Where dashed line denotes the marginal probability of the severity, solid line the conditional probability of each syndrome. Row points that are apart from each

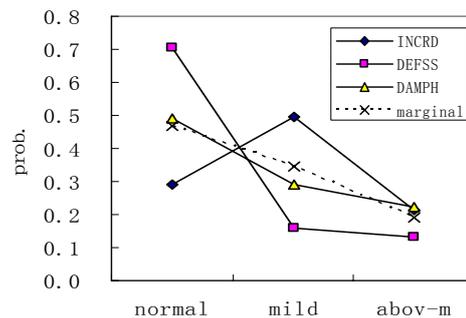


Fig. 4. Syndrome by belch severity profile.

other indicate the row categories are related to the column categories.

At level "normal" and "mild" in Figure 4, there is a maximum apart from dashed to solid line (DEFSS), nearly up to 0.2, therefore we infer there is less probability of belch severity  $\geq 1$  for DEFSS compared to DAMPH.

So belch severity associates with syndromes. We use chi-squared test for independence to test if syndrome is associated with the belch severity, the  $\chi^2$ -value would be 33.26, and  $P=1.06E-6$ .

**5.2. Syndromes and weakness severity profile**

Table 9 contains the frequencies of J=4 different levels of weakness severity (0-3: normal, mild, moderate, severe) found at I=3 syndromes. The large-sample theory for score test applies for contingency table when the fitted counts mostly exceed about 5.

Table 9. Frequencies of levels of weakness severity.

Syndrome	Severity				Total
	0	1	2	3	
INCRD	67	35	5	0	107
DEFSS	13	33	27	2	75
DAMPH	25	14	6	0	45
Total	105	82	38	2	227

Seldom have severity "3", so we amalgamate (i.e., add up) "2" and "3" to "2" (above moderate). If we divide the frequencies in each row (syndrome) by the corresponding row total, we obtain Table 10 which shows 3x3 contingency table with relative frequencies.

Table 10. Syndrome x weak severity: Row proportions.

Weakness syndrome	Severity		
	0	1	2
INCRD	0.626	0.327	0.047
DEFSS	0.173	0.440	0.387
DAMPH	0.556	0.311	0.133
marginal prob.	0.463	0.361	0.176

Utilizing this data to plot  $p_{ij}/p_i$  and  $p_j$ , we obtain a profile of severities of symptom. The profiles for the different syndrome (rows) are shown in a line graph in Figure 5, denoting marginal probability and conditional probability of weakness severity. Row points that are apart from each other indicate the row categories are related to the column categories.

At level "normal" in Figure 5, there is a maximum apart from dashed to solid line(DEFSS), nearly up to 0.2, therefore we infer there is greater probability of severity  $\geq 1$  for DEFSS compared to DAMPH.

So weakness severity associates with syndromes. We utilize chi-squared test for independence to confirm

results of Figure 5, receiving chi-squared value 52.04, with P-value=1.35E-10, indicating syndromes are all

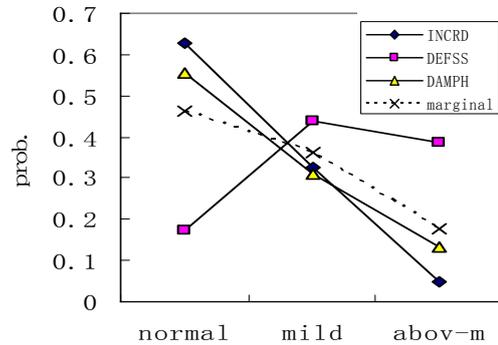


Fig. 5. Syndrome by weakness severity profile.

significantly associated with weakness severity.

### 5.3. Proportional odds model

Miller et al.<sup>21</sup> extended Generalized estimating equations (GEE), proposed by Liang and Zeger,<sup>22</sup> to correlated nominal and ordinal categorical data; in particular, they used GEE for fitting McCullagh's<sup>23</sup> proportional odds model. Stiger et al.<sup>24</sup> consider score test for assessing the assumption of proportional odds in the proportional odds model fitted with GEE. The score test requires fitting just the proportional odds model.

Ordered categorical outcomes frequently arise in the social sciences and medical research. Within the large class of generalized linear models there are a number of regression models for a multinomial response, say Y, in which there are C ordered categories with corresponding probabilities  $\pi_1, \pi_2, \dots, \pi_c$ , that is,  $Pr(Y=c)=\pi_c$ . Many of these models are based on the cumulative probabilities,  $\phi_c = \pi_1 + \dots + \pi_c$ , for  $c=1$  to  $C-1$ . The logit function is often used to relate the cumulative probabilities to a linear function of P explanatory variables X. The simplest model involving the cumulative logits is written for the ith individual as follows:

$$\text{logit } \phi_c \equiv \log(\phi_c / 1 - \phi_c) = \theta_c + \beta^T x,$$

where  $\theta_1 \leq \theta_2 \leq \dots \leq \theta_{c-1}$  are called cutpoint parameters, and  $\beta^T$  is the common vector of slope parameters across all levels of c. Since

$$\{\phi_{ic} / (1 - \phi_{ic})\} / \{\phi_{jc} / (1 - \phi_{jc})\} = \{(\beta_i - \beta_j)x\},$$

this model is referred to as the proportional odds model; that is, the ratio of the odds of being in the first c categories given  $\beta_i$ , to the odds of being in the first c

categories given  $\beta_j$ , is proportional, on an exponential scale, to the distance between  $\beta_i$  and  $\beta_j$  for all c. This model is often cited in both the statistical and medical literature and is frequently used in practice by statisticians analyzing medical data.

This section shows how to use the ordinal response levels of severity. Modeling ordinal score data that are repeated over time is common to a wide range of problems and has been studied by medical researchers. In the article the response variable were grouped into four categories: 'normal', 'mild', 'moderate' and 'severe'. We hope to find the factor influencing disease, proportional odds model is commonly used for the analysis of ordinal data, therefore we utilize it to discover the influence disease's factor. We consider the following models

$$l_j(x) = \theta_j + \beta^T x, \quad j = 1, 2, \dots, k - 1 \quad (12)$$

where  $\theta_1 \leq \theta_2 \leq \dots \leq \theta_{k-1}$ .

Given  $l(x)$  be the likelihood function of the parameters to be estimated. Suppose there is only one covariate X in the model, model denotes straight line relate to the logarithm of the accumulated odds of level j and covariate variable x, k-1 parallel regression lines are assumed for k categories of the response variable.<sup>25-27</sup>

One predictor variable was included in the study, denoting known syndromes. Syndrome status is a categorical variable with three levels. It is represented by two indicator variables ( $x_1$  and  $x_2$ ), as follows: INCRD by (1, 0), DEFSS by (0, 1), and DAMPH by (0, 0). Note the use of the indicator variables as just explained for the categorical variable. The primary purpose of the study was to assess the strength of the association between each of the syndromes. Then

$$l_j((x_1, x_2)) = \theta_j + \beta_1 x_1 + \beta_2 x_2, \quad j = 1, 2. \quad (13)$$

use score test<sup>28-29</sup> for  $b_1 = b_2 = 0$  hypothesis, the explanation of score tests is as follows: For the likelihood-ratio approach, denote the maximized value of the likelihood function by  $l_0$  under  $H_0: \beta=0$  and by  $l_1$  when  $\beta$  need not equal 0. The likelihood-ratio test statistic equals

$$D2 = -2 \log(l_0 / l_1) = -2(L_0 - L_1). \quad (14)$$

where  $L_0$  and  $L_1$  denote the maximized log-likelihood functions. Under  $H_0: \beta=0$ , the test statistic has a large-sample chi-squared distribution with  $df = 1$ .

The score test is based on model (14) at the null value for  $\beta$  of 0. If reject the hypothesis, furthermore, to

test if  $b_1=0$  or  $b_2=0$ . The name proportional odds is given because the ratio of the odds of the event  $Y \leq j$  at two values of X is independent of the choice of category j. Therefore, the odds of severity  $\leq 1$  (normal or mild) instead of severity  $> 1$  (moderate or severe) for INCRD are about  $\exp(b_1)$  the odds of severity  $\leq 1$  for DAMPH; the odds of severity  $\leq 0$  instead of severity  $> 0$  for INCRD are also  $\exp(b_1)$  the odds of severity  $\leq 0$  for DAMPH. Now according to formula,  $b_1$  is the logarithm of the estimated odds of severity  $\leq j$  instead of severity  $> j$  for INCRD and DAMPH. The odds for DEFSS are about  $\exp(b_2)$  the odds for DAMPH. The odds for INCRD are  $\exp(b_1 - b_2)$  the odds for DEFSS.

#### 5.4. Associated analysis of syndromes and belch severity

Combining severity "2" and "3" to "2"(above moderate), we utilize proportional odds model (13) to calculate,  $L_1$  equal to -231.81, if  $b_1 = b_2 = 0$ ,  $L_0 = -235.59$ ,  $-2(L_0 - L_1)$  results D2 equal 7.56, the related  $\chi^2$ -critical value will be test  $H_0: \beta_1 = \beta_2 = 0$ , we find  $\chi^2_{0.05} = 7.56$ , so  $P=0.0229$ . The null hypothesis is rejected and we conclude that  $\beta_1$  and  $\beta_2$  are not zero simultaneously. Furthermore, we analyze  $\beta_1 = 0$  or  $\beta_2 = 0$  or both not equal to zero. Table 11 is the results by maximum likelihood estimates (MLE).

Table 11. MLE – syndrome and belch severity.

Effect	B	Coefficient	Std Error	z-statistic	Sig
Intercept	$\hat{\theta}_1$	-0.167	0.292	-0.570	0.569
	$\hat{\theta}_2$	1.545	0.315	4.910	0
syndrome	$\hat{\beta}_1$	-0.521	0.337	1.545	0.122
	$\hat{\beta}_2$	0.954	0.385	-2.475	0.013

From Table 11, we accept  $\beta_1=0$  and reject  $\beta_2=0$ , P-value are 0.1223 and 0.0133, respectively. This denotes the syndromes are related to belch severity, and INCRD compare with DEFSS as indicator variables x is (1, 0) and (0, 1) respectively, so

$$L_j(1,0) = \theta_j + \beta_1 = \theta_j$$

and

$$L_j(0,1) = \theta_j + \beta_2$$

are the logarithm of odds ratio for INCRD and DEFSS respectively, then the logarithm of odds ratio for DAMPH is

$$L_j(0,0) = \theta_j$$

$\beta_1=0$  represents the logarithm of odds ratio for INCRD is the same for DAMPH.  $\hat{\beta}_2 > 0$  represents the logarithm of odds of severity  $\leq j$  for DEFSS is greater than the odds for DAMPH. In other words, DEFSS patient's belch symptoms are usually less inclined to serious. These findings are consistent with the profile pictured in Figure 4.

This odds ratio ( $\exp(0.954)$ ) tells us that the odds of severity  $\leq 0$  instead of severity  $> 0$  for DEFSS are about 2.596 times the odds of severity  $\leq 0$  for DAMPH.

5.4.1. Chi-square Goodness of fit test

The appropriateness of the fitted logistic regression model needs to be examined before it is accepted for use, as is the case for all regression models. In particular, we need to examine whether the estimated response function for the data is monotonic and sigmoidal in shape, key properties of the logistic response function. Chi-square goodness test assumes only that the  $Y_i$  observations are independent and that the sample size is reasonably large. The test can detect major departures from a logistic response function. The alternatives of interest are:

$$\begin{aligned} H_0 : E\{Y\} &= [1 + \exp(-\beta' X)]^{-1} \\ H_a : E\{Y\} &\neq [1 + \exp(-\beta' X)]^{-1} \end{aligned} \quad (15)$$

If the logistic response function is appropriate, the expected numbers of cases with  $Y_i=1$  and  $Y_i=0$  for the  $j$ th class are estimated to be:

$$E_{j1} = \sum \hat{\pi}_i \quad (16)$$

$$E_{j0} = \sum (1 - \hat{\pi}_i) = n_j - E_{j1} \quad (17)$$

where each summation is over the  $n_j$  cases in the  $j$ th class,  $E_{j1}$  denotes the estimated expected number of cases with  $Y_i=1$  in the  $j$ th class, and  $E_{j0}$  denotes the estimated expected number of cases with  $Y_i=0$ . The test statistic is the usual chi-square goodness of fit test statistic:

$$X^2 = \sum_{j=1}^c \sum_{k=0}^1 \frac{(O_{jk} - E_{jk})^2}{E_{jk}} \quad (18)$$

Since the sample size  $n=227$  is large and all expected frequencies are greater than 5, the chi-square test is appropriate here. For  $\alpha = .05$  and  $c-2=7$ , we require  $\chi^2(.95;7)=14.1$ . Since  $X^2 = 12.9 \leq 14.1$ , we conclude  $H_0$ , that the logistic response function is appropriate. The P-value of the test is 0.07.

5.4.2. Index plot of deviance residuals.

A simple graphic display of the deviance residuals is to present them in an index plot, where  $dev_i$  is plotted against its index  $i$ . Index plots help to identify outlying residuals, Figure 6 contains an index plot of the belch deviance residuals. Note that the largest absolute deviance residuals do not stand apart and we have no guidance as to whether these cases should be considered to be outliers. Obviously, index plot of other symptoms' deviance residuals are same as well.

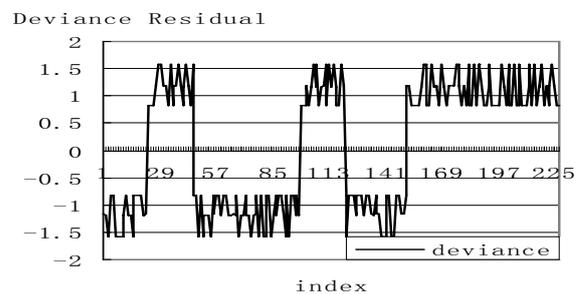


Fig. 6. Index plot of deviance residuals – belch.

5.4.3. Half-normal probability plot.

A half-normal probability plot with a simulated envelope is useful for identifying deviance residuals that are outlying and for examining the adequacy of the linear part of the logistic regression model. A half-normal probability plot helps to highlight outlying deviance residuals even though the residuals are not normally distributed. In a half-normal probability plot, the  $k$ th ordered absolute residual is plotted against:

$$z\left(\frac{k+n-1/8}{2n+1/2}\right) \quad (19)$$

Outliers will appear at the top right of a half-normal probability plot as points separated from the others. However, a half-normal plot of the absolute residuals will not necessarily give a straight line even when the fitted model is actually correct.

To identify outlying deviance residuals, we combine a half-normal probability plot with a simulated envelope. This envelope constitutes a band such that the plotted residuals are all likely to fall within the band if the fitted model is correct.<sup>30</sup>

Figure 7 contains a percentile plot of the belch deviance residuals. Note that small deviations of points from the means of the simulated values and no

occurrence of points near to or outside the simulated envelope are indications that the fitted model is appropriate.

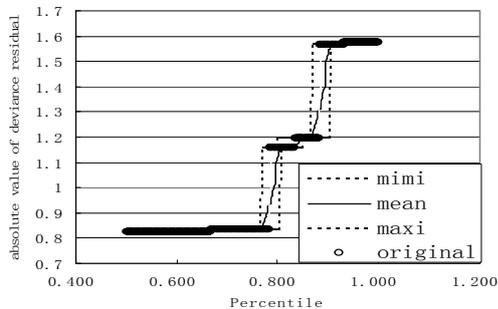


Fig. 7. Half-normal probability plot with simulated envelope – belch.

**5.5. Associated analysis of syndromes and weakness severity**

Combining severity "2" and "3" to "2"(above moderate), we utilize proportional odds model (13) to compute,  $L_1$  equal to -221.31, if  $b_1 = b_2 = 0$ ,  $L_0 = -262.8$ ,  $-2(L_0 - L_1)$  results D2 equal 83.00, the related  $\chi^2$ -critical value will be test  $H_0 : \beta_1 = \beta_2 = 0$ . Since  $\chi^2_{0.05} = 7.56$ ,  $P < 0.0001$ . The null hypothesis is rejected and we conclude that  $\beta_1$  and  $\beta_2$  are not zero simultaneously. Furthermore, we analyze  $\beta_1 = 0$  or  $\beta_2 = 0$  or both not equal to zero. Table 12 is the results by MLE.

From Table 12, we accept  $\beta_1 = 0$ , representing the logarithm of odds ratio for INCRD is the same as that for DAMPH, and reject  $\beta_2 = 0$ , since P-value is 0.257 and 0.0001 respectively. It shows the syndromes are related to weakness severity.  $\beta_2 < 0$  represents the logarithm of odds ratio of weakness  $> 1$  for DEFSS is higher than DAMPH.

Table 12. MLE – syndrome and weakness severity.

Effect	B	Coefficient	Std Error	z-statistic	Sig
Intercept	$\hat{\theta}_1$	0.148	0.297	0.498	0.6187
	$\hat{\theta}_2$	2.231	0.347	6.422	0
Syndrome	$\hat{\beta}_1$	0.403	0.355	-1.134	0.257
	$\hat{\beta}_2$	-1.751	0.385	4.545	0

This odds ratio ( $\exp(-1.751)$ ) tells us that the odds of weakness severity  $\leq 1$  instead of severity  $> 1$  for DEFSS

are about 0.174-fold the odds of severity  $\leq 1$  for DAMPH, and DEFSS patient's weakness symptoms are usually more inclined to serious. These findings are consistent with the profile pictured in Figure 5.

**5.6. Odds ratio**

The severity of stomachache, stomach bloated is similar in the behavior under three kinds of syndrome: INCRD, DEFSS and DAMPH. Severity of belch and weakness are related to syndromes.

The odds ratio of belch severity  $\leq 1$  is about 0.229-fold for INCRD compared to DEFSS patients. Thus, the 95% confidence intervals are (-2.088, -0.862) for  $\beta_1$ . The 95% confidence intervals for the odds ratio are (-2.088, 0.422). The odds ratio of belch severity  $\leq 1$  is about 0.594-fold for INCRD compared to DAMPH patients. Thus, with 95% confidence,  $0.521 \pm 1.96$  (0.337) contains  $\beta_1$  or  $-1.182 \leq \beta_1 \leq 0.140$ . The 95% confidence limits for the odds ratio are (0.307, 1.150). The odds ratio of belch severity  $\leq 1$  is about 2.596-fold for DEFSS compared to DAMPH patient, we can conclude with 95% confidence that  $\beta_1$  is between 0.199 and 1.709. The 95% confidence limits for the odds ratio are (1.221, 5.521). The odds ratio of weakness severity  $\leq 1$  is about 8.619-fold for INCRD compared to DEFSS patient, we can conclude with 95% confidence that  $\beta_1$  is between 1.526 and 2.780. The 95% confidence intervals for the odds ratio are (4.600, 16.126).

From the discussion above, we see both belch and weakness (severity) are related to chronic gastritis syndrome. The doctors may get benefit out of it for judging the syndromes.

**6. Conclusions and Future Work**

We propose a general approach to create TCM syndrome model that includes three phases such as symptom selections, syndrome similarity measures, and empirical analysis.

The syndromes represent TCM's concise and brilliant summarization which can identify if the relationships among a set of symptoms are caused by some syndromes.

A graph model for TCM syndromes gives intuitive mathematical insight into its use and computation. Similarity among syndromes explains that the model we construct is altogether fitting and proper. Experimental results using dimension reduced concept over one real dataset that show the effectiveness and feasibility of

dimension reduced concept. Thus syndrome model is created for the syndrome theory of TCM along with a priori, objective, and full of quantification and standardization.

As mentioned in Section 2.1, constructing syndrome-formula model and integrate with syndrome-symptom model is the next step when syndrome-symptom models are constructed.

### Acknowledgements

We would like to thank referees for their valuable suggestions and remarks, which improve the presentation of this paper.

### References

1. H. Z. Niu, R. X. Wang, S. M. Lan, W. L. Xu, Thinking and approaches on treatment of chronic gastritis with integration of traditional Chinese and western medicine, *Shandong Journal of Traditional Chinese Medicine*, **20**(2), (2001) 70-72.
2. B.H. Wei, The prospect of diagnosis to chronic gastritis in integrated traditional and western medicine, *Chinese Journal of Digestion*, **20**(5), (2000) 295-296.
3. Q. Chen, C. Shen, H. Zhang, et al., Application of Structural Equation Model in Standardization of Syndrome Differentiation, *Chinese Journal of Health Statistics*, **22**(1), (2005) 2-4.
4. G. C. LI, C. T. LI, L. P. Huang, Z. W. Shan and Q. G. Chen. An Investigation into Regularity of Syndrome Classification for Chronic Atrophic Gastritis Based on Structural Equation Model. *Journal of Nanjing TCM University*, **22**(4) (July.2006) 217-220.
5. X. Yu, C. J. Tang, H. Zhang. et al. Mining Formula-Syndrome Relationship in Traditional Chinese Medicine with Gene Expression Programming. *Computer Applications*, **25**(11), (2005) 2679-2680.
6. F. Lu, S. Liu, et.al., Study of evaluation method based on measures the syndrome by the medicine through fever syndrome model, *Information on Traditional Chinese Medicine* **25**(6), (2008) p61-63.
7. M. Peng, et al., A Model Based on TCM Pattern Class Evaluation of the Curative Effect, *Journal of Mathematical Medicine* **22**(4), (2009) 379-382.
8. N. L. Zhang, S. Yuan, T. Chen, and Y. Wang, Latent tree models and diagnosis in traditional Chinese medicine, *Artificial Intelligence in Medicine*, **42**(3), (March 2008) 229-245.
9. J. Chen, J. Chen, G. Xi, J. Yi, and D. Zhao, One entropy partition method for TCM, *Int. Conf. on Comput. Intelligence and Security*, ICCIAS, **1**, (2006) 823-827.
10. Y. Wang, L. Ma, X. Liao, P. Liu, Decision tree method to extract syndrome differentiation rules of posthepatitic cirrhosis in traditional Chinese medicine, *Proc. of Int. Symposium on IT in Medicine and Education, ITME2008*, (2008) 744-748.
11. M. Shi and C. Zhou, An approach to syndrome differentiation in traditional chinese medicine based on neural network, *Proc. of the 3rd Int. Conf. on Natural Computation, ICNC 2007*, **1**, (2007) 376-380.
12. N. Friedman, M. Goldszmidt. Learning Bayesian networks with local structure, Portland: in *Proc.12th Int. Conf. on Uncertainty in Artificial Intelligence*, (1996).
13. D. Heckerman, D. Geiger, D. M. Chickering, Learning Bayesian networks: The combination of knowledge and statistical data, *Machine Learning*, **20**(3), (1995) 197-243.
14. J. Cheng, R. Greiner, J. Kelly, D. A. Bell, W. Liu, Learning Bayesian networks from data: An information theory based approach, *Artificial Intelligence*. **137**, (2002) 43-90.
15. J. Cheng, Power Predictor System software, <http://www.cs.ualberta.ca/~jcheng/bnpp.htm>, (2006).
16. L. Marínez, M. J. Barranco, L. G. Pérez, M. Espinilla, A Knowledge Based Recommender System with Multigranular Linguistic Information, *Int. J. of Comput. Intelligence Systems*, **1**(3) (August, 2008) 225-236.
17. S. Upadhyaya, A. Arora, R. Jain, Reduct driven pattern extraction from clusters, *Int. J. of Comput. Intelligence Systems*, **2**(1) (March, 2009) 10-16.
18. G. Kong, D. Xu and J. Yang, Clinical decision support systems: a review on knowledge representation and inference under uncertainties, *Int. J. of Comput. Intelligence Systems*, **1**(2) (May, 2008) 159-167.
19. G. Jeh and J. Widom. SimRank: A measure of structural-context similarity, *Technical report, Stanford University Database Group*, <http://dbpubs.stanford.edu/pub/2001-41>. (2001).
20. J. D. Jobson, *Applied Multivariate Data Analysis Volume II: Categorical and Multivariate Methods*, (Springer-Verlag New York, 1992).
21. M. E. Miller, C. S. Davis, and J. R. Landis, The analysis of longitudinal polytomous data: generalized estimating equations and connections with weighted least squares, *Biometrics*, **49**, (1993) 1033-1044.
22. K. Y. Liang, and S. L. Zeger, Longitudinal data analysis using generalized linear models, *Biometrika*, **73**, (1986) 13-22.
23. P. McCullagh, Regression models for ordinal data, *J. of the Royal Statistical Society, Series B*, **42**, (1980) 109-142.
24. T. R. Stiger, H. X. Barnhart. and J. M. Williamson, Testing proportionality in the proportional odds model fitted with gee, *Statistics In Medicine*, **18**, (1999) 1419-1433.
25. P. J. Diggle, P. Heagerty, K. Y. Liang and S. L. Zeger, *Analysis of Longitudinal Data*, (Oxford Press. 2002).
26. P. McCullagh and J. A. Nelder, *Generalized Linear Models*, 2nd ed. (Chapman and Hall, New York, 1989).
27. N. Cressie and T. R. C. Read, Pearson's X<sup>2</sup> and the loglikelihood ratio statistic G<sub>2</sub>: A comparative review, *Int. statistical review*, **57**( 1), (1989) 19-43.

28. R. M. Heiberger and B. Holland. *Statistical analysis and data display: an intermediate course with examples in S-plus, R, and SAS*, (Springer-Verlag, New York, 2004).
29. W. N. Venables, and B. D. Ripley, *Modern applied statistics with S-PLUS*, (Springer-Verlag, New York, 1999).
30. A. C. Atkinson, *Two Graphical Displays for Outlying and Influential Observations in Regression*, *Biometrika*, **68**, (1981) 13-20.