

Research Article

Research on Zheng Classification Fusing Pulse Parameters in Coronary Heart Disease

Rui Guo,¹ Yi-Qin Wang,¹ Jin Xu,¹ Hai-Xia Yan,¹ Jian-Jun Yan,² Fu-Feng Li,¹ Zhao-Xia Xu,¹ and Wen-Jie Xu¹

¹ Laboratory of Information Access and Synthesis of TCM Four Diagnostic, Center for TCM Information Science and Technology, Shanghai University of Traditional Chinese Medicine, Shanghai 201203, China

² Center for Mechatronics Engineering, East China University of Science and Technology, Shanghai 200237, China

Correspondence should be addressed to Yi-Qin Wang; wangyiqin2380@sina.com and Jian-Jun Yan; jjyan@ecust.edu.cn

Received 11 January 2013; Revised 28 March 2013; Accepted 6 April 2013

Academic Editor: Aiping Lu

Copyright © 2013 Rui Guo et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This study was conducted to illustrate that nonlinear dynamic variables of Traditional Chinese Medicine (TCM) pulse can improve the performances of TCM Zheng classification models. Pulse recordings of 334 coronary heart disease (CHD) patients and 117 normal subjects were collected in this study. Recurrence quantification analysis (RQA) was employed to acquire nonlinear dynamic variables of pulse. TCM Zheng models in CHD were constructed, and predictions using a novel multilabel learning algorithm based on different datasets were carried out. Datasets were designed as follows: *dataset1*, TCM inquiry information including inspection information; *dataset2*, time-domain variables of pulse and *dataset1*; *dataset3*, RQA variables of pulse and *dataset1*; and *dataset4*, major principal components of RQA variables and *dataset1*. The performances of the different models for Zheng differentiation were compared. The model for Zheng differentiation based on RQA variables integrated with inquiry information had the best performance, whereas that based only on inquiry had the worst performance. Meanwhile, the model based on time-domain variables of pulse integrated with inquiry fell between the above two. This result showed that RQA variables of pulse can be used to construct models of TCM Zheng and improve the performance of Zheng differentiation models.

1. Introduction

Traditional Chinese Medicine (TCM) has made great contribution worldwide. In long term medical practice, Chinese physicians have realized the close relationship between the external expressive form of the human body and the essence of diseases; they have diagnosed diseases through “Zheng differentiation.” The Zheng differentiation of TCM considers the etiology, location, nature, and condition of a disease, as well as the struggle between body resistance and pathogenic factors, during a specific stage of the disease process. Zheng is a unique TCM concept. It is an abstractive collection of various signs and symptoms and is a pathological summary of a specific stage during the course of a disease. The signs and symptoms can be captured by doctors through four diagnostic methods, namely, inspection, auscultation and olfaction, inquisition, and pulse taking. TCM pulse taking has been proven to be clinically valid for over 2000 years. During

pulse taking, TCM doctors place their fingers on the radial artery, from which various physiological and pathological conditions can be detected. Traditional pulse taking has important clinical value on the diagnosis and prognosis of diseases, especially angiocardiopathy. Accurate pulse taking can only be done by TCM doctors after years of experience. Therefore, objective and digital pulse diagnosis is highly desirable.

For many years, the quantification problem of TCM pulse taking has been a hot subject in multidisciplinary research. Currently, numerous researchers in China and abroad have put forward different methods for pulse analysis, such as time-domain, frequency-domain, and time-frequency joint analyses [1–4]. However, these traditional methods based on linear concepts are restricted to system stability. They are not particularly sensitive to changes in pathological state and are insufficient in characterizing the complex dynamics

of a nonlinear system. Nonlinear dynamics methods for time series prediction can obtain nonlinear information on pulse, which cannot be achieved by traditional analytical methods. These nonlinear dynamics methods have wide applications in physiological signals, such as heart rate variability, electroencephalogram, electrocardiogram, and electromyography [5–7].

Numerous methods (e.g., K-S entropy, correlation dimension, the largest Lyapunov exponent, etc.) that can disclose nonlinear characteristics of physiological signals have been established. However, these aforementioned methods require large data and are susceptible to noise. Actual measured physiological signals often fail to meet these requirements. To overcome these problems, Eckmann et al. [8] introduced a tool that enables us to investigate the m -dimensional phase space trajectory through a two-dimensional representation of its recurrences. This representation is called a recurrence plot (RP). RPs can also be applied to rather short and even nonstationary data. Webber and Zbilut [9, 10] developed recurrence quantification analysis (RQA), a tool that quantifies structures in RPs. Marwan et al. [11, 12] supplemented RQA variables to quantify RPs. In the present study, RQA was employed to acquire the nonlinear dynamic features of the wrist pulse of coronary heart disease (CHD) patients and normal subjects. For Zheng pattern classification, RQA variables of pulse in CHD patients were analyzed.

Zheng is the pathological summary of a disease during a certain stage. The symptoms constituting both Zheng and disease are the most essential elements for disease identification and Zheng differentiation. The different stages of a disease may have various Zheng, and a similar Zheng may occur in other diseases. Deficiency and excess are the two main principles in differentiating the preponderance or decline of pathogenic factors and healthy-qi. Deficiency Zheng and excess Zheng can reflect the two aspects of struggle between body resistance and pathogenic factors during a disease process. The former refers to deficiency of body resistance, whereas the latter refers to excessive and hyperactivity of pathogenic factors. In clinical practice, the Zheng of mixed deficiency and excess, which refers to a complex condition of concurrent deficiency and excess, is common. For example, the pattern of heart-qi deficiency (deficiency Zheng) in CHD patients is accompanied by the obstruction pattern of heart vessels by phlegm (excessive Zheng). Thus, one case is often associated with multiple Zheng patterns. Therefore, TCM Zheng patterns can be classified as a multi-label problem that traditional single-label learning algorithms cannot solve. The multi-label learning (MLL) algorithm REAL is suitable for solving multi-label recognition of TCM Zheng patterns [13]. This algorithm was applied to construct Zheng classification models based on different datasets. The performances of models were compared.

The remainder of the paper was organized as follows. In Section 2, we described the collection of clinical data, including inquiry, inspection, and pulse. In Section 3, we introduced the pulse data processing method through RQA, the recognition model of Zheng pattern-based REAL, and the evaluation measures for the recognition performance of

the Zheng model. In Section 4, we summarized the results. A statistical analysis of RQA variables was made between CHD patients and normal subjects to establish the physiological and pathological significance of RQA variables. Zheng classification models of CHD were constructed based on REAL using different datasets. The datasets were designed as follows: *dataset1*, inquiry and inspection information; *dataset2*, time-domain variables integrated with *dataset1*; *dataset3*, RQA variables of pulse integrated with *dataset1*; and *dataset4*, major principal components of RQA variables and *dataset1*. Moreover, the recognition performances of different Zheng models were compared according to evaluation measures, such as *average precision*, *overage*, *hamming loss*, and *one error*, and *rank loss*. In Section 5, we discussed the effect of the RQA variables of pulse on the recognition performance of TCM Zheng in CHD patients.

2. Material

2.1. Collected Material. Patients who met the diagnostic criteria of CHD and those who provided informed consent were included in the present study. The diagnostic criteria of patients were based on western medicine and TCM. The diagnostic criteria based on western medicine was adopted from “naming and diagnosis criteria of ischemic heart disease” issued by the International Society of Cardiology and the Joint Subject Team on standardization of clinical naming in the World Health Organization [14]. The diagnostic criteria based on TCM were according to the “differentiation standards for symptoms and signs of coronary heart disease and angina pectoris in Traditional Chinese Medicine” in the “standards for differentiation of chest pain, chest distress, palpitation, short breath, or debilitation for coronary heart disease in Traditional Chinese Medicine” modified by the China Society of Integrated Traditional Chinese and Western Medicine in 1990, the “Guideline for clinical study of new drugs in Chinese herbs,” and the standards in textbooks [15]. Patients with mental diseases or other severe diseases, as well as those who cannot express their feelings clearly and did not provide informed consent, were excluded in the present study.

456 CHD cases including pulse recordings were collected. According to the inclusion and exclusion criteria of the patients, 334 CHD cases were included, and 121 CHD cases were excluded in this study. Pulse recordings of 117 normal subjects (control group) were also collected. The CHD patients (age, 63.00 ± 10.74) were those admitted to Longhua Hospital and Shuguang Hospital, which are affiliated to Shanghai University of Traditional Chinese Medicine. The healthy subjects (age, 52.17 ± 11.00) were players of the “2010 Zhangjiang ball game competition for the elderly” and the faculty of the Shanghai University of Traditional Chinese Medicine. The CHD patients were clinically differentiated using Zheng patterns by integrating inspection, pulse feeling, and inquiry information. Each of the 334 patients has more than one Zheng pattern and occurred mostly in the pattern of heart-qi deficiency (172 cases), heart-yang deficiency (72 cases), heart-yin deficiency (219 cases), and heart-vessel

obstruction by phlegm (170 cases). Therefore, these four Zheng patterns were selected in this study. The pattern of heart-qi deficiency refers to the weakness of heart-qi to pump blood, that is, blood circulation. Its main symptoms include palpitations, mental fatigue, and general symptoms of qi deficiency Zheng. The pattern of heart-yang deficiency refers to the failure of heart-yang to warm and circulate blood and to the internal production of deficient cold. Its main symptoms include palpitations, chest oppression, and general symptoms of deficient cold Zheng. The pattern of heart-yin deficiency refers to the failure of yin-fluid consumption to nourish the heart and heart-mind and to the internal disturbance of deficient-heat. Its main symptoms include palpitations, restlessness, insomnia, and general symptoms of deficient-heat Zheng. The pattern of heart-vessel obstruction by turbid phlegm refers to the condition of turbid-phlegm obstructing the heart vessel. It is mainly include mild or severe palpitation, chest oppression, and pain.

2.2. Acquisition of Inquiry and Inspection Information. In this study, inquiry symptoms were quantified according to the scale of inquiry diagnosis for the heart system [16]. The symptoms collected for inquiry diagnosis involve temperature, sweating, head, body, chest and abdomen, urine and stool, appetite, sleeping, mood, gynecology, tongue color, and face color, a total of 125 symptoms. The symptom was assigned with either “1” or “0,” referring to “existence” or “nonexistence,” respectively. The symptoms were removed when their frequencies of existence were not more than 10 to avoid interference of data redundancy. Thus, a total of 79 symptoms were selected for Zheng modeling in CHD.

2.3. Acquisition of Pulse Recordings. Pulse recordings were acquired using a Z-BOX type pulse measurement device and classical time-domain variables $h_1, h_3, h_4, h_5, t_1, t_4, t_5, t, w, h_3/h_1, h_4/h_1, h_5/h_1, w/t, A_s$, and A_d were calculated using the software included in the Z-BOX type pulse measurement device. Each subject was asked to relax for more than 3 min before pulse acquisition. The pulse waveforms of all subjects were captured for 60 s at a sampling rate of 720 Hz.

The time-domain method is most commonly used in pulse waveform analysis in TCM. This method has many applications in clinical practices. The classical time-domain variables are described through some characteristic points on the pulse wave (Figures 1 and 2) [17].

3. Methods

3.1. Extraction of RQA Variables of Pulse. Phase space reconstruction is the basis for the nonlinear time series analysis. It can be used to estimate the characteristic of the dynamic system. Usually, the phase space has to be reconstructed from the original one-dimensional time series [18, 19]. The time delay method is frequently used for reconstruction. For one-dimensional time series of length M , a trajectory X_i was

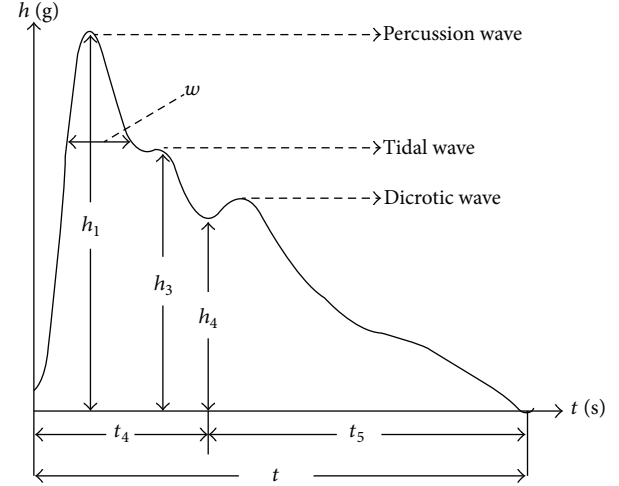


FIGURE 1: Height and time variables of a typical pulse cycle. h_1 : height of percussion wave, h_3 : height of tidal wave, h_4 : height of dicrotic notch, t_4 : time distance between the starting point of pulse chart and dicrotic notch, t_5 : time distance between dicrotic notch and the ending point of pulse chart, t : time distance between the starting point and the ending point, w : width of percussion wave in its 1/3 height position.

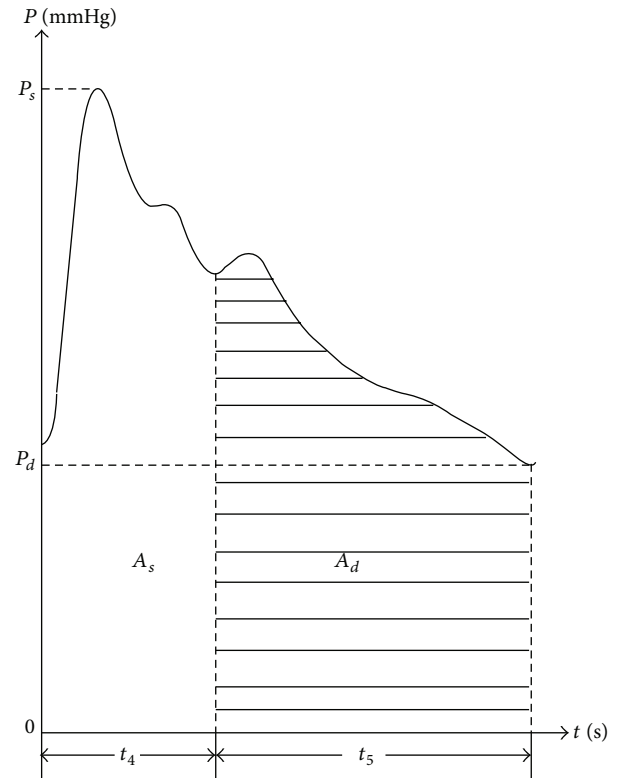


FIGURE 2: Area variables of a typical pulse cycle. P_s : systolic pressure, P_d : diastolic pressure, A_s : area of contraction phase, and A_d : area of diastolic phase.

reconstructed as $X_i = (x(i), x(i + \tau), \dots, x(i + (m - 1)\tau))^T$. The length of X_i was $N = M - (m - 1) \times \tau$, where the embedding dimension m can be estimated with the method of false nearest neighbors [20, 21], and the time delay τ can be estimated with the method of mutual information [22]. All these marked time trajectories $\{X_i, i = 1, 2, \dots, N\}$ made up the m -dimension phase space orbits of the system. Figure 3(a) showed an example of phase space reconstruction of one-dimensional pulse data.

RPs are a two-dimensional squared matrix with black and white dots, where black dots mark a recurrence and both axes were time axes Figure 3(b) [12]. RPs visualize the recurrent behavior of dynamical systems and can be mathematically expressed as [11, 12]

$$R_{i,j} = \Theta(\varepsilon_i - \|\vec{x}_i - \vec{x}_j\|), \quad i, j = 1, 2, \dots, N, \quad (1)$$

where N is the number of considered states x_i , ε_i is a threshold distance, $\|\cdot\|$ is a Euclidean norm, and $\Theta(\cdot)$ is the Heaviside function.

As shown in Figure 3, RP was composed of dots and lines in the diagonal and vertical structures. The diagonal structures mean that the evolution of states is similar at different times and that the process can be deterministic. The vertical structures describe the stability of the system. The visual interpretation of RPs only makes a qualitative analysis on the dynamic characteristics of the system. Thus, quantitative analysis of RPs was developed. Webber and Zbilut [9, 10] defined the measures of complexity using the recurrence point density and diagonal structures in the RPs. Gao [23] defined the measures of vertical structures. In the present study, we calculated the RQA measures of pulse using the following variables: recurrence rate (RR), determinism (DET), average diagonal line length (L), maximum diagonal structure length (L_{\max}), Shannon entropy of the frequency distribution of diagonal line length (ENTR), laminarity (LAM), the average length of vertical structures (TT), and the maximal length of vertical structures (V_{\max}).

Pulse morphology refers to blood pressure, vascular resistance, artery compliance, and so forth, which are important variables in assessing the cardiovascular system. Pulse morphological variation refers to the alterations and variations in the morphology of pulse waveform [24]. In this study, RQA was used to analyze pulse morphological variation. In this case, embedding dimension $m = 3$, time delay $\tau = 5$, and distance cutoff $\varepsilon = 0.3$. Each pulse recording was segmented using a moving window with a size of 1000 sampling points. The window moves forward by 300 sampling points each time. Here, we applied eight RQA measures: $\{RR, DET, L, L_{\max}, ENTR, LAM, TT, \text{ and } V_{\max}\}$. For each RQA measure during window movement, each pulse recording should produce a series of window RQA variables. The mean of the window RQA variables (MWRQA) and standard deviation of the window RQA variables (SWRQA)

for each RQA measure were calculated based on the following mathematical expressions:

$$\begin{aligned} \text{MWRQA}_j &= \frac{1}{\text{NW}} \sum_{i=1}^{\text{NW}} \text{RQA}_i^j, \\ \text{SWRQA}_j &= \sqrt{\frac{1}{\text{NW}} \sum_{i=1}^{\text{NW}} (\text{RQA}_i^j - \text{MWRQA}_j)^2}, \end{aligned} \quad (2)$$

where NW is the number of windows, RQA_i^j is the j th RQA variables of the i th windows of one pulse recording, MWRQA_j is the mean of the j th RQA variables of all windows of one pulse recording, and SWRQA_j is the standard deviation of j th RQA variables of all windows of one pulse recording.

In this study, we selected eight RQA measures and then applied eight MWRQAs and eight SWRQAs to describe the morphological variation of a pulse recording.

3.2. Construction of a Zheng Classification Model Using REAL. Multi-label learning deals with objects having multiple labels simultaneously. The TCM Zheng pattern belongs to such a problem. Formally, let $\chi = \mathcal{R}^d$ be the d -dimensional input space, and let $y = \{l_1, l_2, \dots, l_q\}$ be the finite set of q possible labels. The task of multi-label learning (or multi-label classification) is to learn a function $h : \chi \rightarrow 2^y$ that maps each instance $x \in \chi$ into a set of proper labels $h(x) \subseteq y$ [25]. REAL, a new multi-label learning algorithm with features selected through the maximization of mutual information, was proposed by our research team and was confirmed suitable for TCM Zheng differentiation [13, 26]. For this method, feature variables associated mostly with Zheng were selected according to maximization of mutual information. This method fully paid attention to the relationship between features and Zheng; thus, it is more suitable for TCM Zheng classification. In this paper, REAL was applied to construct a model of TCM Zheng based on different datasets.

To evaluate the performance of Zheng models based on different datasets, five evaluation measures [27, 28] especially designed for multi-label learning were used.

Average precision evaluates the average fraction of proper labels ranked above a particular label set. For this evaluation measure, higher values mean better classifier performance.

Coverage evaluates how far we need, on the average, to go down the list of labels to cover all the proper labels of the multi-label example. For this evaluation measure, smaller values mean better performance.

Hamming loss evaluates how many times an example label pair was misclassified. For this evaluation measure, smaller values of hamming loss mean better performance.

One error evaluates how many times the top-ranked label was not in the set of proper labels of the multi-label example. For this evaluation measure, smaller values mean better performance.

Ranking loss evaluates the average fraction of Zheng label pairs that were disordered for the multi-label example. For this evaluation measure, smaller values of ranking loss mean better performance.

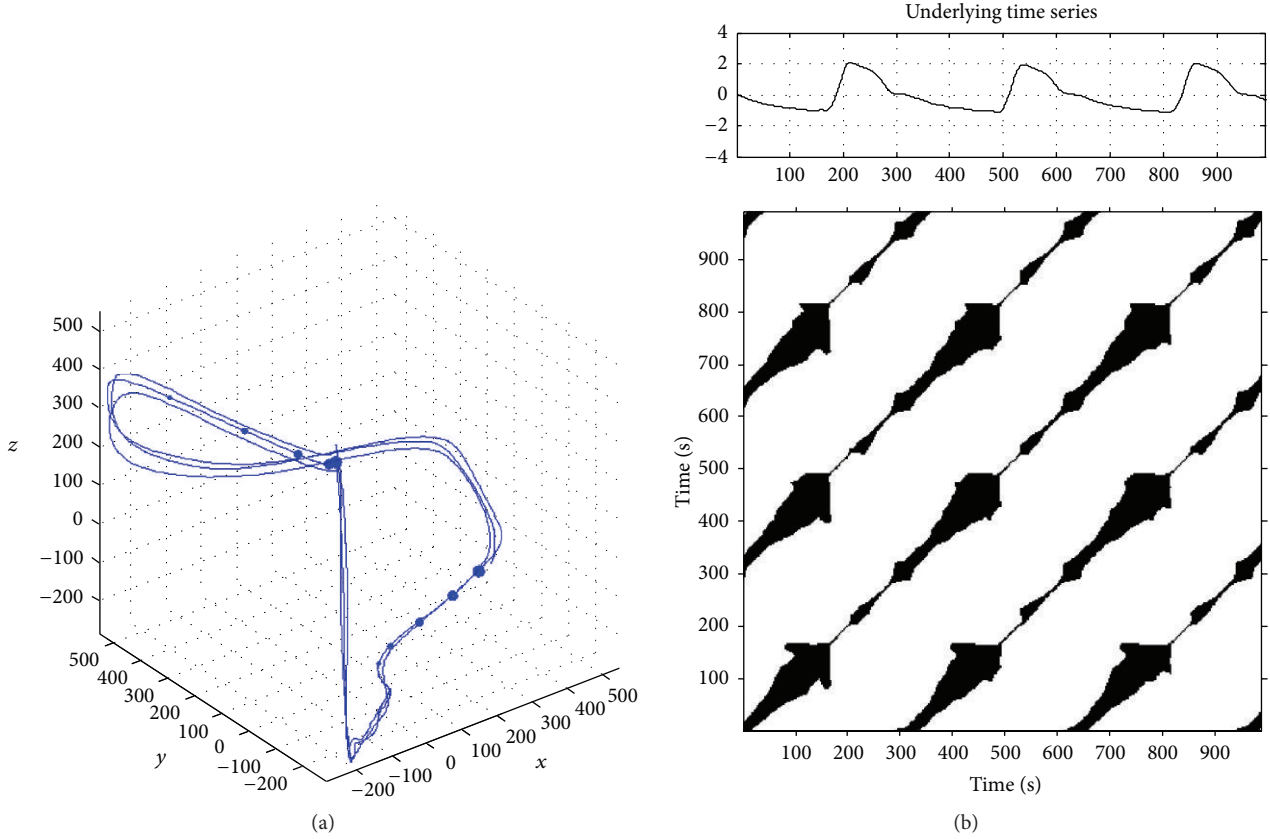


FIGURE 3: Reconstructed phase spaces and corresponding RP of the pulse recording of a CHD patient. (a) A segment from phase space trajectory of the pulse recording of a 60-year-old patient ($m = 3$, $\tau = 5$, $\varepsilon = 0.2\sigma$, σ represents the variance of time series, $y = x + \tau$, and $z = x + 2 * \tau$). (b) Corresponding RP of (a).

4. Results

4.1. Comparison of RQA Variables of Pulse between the CHD Patients and Normal Subjects. The physiological and pathological significances of RQA variables were discussed in Section 5. In this section, a statistical analysis of pulse RQA variables using independent sample t -test and rank-based ANOVA was made between the CHD patients and healthy subjects. Independent sample t -test was applied to analyze the RQA variables with normal distribution and homogeneous variance. Rank-based ANOVA was applied for analysis of RQA variables without these aforementioned requirements. Age variable as covariant was included in the statistical model to correct the effect of age. In Table 1, P values of group effects were calculated after age effects were corrected. As shown in Table 1, the RQA variables of two groups, except for the MWRQAs of L , SWRQAs of LAM, and SWRQAs of DET, had not significantly difference. The MWRQAs of $\{RR, DET, L_{max}, ENIR, LAM, TT, \text{ and } V_{max}\}$ and SWRQAs of $\{RR, L, L_{max}, ENIR, TT, \text{ and } V_{max}\}$ of the CHD patients were significantly higher than those of the healthy subjects.

4.2. Principal Components Analysis of RQA Variable. Traditionally, principal component analysis (PCA) has been the

standard approach to reduce the high-dimensional original pattern vector space into low-dimensional feature vector space. It has wide applications in feature identification. RQA variables contained much redundancy because some variables may be highly dependent on each other. In the present study, PCA extracted four major principal components of RQA variables that represent 83.24% of those that RQA variables can represent, which may be very useful for purpose of classification.

4.3. Comparison of Model Performance for Zheng Differentiation with/without RQA Variables. In this study, each CHD case was simultaneously associated with multiple Zheng, namely, Heart-qi deficiency, heart-yang deficiency, heart-ying deficiency, and heart-vessel obstruction by turbid phlegm. Thus, Zheng classification models were constructed, and predictions were made using the MLL algorithm REAL. For comparison, datasets for the construction of the models were designed as follows: (1) *dataset1*, only inquiry and inspection information; (2) *dataset2*, time-domain variables of pulse and *dataset1*; (3) *dataset3*, RQA variables of pulse and *dataset1*; and (4) *dataset4*, the four major principal components of RQA variables and *dataset1*. A tenfold cross-validation was performed on each experimental dataset. Up

TABLE 1: RQA variables show significant differences, as determined by independent sample t -test and rank-based ANOVA.

Measures	Groups		P
	Patient with CHD	Normal subjects	
MWRQAs of RR	$0.07 \pm 0.0141^{*\square}$	$0.0624 \pm 0.008^{\square}$	$P = 0.015$
MWRQAs of DEI	$0.998 \pm 0.0018^{*\square}$	$0.993 \pm 0.0102^{\square}$	$P < 0.001$
MWRQAs of L	$28.714 \pm 8.582^{\square}$	$25.201 \pm 7.148^{\square}$	$P = 0.198$
MWRQAs of L_{\max}	$557.941 \pm 257.692^{*\square}$	$400 \pm 197.999^{\square}$	$P < 0.001$
MWRQAs of ENTR	$4.036 \pm 0.275^{*\square}$	$3.893 \pm 0.285^{\square}$	$P = 0.044$
MWRQAs of LAM	$0.996 \pm 0.0021^{*\square}$	$0.995 \pm 0.0034^{\square}$	$P = 0.001$
MWRQAs of TT	$16.606 \pm 3.980^{*\square}$	$13.947 \pm 2.587^{\square}$	$P = 0.001$
MWRQAs of V_{\max}	$70.616 \pm 18.727^{*\square}$	$55.198 \pm 9.759^{\square}$	$P = 0.001$
SWRQAs of RR	$0.011 \pm 0.005^{*\square}$	$0.007 \pm 0.003^{\square}$	$P < 0.001$
SWRQAs of DET	$230 (102.5-336.5)^{\Delta}$	$210 (131-358)^{\Delta}$	$P = 0.102$
SWRQAs of L	$4.574 \pm 1.936^{*\square}$	$3.316 \pm 1.443^{\square}$	$P < 0.001$
SWRQAs of L_{\max}	$138.739 \pm 89.413^{*\square}$	$94.483 \pm 90.881^{\square}$	$P < 0.001$
SWRQAs of ENT'T	$0.165 \pm 0.087^{*\square}$	$0.133 \pm 0.002^{\square}$	$P < 0.001$
SWRQAs of LAM	$236 (106.5-347)^{\Delta}$	$219 (113.75-331)^{\Delta}$	$P = 0.362$
SWRQAs of TT	$2.208 \pm 1.165^{*\square}$	$1.386 \pm 0.602^{\square}$	$P < 0.001$
SWRQAs of V_{\max}	$15.383 \pm 9.035^{*\square}$	$7.788 \pm 4.126^{\square}$	$P < 0.001$

Compared with normal group, * difference was significant. \square Analyzed by independent sample t -test (mean \pm standard deviation); Δ analyzed by rank-based ANOVA (M (QL-QU)).

TABLE 2: Performance comparison of Zheng classification models based on different datasets.

Group	Dataset1	Dataset2	Dataset3	Dataset4 (PCA)
Average precision	0.8423	0.8534	0.8626	0.8735
Coverage	0.3932	0.3925	0.3786	0.3604
Hamming loss	0.2621	0.2684	0.2428	0.2228
One error	0.2454	0.2089	0.1996	0.1982
Rank loss	0.1944	0.1875	0.1768	0.1543

Note: dataset1, inquiry information; dataset2, time-domain variables of pulse and dataset1; dataset3, RQA variables of pulse and dataset1; dataset4, four major principal components of RQA variables and dataset1.

to 90% of the samples were randomized as the training set; the other 10% were randomized as the test set. Prediction analysis for TCM Zheng was performed after re-testing the models 10 times and taking the mean value.

The performances of the different models based on different datasets were compared in terms of different evaluation measures (Table 2). For *average precision*, higher values meant better performance. For the other four evaluation measures (*coverage*, *hamming loss*, *one error*, and *rank loss*), smaller values meant better performance. As shown in Table 1, for *dataset1*, the value of *average precision* was the lowest. The values of the other four evaluation measures were the highest. These results showed that the performance of the model based on *dataset1* without pulse information was the worst. For *dataset2* with time-domain variables of pulse, the performance of the model was superior to that in *dataset1* but was inferior to those in *dataset3* and *dataset4* including information that RQA variables of pulse can represent. The model based on *dataset4* had the best performance owing to major principal component of RQA variables that maintain

the correlated part of information discarding the minor noise-dominated part of information, which can improve the recognition ability of the model.

5. Discussion

Pulse diagnosis is one of the most important diagnostic methods in TCM. The pulse driven by the heart propagates through the arteries. Hence, the pulse wave should contain most of the information about the cardiovascular system. Many clinical experiments have confirmed that pulse condition is closely related to the cardiovascular system. The pulse wave form (shape), velocity (fast or slow), period (rhythm), and swing (intensity) are affected by the physiological and pathological features of the cardiovascular system [29]. The traditional pulse analytical methods include time-domain, frequency-domain, and time-frequency joint analyses. However, these methods based on linear concepts cannot be used to extract nonlinear features of the pulse. Therefore, in this study, the nonlinear dynamics method RQA was

applied to extract nonlinear features of the pulse. RQA represents recurrence phenomenon, a fundamental property of deterministic dynamic systems. Different RQA variables represent different nonlinear dynamic characteristics. RR is a measure of recurrence density. High RR values indicate the presence of a strong cyclical process. L and DET are measures of system determinism. High L and DET values suggest stronger system determinism. ENTR is a measure of the complexity of the deterministic structure in the system. LAM, TT, and V_{\max} mark a time interval, wherein a state does not change or changes very slowly. Therefore, higher LAM, TT, and V_{\max} values indicate higher system stability. As shown in Table 1, the MWRQAs of {RR, DET, L , L_{\max} , ENIR, LAM, TT, and V_{\max} } of the CHD patients were significantly higher than those of the healthy subjects, and the mean L of the healthy subjects was slightly higher than that of the CHD patients. This finding indicates that the cardiovascular system of the CHD patients, closely connected with pulse, had greater regularity and stability than that of the normal subjects, which is consistent with the discovery by many researches that system in pathological state is usually more stable. Moreover, the SWRQAs of {RR, L , L_{\max} , ENTR, LAM, TT, and V_{\max} } of the CHD patients were significantly higher than those of the healthy subjects. This finding suggests that the pulse morphology of the CHD patients has larger variability than that of the normal subjects. Hence, we speculate that the CHD patients show weaker ability in regulating the cardiovascular system than the healthy subjects.

CHD is a complex process expressed as different Zheng patterns in its different stages. CHD can disrupt the pumping function of the heart, causing changes in both the structure and function of blood vessels and inevitably giving rise to changes in pulse. These changes vary with different Zheng patterns. Therefore, in this study, the RQA variables of pulse were used to characterize the changes in Zheng classification. In clinical practice, one patient is often associated with several Zheng patterns. Zheng classification belongs to a multi-label problem. Therefore, Zheng models were constructed, and predictions were carried out using MLL algorithm REAL. The performances of models based on different datasets were compared (Table 2). According to the five evaluation measures, the model for Zheng differentiation based on RQA variables integrated with inquiry had the best performance, whereas that based only on inquiry information had the worst performance. Meanwhile, the model based on time-domain variables of pulse integrated with inquiry fell between the above two. Therefore, the RQA variables can improve the ability of a model for Zheng classification. Moreover, the model based on *dataset4* had the best performance owing to major principal component of RQA variables that maintain the correlated part of information discarding the minor noise-dominated part of information, which can improve the recognition ability of the model.

The linear method represented by the classical time-domain method is widely used to analyze pulse. However, the nonlinear dynamics method looks bright and promising in revealing the inside information and dynamic properties of systems. The RQA variables of pulse can be used to construct

models of TCM Zheng and improve the performance of Zheng differentiation models.

Conflict of Interests

The authors declare that they have no financial and personal relationships with other people or organizations that can inappropriately influence our work and there was no professional or other personal interest of any nature or kind in any product, service, and/or company that could be construed as influencing the position presented in, or the review of, the paper.

Acknowledgments

This work was supported by Innovation Program of Shanghai Municipal Education Commission under Grant no. 11YZ71, the National Natural Science Foundation of China under Grant no. 81173199. Thanks are due to all of volunteers for providing invaluable pulse data.

References

- [1] H. X. Yan, Y. Q. Wang, and F. F. Li, "Research and application of pulse analysis," *Chinese Archives of Traditional Chinese Medicine*, vol. 23, no. 1, pp. 129–130, 2005.
- [2] B. Jiang and Z. C. Song, "Spectral analysis of pulse signals," *Techniques of Automation and Application*, vol. 26, no. 8, pp. 38–40, 2007.
- [3] F. F. Li, Y. Q. Wang, R. Sun, S. Xue, D. Yao, and H. D. Sheng, "Application of the HHT & time domain method to the wrist-pulse-signal analysis in patients of coronary heart disease," *Chinese Journal of Cardiovascular Review*, vol. 6, no. 4, pp. 244–246, 2008.
- [4] L. Cheng and L. J. Yang, "Identifying pulse signals based on fractal theories and wavelet," *Aeronautical Computing Technique*, vol. 38, no. 4, pp. 41–44, 2008.
- [5] D. Gallez and A. Babloyantz, "Predictability of human EEG: a dynamical approach," *Biological Cybernetics*, vol. 64, no. 5, pp. 381–391, 1991.
- [6] H. W. Lu and Y. Z. Chen, "Fine graining complexity measure algorithm and its application in cardiac RR intervals analysis," *Space Medicine and Medical Engineering*, vol. 17, no. 6, pp. 444–447, 2004.
- [7] C. G. Shao and S. T. Tong, "Nonlinear dynamic fast analysis and analysis of epileptic's EEG," *Journal of Biomedical Engineering Research*, vol. 27, no. 3, pp. 154–157, 2008.
- [8] J. P. Eckmann, S. O. Kamphorst, and D. Ruelle, "Recurrence plots of dynamical systems," *Europhysics Letter*, vol. 4, no. 9, pp. 973–977, 1987.
- [9] C. L. Webber and J. P. Zbilut, "Dynamical assessment of physiological systems and states using recurrence plot strategies," *Journal of Applied Physiology*, vol. 76, no. 2, pp. 965–973, 1994.
- [10] J. P. Zbilut and C. L. Webber, "Embeddings and delays as derived from quantification of recurrence plots," *Physics Letters A*, vol. 171, no. 3–4, pp. 199–203, 1992.
- [11] N. Marwan, M. C. Romano, M. Thiel, and J. Kurths, "Recurrence plots for the analysis of complex systems," *Physics Reports*, vol. 438, no. 5–6, pp. 237–329, 2007.

- [12] N. Marwan, N. Wessel, U. Meyerfeldt, A. Schirdewan, and J. Kurths, "Recurrence-plot-based measures of complexity and their application to heart-rate-variability data," *Physical Review E*, vol. 66, no. 2, Article ID 026702, 8 pages, 2002.
- [13] G. P. Liu, J. J. Yan, Y. Q. Wang et al., "Application of multilabel learning using the relevant feature for each label in chronic gastritis syndrome diagnosis," *Evidence-Based Complementary and Alternative Medicine*, vol. 2012, Article ID 135387, 9 pages, 2012.
- [14] "Nomenclature and criteria for diagnosis of ischemic heart disease. Report of the Joint International Society and Federation of Cardiology/World Health Organization task force on standardization of clinical nomenclature," *Circulation*, vol. 59, no. 3, pp. 607–609, 1979.
- [15] Y. Q. Wang, *Diagnostics of TCM*, Chinese Medicine Science and Technology Press, Beijing, China, 2004.
- [16] G. P. Liu, Y. Q. Wang, and Y. Dong, "Development and evaluation of scale for heart system Inquiry of TCM," *Journal of Chinese Integrative Medicine*, vol. 7, no. 1, pp. 1222–1225, 2009.
- [17] Z. F. Fei, *TCM Pulse Diagnosis*, People's Medical Publishing House, Beijing, China, 2003.
- [18] J. B. Gao, "Detecting nonstationarity and state transitions in a time series," *Physical Review E*, vol. 63, no. 6, Article ID 066202, 8 pages, 2001.
- [19] N. Marwan and J. Kurths, "Nonlinear analysis of bivariate data with cross recurrence plots," *Physics Letters A*, vol. 302, no. 5–6, pp. 299–307, 2002.
- [20] M. B. Kennel, R. Brown, and H. D. I. Abarbanel, "Determining embedding dimension for phase-space reconstruction using a geometrical construction," *Physical Review A*, vol. 45, no. 6, pp. 3403–3411, 1992.
- [21] L. Cao, "Practical method for determining the minimum embedding dimension of a scalar time series," *Physica D*, vol. 110, no. 1–2, pp. 43–50, 1997.
- [22] A. M. Fraser and H. L. Swinney, "Independent coordinates for strange attractors from mutual information," *Physical Review A*, vol. 33, no. 2, pp. 1134–1140, 1986.
- [23] J. B. Gao, "Recurrence time statistics for chaotic systems and their applications," *Physical Review Letters*, vol. 83, no. 16, pp. 3178–3181, 1999.
- [24] L. Xu, M. Q. H. Meng, X. Qi, and K. Wang, "Morphology variability analysis of wrist pulse waveform for assessment of arteriosclerosis status," *Journal of Medical Systems*, vol. 34, no. 3, pp. 331–339, 2010.
- [25] M. L. Zhang, "LIFT: multi-label learning with label-specific features," in *Proceedings of the 22nd International Joint Conference on Artificial Intelligence*, pp. 1609–1614, 2011.
- [26] Y. Q. Wang, H. X. Yan, R. Guo et al., "Study on intelligent syndrome differentiation in traditional Chinese medicine based on multiple information fusion methods," *International Journal of Data Mining and Bioinformatics*, vol. 5, no. 4, pp. 369–382, 2011.
- [27] R. E. Schapire and Y. Singer, "BoosTexter: a boosting-based system for text categorization," *Machine Learning*, vol. 39, no. 2, pp. 135–168, 2000.
- [28] G. Tsoumakas, M. L. Zhang, and Z. H. Zhou, *Tutorial on Learning from Multi-Label Data*, ECML/PKDD, Bled, Slovenia, 2009, <http://www.ecmlpkdd2009.net/wp-content/uploads/2009/08/learning-from-multi-label-data.pdf>.
- [29] Z. C. Luo, S. Zhang, and Y. M. Yang, *Engineering Analysis for Pulse Wave and Its Application in Clinical Practice*, Science Press, Beijing, China, 2006.