



Breast Cancer Index: A Perspective on Tongue Diagnosis in Traditional Chinese Medicine

Lun-Chien Lo^{1,2}, Tsung-Lin Cheng², John Y. Chiang³, Natsagdorj Damdinsuren⁴

¹Department of Traditional Chinese Medicine, Changhua Christian Hospital, Changhua, Taiwan.

²Graduate Institute of Statistics and Information Science, National Changhua University of Education, Changhua, Taiwan.

³Department of Computer Science and Engineering, National Sun Yat-sen University, Kaohsiung, Taiwan.

⁴Manba Datsan, Clinic and Training Centre for Mongolian Traditional Medicine. 2nd Khoroo, Bayanzurkh district, Ulaanbaatar, Mongolia.

ABSTRACT

Breast cancer (BC) ranks second in the cancer fatality rate among females worldwide. Mammogram, ultrasound, magnetic resonance imaging (MRI), blood testing, and fine needle aspiration biopsy are usually applied to discriminate BC patients from normal persons. False-negative results, undetectable calcifications, movement-incurred blurry image, infection, and sampling error are commonly associated with these traditional means of diagnosis. Traditional Chinese medicine (TCM) covers a broad range of medical practices sharing common theoretical concepts. Tongue diagnosis plays an important role in TCM. Organ conditions, properties, and variation of pathogens can be revealed through observation of tongue. In light of this observation, this paper investigates discriminating tongue features to distinguish between BC patients and normal people, and establishes differentiating index to facilitate the non-invasive detection of BC. The tongue features for 60 BC patients and 70 normal persons were extracted by the Automatic Tongue Diagnosis System (ATDS). The Mann–Whitney test showed that the amount of tongue fur ($P = 0.007$), tongue fur in the spleen-stomach area, maximum covering area of tongue fur, thin tongue fur, the number of tooth marks, the number of red dots, red dot in the spleen-stomach area, red dot in the liver-gall-left area, red dot in the liver-gall-right area, and red dot in the heart-lung area demonstrated significant differences ($P < 0.05$). The tongue features of the testing group were employed to test the power of significant tongue features identified in predicting BC. An accuracy of 80% was reached by applying the seven significant tongue features obtained through Mann–Whitney test. To the best of our knowledge, this is the first attempt in applying TCM tongue diagnosis to the discrimination of BC patients and normal persons.

Key words: Automatic tongue diagnosis system, Breast cancer, Logistic regression, Mann–Whitney test

INTRODUCTION

Breast cancer (BC) occurs when breast cells develop abnormally and grow out of control, forming a malignant tumor that can spread to other parts of the body.^[1] BC that started off in the lobules is known as lobular carcinoma, while the one that developed from the ducts is called ductal carcinoma.^[2] Mam-

mogram, ultrasound, magnetic resonance imaging (MRI), blood testing, and fine needle aspiration biopsy are usually applied to discriminate BC patients (BCPs) from normal persons (NPs).^[3,4] However, other than being invasive or radioactive, some disadvantages are associated with these traditional BC diagnoses, for example, false-negative results of mammogram, undetectability of breast calcifications with ultrasound, unclear image caused by

Correspondence to:

Prof. John Y. Chiang, Department of Computer Science and Engineering, National Sun Yat-sen University, Kaohsiung, Taiwan,
Tel: 00886-7-5252000 ~ 4343, Fax: 00886-7-5254301, Email: chiang@cse.nsysu.edu.tw

DOI: 10.4103/2225-4110.114901

movement during MRI scanning, infection of blood testing, and cancer missing in fine needle aspiration biopsy if the needle is not placed among the cancer cells. Due to the above pitfalls, the early diagnosis of BC may be difficult.

Nicolao *et al.*, (2010) reported that the traditional Chinese medicine (TCM) practices, such as acupuncture and phytotherapy, are considered as the most popular disciplines requested by both medical experts and students in Switzerland.^[5] They have received wider acceptance from western medicine in recent years. The essence of TCM hinges on “correct differentiation of symptoms for proper means of treatment.” Correct diagnosis is a prerequisite for effective medical treatment. TCM diagnosis is generally based on four standard but not validated approaches, that is, observation, smelling/listening, inquiry, and palpation. Observation tops the four ways of TCM diagnosis and the tongue corresponds to the major subject of focus during observation.^[6] Therefore, tongue diagnosis plays an important role in TCM.^[7]

It is widely believed that the tongue is connected to the internal organs through meridians; thus, the conditions of organs, qi, blood, and body fluids, as well as the degree and progression of disease are all reflected on the tongue.^[8,9] Organ conditions, properties, and variations of pathogens can be found through observation of tongue. For example, changes in the tongue property primarily reflect organ status and the flow of qi and blood; variations in tongue fur can be employed to determine the impact of exogenous pathogenic factors and the flow of stomach qi. In clinical practice of TCM, practitioners observe the characteristics of tongue, such as the color, shape, and the amount of saliva, before deducing the primary ailment of a patient. However, observation diagnosis is often biased by subjective judgment, originating from personal knowledge, experience, thinking patterns, diagnostic skills, and color perception or interpretation. There are no precise or quantifiable standards existing. Different practitioners may pass varying judgments on the same tongue, while a practitioner may even reach inconsistent diagnoses on the identical tongue if examined at different time.

Previous studies that have been conducted on the issue regarding consistency of TCM diagnosis^[10-20] as well as herbal prescription or treatment^[14,15,17] indicated that inter- and intra-observer agreements are low. The inconsistency of subjective diagnosis and treatment can be improved by the development of validated instruments such as standard questionnaire for inquiries^[6] and manual for guiding treatments.^[21] Recently, experiments on intra- and inter-observer agreements of the automatic tongue diagnosis system (ATDS) and TCM practitioners have been conducted in our laboratory.^[20] The results demonstrate that the ATDS is very consistent even in the face of variations of environmental lighting and extruding tongue, with an intra-observer agreement being significantly higher than that of the TCM doctors, while the inter-observer agreements between the ATDS and a group of TCM doctors and among the TCM doctors are both moderate.^[20,22-24] ATDS serves not only as clinical equipment in providing doctors with consistent tongue features of patients but also as feasible teaching and evaluation means for students learning tongue diagnosis.

This paper investigates discriminating tongue features to distinguish between BCPs and normal people, and establishes

differentiating index to facilitate the non-invasive detection of BC. Nine tongue features, namely tongue color, tongue quality, tongue fissure, tongue fur, red dot, ecchymosis, tooth mark, saliva, and tongue shape, were extracted for BCPs and NPs by the ATDS.^[22-24] Features identified were further subdivided according to the areas located (i.e., spleen-stomach, liver-gall-left, liver-gall-right, kidney, and heart-lung area) as shown in Figure 1. The Mann–Whitney test was performed based on the data collected to induce significant tongue features ($P < 0.05$) to discriminate BCPs from NPs.^[25,26] Ten tongue features with significant differences, identified by the Mann–Whitney test, were further employed as factors in the logistic regression to derive features with independently significant meaning. Three logistic regression models were constructed to predict the probability of getting BC. An accuracy of 80%, 80%, and 67% was achieved, respectively. To the best of our knowledge, this is the first attempt in applying TCM tongue diagnosis to the discrimination of BCPs from NPs.

MATERIALS AND METHODS

Participants

Two groups of the tongue images, that is, experimental and control groups, were collected through the outpatients of Department of Traditional Chinese Medicine of Changhua Christian Hospital (CCH) in Taiwan (IRB no. 120512). The tongue features for 60 BCPs in the experimental group and 70 NPs in the control group were extracted by the ATDS, respectively. The inclusion and exclusion criteria for the subjects of the experimental group are as follows:

- *Inclusion criteria*
 - Diagnosed as BCP (ICD-9-CM 174-174.9) by a specialist
 - Both males and females were included
 - Patients who had signed IRB agreement.
- *Exclusion criteria*
 - Patients who had not signed IRB agreement
 - Pregnant women
 - Patients with acute infection
 - Cognitive impaired, for example, imbecility dementia or delirium, caused by cancer metastasis to brain
 - Patients unable to protrude the tongue stably.

Figure 2 shows the exemplary tongue images of the BCPs and NPs recruited.

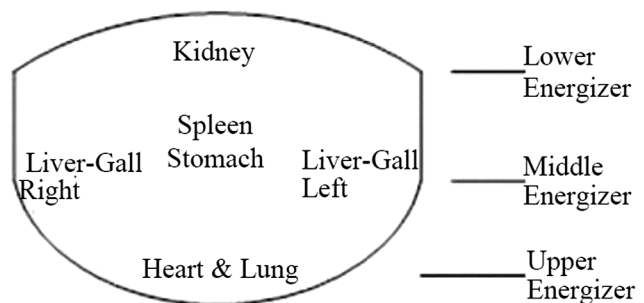


Figure 1. The tongue is subdivided into areas corresponding to different internal organs

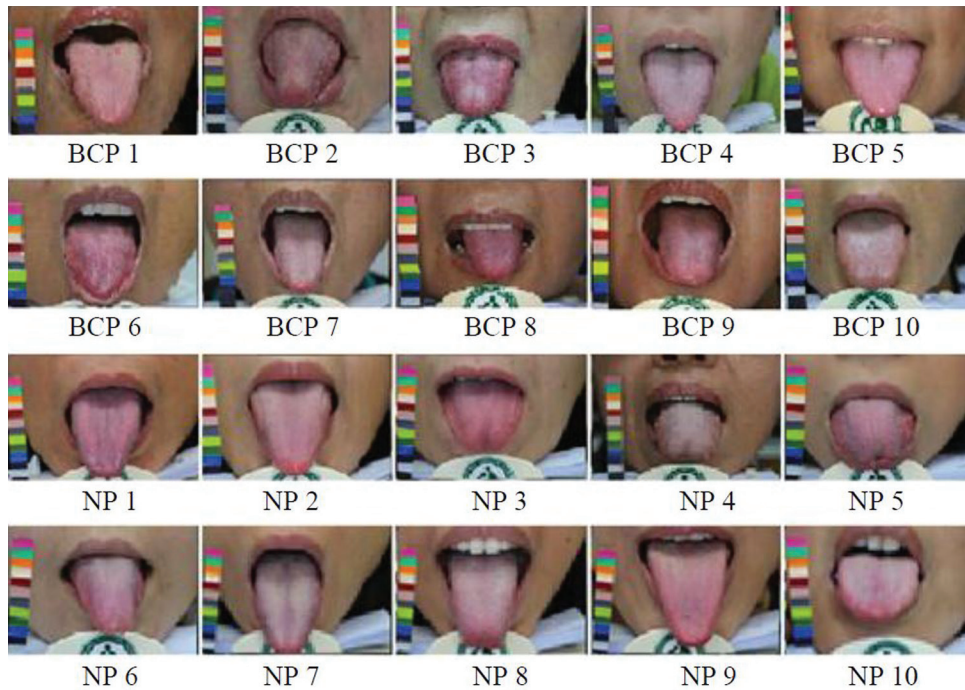


Figure 2. Tongue images of breast cancer patients and normal persons

Automatic tongue diagnosis system

As shown in Figure 3a, the ATDS was developed to capture the tongue images and extract features reliably to assist the diagnosis of TCM practitioners.^[20] Figure 3b demonstrates the steps in three major functions, that is, image capturing and color calibration, tongue areas' segmentation, and tongue feature extraction, included in the ATDS.^[20,23]

Variations in background lighting may change the color and brightness of the acquired images, greatly affecting the consistency and stability of the extracted tongue features. The consistency and stability of the tongue images captured and the features extracted were achieved by calibrating brightness and color to compensate the variations in intensity and color temperature of the light source and the imaging hardware. The ATDS developed can automatically correct lighting and color deviation caused by the change of background lighting with a color bar placed beside the subject. Color calibration utilizes the information provided by the color bar to make sure the image quality is consistent even when taken in different circumstances. Figure 4a and b displays the images taken at T_1 before and after color calibration, respectively, whereas Figure 4c demonstrates the image taken at T_2 after calibration. The second and third rows show the color bars clipped from the tongue images and their corresponding histograms. ATDS automatically compensates the color deviation of the original image [Figure 4a, with a mean gray level 59.5] to allow colors in images taken at different time intervals consistent with each other [Figure 4b and c, with mean gray levels 66.1 and 67.3, respectively].

Tongue images were analyzed by first isolating the tongue region within an image to eliminate irrelevant lower facial portions and background surrounding the tongue, thereby facilitating feature identification and extraction, and then extracting the tongue features by employing criteria such as the aspect ratio, color composition, location, shape, and color distribution of the tongue, as well as the

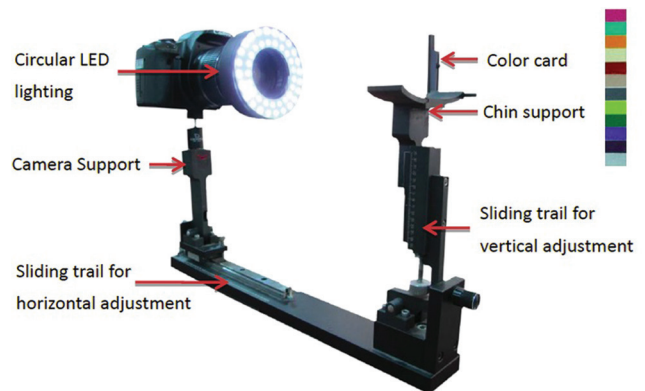


Figure 3a. Components of the ATDS

values of neighboring pixels. Nine primary features including tongue color, fur color, fur thickness, saliva, tongue shape, tongue fissure, red dot, ecchymosis, and tooth marks were extracted to generate detailed information regarding length, area, moisture, and the number of fissures, tooth marks, and red dots to be employed in tongue diagnosis, as depicted in Figure 5.^[18] Features identified were further subdivided according to the areas located (i.e., spleen-stomach, liver-gall-left, liver-gall-right, kidney, and heart-lung area), as shown in Figure 1.

A complete listing of the tongue features extracted is summarized below:

1. Tongue color: Includes slightly white, slightly red, red, dark red, and dark purple
2. Tongue shape: Includes shape (thin and small, moderate, fat and large) and tongue body (normal, tilted to the left, tilted to the right)
3. Saliva: Includes total area and the amount of saliva (none, little, normal, excessive)

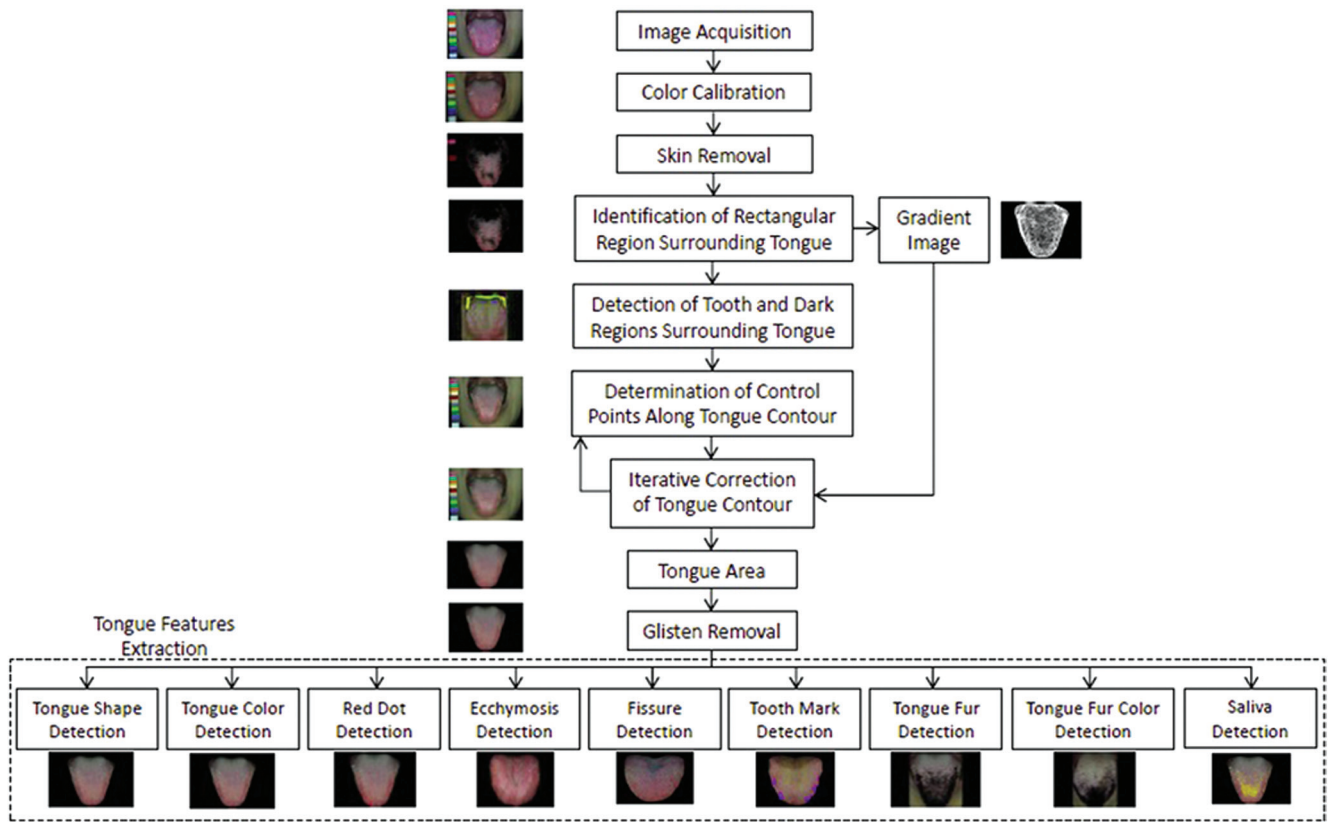


Figure 3b. The processing steps of ATDS analysis

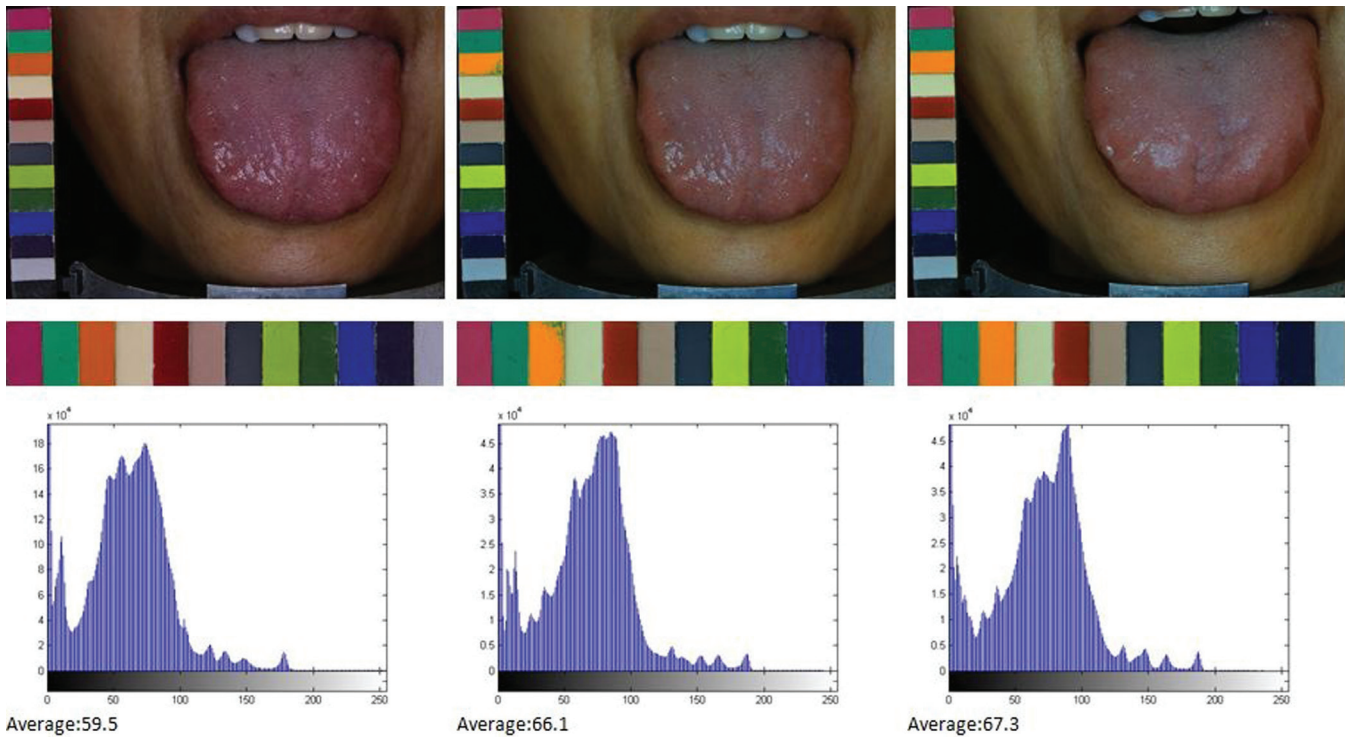


Figure 4. Calibration of image color using the color bar accompanying the subject to make the image quality consistent for images taken in different circumstances; a) T_1 before color calibration, and the corresponding color bar histogram; b) T_1 after color calibration, and the corresponding color bar histogram; c) T_2 after calibration, and the corresponding color bar histogram

4. Tongue fur: Includes fur color (white, yellow, dye), amount, average covering area, maximum covering area, minimum covering area, degree of thickness (none, thin, thick), and organs corresponding to the covering area (spleen-stomach, liver-gall-left, liver-gall-right, kidney, heart-lung areas)
5. Tongue quality: Organs corresponding to the covering area (spleen-stomach, liver-gall-left, liver-gall-right, kidney, heart-lung areas)
6. Tongue fissure: Includes amount, average covering area, shortest length, longest length
7. Ecchymosis: Amount, average covering area, maximum covering area, minimum covering area, and organs corresponding to the covering area (spleen-stomach, liver-gall-left, liver-gall-right, kidney, heart-lung areas)
8. Tooth mark: Includes number, average covering area, maximum covering area, minimum covering area, and organs corresponding to the covering area (spleen-stomach, liver-gall-left, liver-gall-right, kidney, heart-lung areas)
9. Red dot: Includes number, average covering area, maximum covering area, minimum covering area, and organs corresponding to the covering area (spleen-stomach, liver-gall-left, liver-gall-right, kidney, heart-lung areas).

The tongue features of BCP 1 are given in Table 1.

Statistical analysis

The tongue features of the subjects participating in the study were extracted by ATDS. The Mann–Whitney test was performed on the data sets acquired in the experimental and control groups to identify the features with significant difference ($P < 0.05$). The Mann–Whitney test is a non-parametric test used to compare two independent groups of sampled data and without the condition of normal distributions. The test statistic for the Mann–Whitney test is U. This value is compared to a table of critical values for U based on the sample size of each group.

Besides, in classifying two distinct groups, logistic regression can be used to predict the outcome of a categorical dependent variable based on one or more predictor variables.^[27] It is used with data in which there is a binary (success–failure) outcome. Let P be the predicted probability of success using covariate x , i.e.,

$$P = \frac{e^{f(x)}}{1 + e^{f(x)}} \tag{1}$$

where $f(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k$ and $\beta_i, i \in Z, i \geq 0$ are the corresponding coefficients. The statistical analysis of this study is twofold: First, to conduct Mann–Whitney test, with respect to the tongue features, for the BC group and the normal group to select significant variables for the following logistic regression and second, to conduct logistic regression based on the selection of the variables in the former Mann–Whitney test to obtain a predicting equation for the probability of with/without BC.

EXPERIMENTAL RESULTS

The tongue features for 60 BCPs and 70 NPs were extracted by

the ATDS.^[22-24] The purpose focuses on inducing significant tongue features ($P < 0.05$) to discriminate BCPs from NPs. The results

Table 1. Tongue features of BCP 1 in Figure 2, extracted by ATDS

Tongue color	Red
Tongue shape	Moderate
Tongue body	Normal
Saliva	Total area: 0.21 cm ² Amount: little
Tongue fur	Color: white Amount: 44% Average covering area: 4.13 cm ² Maximum covering area: 15.77 cm ² Minimum covering area: 0.31 cm ² Thick: 57% Thin: 43% In the spleen-stomach area: 54% In the liver-gall-left area: 18% In the liver-gall-right area: 37% In the kidney area: 98% In the heart-lung area: 6%
Tongue quality	In the spleen-stomach area: 46% In the liver-gall-left area: 82% In the liver-gall-right area: 63% In the kidney area: 2% In the heart-lung area: 94%
Tongue fissure	Amount: 0 Average covering area: 0 cm ² Shortest length: 0 cm Longest length: 0 cm
Ecchymosis	Amount: 0 Average covering area: 0 cm ² Maximum covering area: 0 cm ² Minimum covering area: 0 cm ² In the spleen-stomach area: 0 In the liver-gall-left area: 0 In the liver-gall-right area: 0 In the kidney area: 0 In the heart-lung area: 0
Tooth mark	Number: 3 Average covering area: 0.12 cm ² Maximum covering area: 0.14 cm ² Minimum covering area: 0.12 cm ² In the spleen-stomach area: 1 In the liver-gall-left area: 1 In the liver-gall-right area: 1 In the kidney area: 0 In the heart-lung area: 0
Red dot	Number: 21 Average covering area: 0 cm ² Maximum covering area: 0.001 cm ² Minimum covering area: 0 cm ² In the spleen-stomach area: 6 In the liver-gall-left area: 10 In the liver-gall-right area: 2 In the kidney area: 0 In the heart-lung area: 3

BCPs: Breast cancer patients; ATDS: Automatic tongue diagnosis system

obtained by applying Mann–Whitney test are listed in Table 2, with 10 features, namely, the amount of tongue fur ($P = 0.007$), tongue fur in the spleen-stomach area ($P = 0.020$), maximum covering area of tongue fur ($P = 0.002$), thin tongue fur ($P = 0.000$), the number of tooth marks ($P = 0.050$), the number of red dots ($P = 0.000$), red dot in the spleen-stomach area ($P = 0.000$), red dot in the liver-gall-left area ($P = 0.002$), red dot in the liver-gall-right area ($P = 0.000$), red dot in the heart-lung area ($P = 0.003$), demonstrating significant differences.

Next, the data collected were classified into two groups. The training group consisted of 55 BCPs and 60 NPs, while the testing group was composed of 5 BCPs and 10 NPs. The logistic regression by utilizing these 10 tongue features with significant

differences in Mann–Whitney test as factors was performed. Table 3 lists the results obtained.

Model I

The tongue features of the testing group were employed in two logistic regression models to test the power of significant tongue features identified in predicting BC. In deriving the first model, the 10 tongue features with significant difference, identified by Mann–Whitney test, were employed as factors $x_i, i \leq 10$ in deciding the subject with or without BC. Let P represent the predicted probability of infecting BC using logistic regression. The coefficients in Equation (1) can be determined as:

$$f(x) = 7.330 + 0.4075x_1 - 0.4741x_2 + 0.2972x_3 - 0.08705x_4 - 0.2347x_5 + 0.0767x_6 - 0.1992x_7 - 0.0166x_8 + 0.0367x_9 - 0.0921x_{10} \tag{2}$$

where x_1 is the amount of tongue fur, x_2 the tongue fur in the spleen-stomach area, x_3 the maximum covering area of tongue fur, x_4 the thin tongue fur, x_5 the number of tooth marks, x_6 the number of red dots, x_7 the red dot in the spleen-stomach area, x_8 the red dot in the liver-gall-left area, x_9 the red dot in the liver-gall-right area, and x_{10} is the red dot in the heart-lung area.

The accuracy of Model I in predicting BC can be tested by substituting the values of tongue features extracted from the testing group into Equation (2). According to Equation (1), the probability of infecting BC P is determined once the value of $f(x)$ is given. A predicted probability larger than 0.5 corresponds to tongue features of a BCP, while that smaller than 0.5 represents those of an NP. Table 4 lists the probability of infecting BC for 5 BCPs and 10 NPs in the testing group according to Model I. Among them, the probability of infecting BC is larger than 0.5 for BCPs 1, 2, 3, and 4 and smaller than 0.5 for NPs 3, 4, 5, 6, 7, 8, 9, and 10. Correct diagnoses are reached for these 12 cases out of a total of 15 ones. An accuracy of 80% is reached by employing Model I in predicting BC through tongue diagnosis.

Model II

Table 3 lists the results by employing logistic regression to

Table 2. Significant tongue features identified by applying Mann-Whitney test to the data sets acquired from the group of normal persons and the group of breast cancer patients

Significant tongue features	Group	Sample size	Median	P value
The amount of tongue fur	NP	70	43.5	0.007
	BCP	60	55.0	
The tongue fur in the spleen-stomach area	NP	70	49.0	0.020
	BCP	60	58.5	
Maximum covering area of tongue fur	NP	70	10.62	0.002
	BCP	60	20.59	
Thin tongue fur	NP	70	68.0	0.000
	BCP	60	52.0	
The number of tooth marks	NP	70	2.5	0.050
	BCP	60	3.0	
The number of red dots	NP	70	137.0	0.000
	BCP	60	63.0	
Red dot in the spleen-stomach area	NP	70	52.0	0.000
	BCP	60	0	
Red dot in the liver-gall-left area	NP	70	2.0	0.002
	BCP	60	10.0	
Red dot in the liver-gall-right area	NP	70	0	0.000
	BCP	60	11.5	
Red dot in the heart-lung area	NP	70	60.0	0.003
	BCP	60	31.5	

NP: Normal persons; BCP: Breast cancer patients

Table 3. The results of applying the logistic regression by utilizing 10 tongue features with significant differences identified in Mann-Whitney test as factors

Predictor	Coefficient	SE coefficient	Z	P	Odds ratio	Odds ratio 95% CI	
						Lower	Upper
Constant	7.330	4.871	1.50	0.132			
The amount of tongue fur	0.4075	0.1794	2.26	0.024	1.50	1.06	2.13
The tongue fur in the spleen-stomach area	-0.4741	0.2137	-2.22	0.026	0.62	0.41	0.95
Maximum covering area of tongue fur	0.2972	0.1159	2.56	0.010	1.35	1.07	1.69
Thin tongue fur	-0.08705	0.03726	-2.34	0.019	0.92	0.85	0.99
The number of tooth marks	-0.2347	0.3116	-0.75	0.451	0.79	0.43	1.46
The number of red dots	0.0767	0.3106	0.25	0.805	1.08	0.59	1.98
Red dot in the spleen-stomach area	-0.1992	0.3080	-0.65	0.518	0.82	0.45	1.50
Red dot in the liver-gall-left area	-0.0166	0.3202	-0.05	0.959	0.98	0.53	1.84
Red dot in the liver-gall-right area	0.0367	0.3063	0.12	0.905	1.04	0.57	1.89
Red dot in the heart-lung area	-0.0921	0.03089	-0.30	0.766	0.91	0.50	1.67

analyze the relationship between dependent and independent variables, under the hypothesis of infecting BC or not. We removed one of the 10 tongue features which did not show the most significant difference ($P > 0.05$) and performed logistic regression three times. Tables 5-7 lists the final results obtained. Among them, the amount of tongue fur ($P = 0.011$), the tongue

fur in the spleen-stomach area ($P = 0.006$), the maximum covering area of tongue fur ($P = 0.003$), thin tongue fur ($P = 0.019$), the number of red dots ($P = 0.017$), red dot in the spleen-stomach area ($P = 0.002$), and red dot in the heart-lung area ($P = 0.016$) exhibit significant differences ($P = 0.05$). Let P represent the probability of infecting BC. Considering the significant vari-

Table 4. The probability of infecting breast cancer by employing Model I to the tongue features of 5 breast cancer patients and 10 normal persons in the testing group

Group	The amount of tongue fur	Tongue fur in the spleen-stomach area	The maximum covering area of tongue fur	Thin tongue fur	The number of tooth marks	The number of red dots	Red dot in the spleen-stomach area	Red dot in the liver-gall-left area	Red dot in the liver-gall-right area	Red dot in the heart-lung area	Model I	Probability of infecting breast cancer
BCP 1	55	60	25.36	57	3	86	0	26	44	16	9.37	0.99
BCP 2	77	83	32.78	61	5	65	0	0	11	54	2.89	0.95
BCP 3	19	16	2.02	64	0	169	0	79	40	50	10.99	0.99
BCP 4	30	40	16.78	73	5	186	0	25	55	106	4.10	0.98
BCP 5	7	9	1.63	92	3	172	0	38	7	127	-1.20	0.23
NP 1	90	91	28.12	91	4	0	0	0	0	0	0.19	0.55
NP 2	50	56	33	70	4	198	0	40	37	121	8.58	0.99
NP 3	39	40	11.83	65	3	193	114	12	7	60	-12.0	0.00
NP 4	50	50	8.9	42	4	21	17	0	0	4	-0.18	0.45
NP 5	19	19	3.03	17	2	192	190	1	0	1	-18.2	0.00
NP 6	51	55	11.77	59	4	45	25	0	0	20	-4.00	0.02
NP 7	72	81	24.72	78	3	49	23	9	0	17	-4.55	0.01
NP 8	58	70	19.37	68	2	113	22	0	0	91	-7.05	0.00
NP 9	26	26	8.78	30	3	288	148	3	9	127	-13.9	0.00
NP 10	11	9	0.62	97	4	100	41	0	8	51	-6.57	0.00

NP: Normal persons; BCP: Breast cancer patients

Table 5. The first results of removing the most insignificant differences of 10 tongue features in Table 3 and applying the logistic regression

Predictor	Coefficient	SE coefficient	Z	P	Odds ratio	Odds ratio 95% CI	
						Lower	Upper
Constant	7.45	4.307	1.73	0.084			
The amount of tongue fur	0.4084	0.1729	2.36	0.018	1.50	1.07	2.11
Tongue fur in the spleen-stomach area	-0.4784	0.1984	-2.41	0.016	0.62	0.42	0.91
Maximum covering area of tongue fur	0.2996	0.1070	2.80	0.005	1.35	1.09	1.66
Thin tongue fur	-0.08755	0.03614	-2.42	0.015	0.92	0.85	0.98
The number of tooth marks	-0.2375	0.3062	-0.78	0.438	0.79	0.43	1.44
The number of red dots	0.06089	0.06023	1.01	0.312	1.06	0.94	1.20
Red dot in the spleen-stomach area	-0.18383	0.07907	-2.33	0.020	0.83	0.71	0.97
Red dot in the liver-gall-right area	0.0513	0.1170	0.44	0.661	1.05	0.84	1.32
Red dot in the heart-lung area	-0.07647	0.06064	-1.26	0.207	0.93	0.82	1.04

Table 6. The second results of removing the most insignificant differences of nine tongue features in Table 5 and applying the logistic regression

Predictor	Coefficient	SE coefficient	Z	P	Odds ratio	Odds ratio 95% CI	
						Lower	Upper
Constant	7.807	4.471	1.75	0.081			
The amount of tongue fur	0.4343	0.1705	2.55	0.011	1.54	1.11	2.16
Tongue fur in the spleen-stomach area	-0.5152	0.1922	-2.68	0.007	0.60	0.41	0.87
Maximum covering area of tongue fur	0.3160	0.1047	3.02	0.003	1.37	1.12	1.68
Thin tongue fur	-0.08701	0.03598	-2.42	0.016	0.92	0.85	0.98
The number of tooth marks	-0.2130	0.3036	-0.70	0.483	0.81	0.45	1.47
The number of red dots	0.08338	0.03462	2.41	0.016	1.09	1.02	1.16
Red dot in the spleen-stomach area	-0.20554	0.06689	-3.07	0.002	0.81	0.71	0.93
Red dot in the heart-lung area	-0.09767	0.03977	-2.46	0.014	0.91	0.84	0.98

Table 7. The third results of removing the most insignificant differences of eight tongue features in Table 6 and applying the logistic regression

Predictor	Coefficient	SE coefficient	Z	P	Odds ratio	Odds ratio 95% CI	
						Lower	Upper
Constant	5.921	3.027	1.96	0.050			
The amount of tongue fur	0.4049	0.1599	2.53	0.011	1.50	1.10	2.05
Tongue fur in the spleen-stomach area	-0.4702	0.1696	-2.77	0.006	0.62	0.45	0.87
Maximum covering area of tongue fur	0.29514	0.09826	3.00	0.003	1.34	1.11	1.63
Thin tongue fur	-0.08049	0.03438	-2.34	0.019	0.92	0.86	0.99
The number of red dots	0.08255	0.03457	2.39	0.017	1.09	1.01	1.16
Red dot in the spleen-stomach area	-0.19592	0.06242	-3.14	0.002	0.82	0.73	0.93
Red dot in the heart-lung area	-0.09414	0.03919	-2.40	0.016	0.91	0.84	0.98

Table 8. The probability of infecting breast cancer by employing Model II to the tongue features of 5 breast cancer patients and 10 normal persons in the testing group

Group	The amount of tongue fur	Tongue fur in the spleen-stomach area	Maximum covering area of tongue fur	Thin tongue fur	The number of red dots	Red dot in the spleen-stomach area	Red dot in the heart-lung area	Model II	Probability of infecting breast cancer
BCP 1	55	60	25.36	57	86	0	16	8.47	0.99
BCP 2	77	83	32.78	61	65	0	54	3.12	0.96
BCP 3	19	16	2.02	64	169	0	50	10.78	0.99
BCP 4	30	40	16.78	73	186	0	106	3.71	0.98
BCP 5	7	9	1.63	92	172	0	127	-0.16	0.46
NP 1	90	91	28.12	91	0	0	0	0.55	0.63
NP 2	50	56	33	70	198	0	121	8.89	0.99
NP 3	39	40	11.83	65	193	114	60	-10.89	0
NP 4	50	50	8.9	42	21	17	4	-0.07	0.48
NP 5	19	19	3.03	17	192	190	1	-17.26	0
NP 6	51	55	11.77	59	45	25	20	-3.63	0.03
NP 7	72	81	24.72	78	49	23	17	-4.06	0.02
NP 8	58	70	19.37	68	113	22	91	-6.81	0.00
NP 9	26	26	8.78	30	288	148	127	-12.78	0
NP 10	11	9	0.62	97	100	41	51	-6.06	0.00

NP: Normal persons; BCP: Breast cancer patients

ables in the above logistic regression, we consider the following model:

$$f(x) = 5.921 + 0.4049x_1 - 0.4702x_2 + 0.29514x_3 - 0.08049x_4 + 0.08255x_5 - 0.19592x_6 - 0.09414x_7 \quad (3)$$

where x_1 is the amount of tongue fur, x_2 the tongue fur in the spleen-stomach area, x_3 the maximum covering area of tongue fur, x_4 the thin tongue fur, x_5 the number of red dots, x_6 the red dot in the spleen-stomach area, and x_7 is the red dot in the heart-lung area.

The accuracy of using Model II in predicting BC can be tested by substituting the values of tongue features extracted from the testing group into Equation (3). According to Equation (1), the probability of infecting BC P is determined once the value of $f(x)$ is given. A predicted probability larger than 0.5 corresponds to tongue features of a BCP, while that smaller than 0.5 represents those of an NP. Table 8 lists the probability of infecting BC for 5 BCPs and 10 NPs in the testing group according to Model II. Among them, the probability of infecting BC is larger than 0.5 for BCPs 1, 2, 3, and 4 and smaller than 0.5 for NPs 3, 4, 5, 6, 7, 8, 9, and 10. Correct diagnoses are reached for these 12 cases out of a total of 15 ones. An accuracy

of 80% is reached by employing Model II in predicting BC through tongue diagnosis.

Model III

Table 7 lists the results by employing logistic regression to analyze the relationship between dependent and independent variables, under the hypothesis of infecting BC or not. We removed the least significant differences, thin tongue fur ($P = 0.019$), of seven tongue features and performed logistic regression. Table 9 lists the final results obtained. Among them, the amount of tongue fur ($P = 0.037$), tongue fur in the spleen-stomach area ($P = 0.005$), the maximum covering area of tongue fur ($P = 0.001$), the number of red dots ($P = 0.008$), red dot in the spleen-stomach area ($P = 0.002$), and red dot in the heart-lung area ($P = 0.005$) exhibit significant difference ($P = 0.05$). Let P represent the probability of infecting BC. Considering the significant variables in the above logistic regression, we consider the following model:

$$f(x) = 2.854 + 0.20646x_1 - 0.2952x_2 + 0.23778x_3 + 0.08462x_4 - 0.18531x_5 - 0.10068x_6, \quad (4)$$

where x_1 is the amount of tongue fur, x_2 the tongue fur in the spleen-stomach area, x_3 the maximum covering area of tongue

Table 9. The results of removing the least significant differences of seven tongue features in Table 7 and applying the logistic regression

Predictor	Coefficient	SE coefficient	Z	P	Odds ratio	Odds ratio 95% CI	
						Lower	Upper
Constant	2.854	2.192	1.30	0.193			
The amount of tongue fur	0.20646	0.09915	2.08	0.037	1.23	1.01	1.46
Tongue fur in the spleen-stomach area	-0.2952	0.1053	-2.80	0.005	0.74	0.61	0.92
Maximum covering area of tongue fur	0.23778	0.07312	325	0.001	1.27	1.10	1.46
The number of red dots	0.08462	0.03175	2.67	0.008	1.09	1.02	1.16
Red dot in the spleen-stomach area	-0.18531	0.05857	-3.16	0.002	0.83	0.74	0.93
Red dot in the heart-lung area	-0.10068	0.03561	-2.83	0.005	0.90	0.84	0.97

Table 10. The probability of infecting breast cancer by employing Model III to the tongue features of 5 breast cancer patients and 10 normal persons in the testing group

Group	The amount of tongue fur	Tongue fur in the spleen-stomach area	Maximum covering area of tongue fur	The number of red dots	Red dot in the spleen-stomach area	Red dot in the heart-lung area	Model III	Probability of infecting breast cancer
BCP 1	55	60	25.36	86	0	16	11.50	0.99
BCP 2	77	83	32.78	65	0	54	6.94	0.99
BCP 3	19	16	2.02	169	0	50	13.23	0.99
BCP 4	30	40	16.78	186	0	106	8.73	0.99
BCP 5	7	9	1.63	172	0	127	5.04	0.99
NP 1	90	91	28.12	0	0	0	6.49	0.99
NP 2	50	56	33	198	0	121	12.76	0.99
NP 3	39	40	11.83	193	114	60	-6.26	0.002
NP 4	50	50	8.9	21	17	4	1.69	0.84
NP 5	19	19	3.03	192	190	1	-16.06	0
NP 6	51	55	11.77	45	25	20	0.20	0.55
NP 7	72	81	24.72	49	23	17	2.16	0.90
NP 8	58	70	19.37	113	22	91	-0.94	0.28
NP 9	26	26	8.78	288	148	127	-10.87	0
NP 10	11	9	0.62	100	41	51	-0.68	0.34

NP: Normal persons; BCP: Breast cancer patients

fur, x_4 the number of red dots, x_5 the red dot in the spleen-stomach area, and x_6 is the red dot in the heart-lung area.

The accuracy of using Model III in predicting BC can be tested by substituting the values of tongue features extracted from the testing group into Equation (4). According to Equation (1), the probability of infecting BC P is determined once the value of $f(x)$ is given. A predicted probability larger than 0.5 corresponds to the tongue features of a BCP, while that smaller than 0.5 represents those of an NP. Table 10 lists the probability of infecting BC for 5 BCPs and 10 NPs in the testing group according to Model II. Among them, the probability of infecting BC is larger than 0.5 for BCPs 1, 2, 3, 4, and 5 and smaller than 0.5 for NPs 3, 5, 8, 9, and 10. Correct diagnoses are reached for these 10 cases out of a total of 15 ones. An accuracy of 67% is reached by employing Model III in predicting BC through tongue diagnosis.

CONCLUSIONS

This study investigates discriminating tongue features to distinguish between BCPs and normal people, and establishes differentiating index to facilitate the non-invasive detection

of BC. The Mann-Whitney test showed that the amount of tongue fur ($P = 0.007$), tongue fur in the spleen-stomach area ($P = 0.020$), maximum covering area of tongue fur ($P = 0.002$), thin tongue fur ($P = 0.000$), the number of tooth marks ($P = 0.050$), the number of red dots ($P = 0.000$), red dot in the spleen-stomach area ($P = 0.000$), red dot in the liver-gall-left area ($P = 0.002$), red dot in the liver-gall-right area ($P = 0.000$), and red dot in the heart-lung area ($P = 0.003$) demonstrated significant differences. Next, logistic regression by utilizing these 10 tongue features with significant differences in Mann-Whitney test as factors was performed. We removed one of the 10 tongue features which did not show the most significant differences ($P > 0.05$) and performed logistic regression three times. Among them, the amount of tongue fur ($P = 0.011$), tongue fur in the spleen-stomach area ($P = 0.006$), the maximum covering area of tongue fur ($P = 0.003$), thin tongue fur ($P = 0.019$), the number of red dots ($P = 0.017$), red dot in the spleen-stomach area ($P = 0.002$), and red dot in the heart-lung area ($P = 0.016$) revealed independently significant meaning. The tongue features of the testing group were employed in two logistic regression models to test the predicting accuracy of significant tongue features identified in predicting BC. An accuracy of 80% was reached through the first logistic regression model by applying

the 10 significant tongue features obtained through Mann–Whitney test. For the second model employing seven tongue features obtained by logistic regression with independently significant meaning, 80% accuracy was achieved. The TCM tongue diagnosis can serve as a preliminary screening procedure in early detection of BC in light of its simple and non-invasive nature, followed by other more accurate testing process. To the best of our knowledge, this is the first attempt in applying non-invasive TCM tongue diagnosis to the discrimination of BCPs and NPs.

ACKNOWLEDGMENTS

We are grateful to Y. E. Kao for his assistance of compiling the BC data.

REFERENCES

1. McCartney M. How useful are lifetime risks of disease? *BMJ* 2011;342:d1046.
2. Wasif N, Maggard MA, Ko CY, Giuliano AE. Invasive lobular vs. ductal breast cancer: A stage-matched comparison of outcomes. *Ann Surg Oncol* 2010;17:1862-9.
3. Armstrong K, Handorf EA, Chen J, Bristol Demeter MN. "Breast cancer risk prediction and mammography biopsy decisions: A model-based study. *Am J Prev Med* 2013;44:15-22.
4. Meade E, Dowling M. Early breast cancer: Diagnosis, treatment and survivorship. *Br J Nurs* 2012;21:S4-7.
5. Nicolao M, Täuberand MG, Heusser P. How should complementary and alternative medicine be taught to medical students in Switzerland? A survey of medical experts and students. *Med Teach* 2010;32:50-5.
6. Zhang GG, Lee W, Bausell B, Lao L, Handwerger B, Berman B. Variability in the Traditional Chinese Medicine (TCM) diagnoses and herbal prescriptions provided by three TCM practitioners for 40 patients with rheumatoid arthritis. *J Altern Complement Med* 2005;11:415-21.
7. Kaptchuk T. *Chinese Medicine: The web that has no weaver*, 2nd ed. London: Rider; 2000.
8. Wiseman N, Ellis A. *Fundamentals of Chinese Medicine*. Zhong Yi Xue Ji Chu. Brookline, MA: Paradigm Publications; 1996.
9. Kirschbaum B. *Atlas of Chinese Tongue Diagnosis*. Seattle: Eastland Press; 2000.
10. O'Brien KA, Abbas E, Zhang J, Guo ZX, Luo R, Bensoussan A, et al. An investigation into the reliability of Chinese medicine diagnosis according to Eight Guiding Principles and Zang-Fu Theory in Australians with hypercholesterolemia. *J Altern Complement Med* 2009;15:259-66.
11. O'Brien KA, Abbas E, Movsessian P, Hook M, Komesaroff PA, Birch S. Investigating the reliability of Japanese Toyohari Meridian therapy diagnosis. *J Altern Complement Med* 2009;15:1099-105.
12. Mist S, Ritenbaugh C, Aickin M. Effects of questionnaire-based diagnosis and training on inter-rater reliability among practitioners of traditional Chinese medicine. *J Altern Complement Med* 2009;15:703-9.
13. O'Brien KA, Birch S. A review of the reliability of Traditional East Asian Medicine diagnoses. *J Altern Complement Med* 2009;15:353-66.
14. Coeytaux RR, Chen W, Lindemuth CE, Tan Y, Reilly AC. Variability in the diagnosis and point selection for persons with frequent headache by traditional Chinese acupuncture. *J Altern Complement Med* 2006;12:863-72.
15. Sung JJ, Leung WK, Ching JY, Lao L, Zhang G, Wu JC, et al. Agreements among traditional Chinese medicine practitioners in the diagnosis and treatment of irritable bowel syndrome. *Aliment Pharmacol Ther* 2004;20:1205-10.
16. Zell B, Hirata J, Marcus A, Ettinger B, Pressman A, Ettinger KM. Diagnosis of symptomatic postmenopausal women by traditional Chinese medicine practitioners. *Menopause* 2000;7:129-34.
17. Zhang GG, Lee WL, Lao L, Bausell B, Berman B, Handwerger B. The variability of TCM pattern diagnosis and herbal prescription on rheumatoid arthritis patients. *Altern Ther Health Med* 2004;10:58-63.
18. Zhang GG, Singh B, Lee W, Handwerger B, Lao L, Berman B. Improvement of agreement in TCM diagnosis among TCM practitioners for persons with the conventional diagnosis of Rheumatoid Arthritis: Effect of training. *J Altern Complement Med* 2008;14:381-6.
19. Kim M, Cobbin D, Zaslowski C. Traditional Chinese Medicine tongue inspection: An examination of inter- and intra-practitioner reliability for specific tongue characteristics. *J Altern Complement Med* 2008;14:527-36.
20. Lo LC, Chen YF, Chen WJ, Cheng TL, Chiang JY. The study on the agreement between automatic tongue diagnosis system and traditional Chinese medicine practitioners. *Evid Based Complement Alternat Med* 2012;2012:505063.
21. Schnyer RN, Allen JJ. Bridging the gap in complementary and alternative medicine research: Manualization as a means of promoting standardization and flexibility of treatment in clinical trials of acupuncture. *J Altern Complement Med* 2002;8:623-34.
22. Lo LC, Chen CY, Chiang JY, Cheng TL, Lin HJ, Chang HH. Tongue diagnosis of traditional Chinese medicine for rheumatoid arthritis. *Afr J Tradit Complement Alternat Med* (In press)
23. Hsu YC, Chen YC, Lo L, Chiang JY. "Automatic Tongue Feature Extraction". *Proc. Int'l Computer Symp. (ICS2010)*, Tainan, Taiwan, Dec. 16-18 2010. p. 936-41.
24. Lo LC, Mark Chun-cheng Hou, Chen YL, Chiang JY, Shyu JJ. "Automatic Tongue Diagnosis System," *Proc. 2nd Int'l Conf. on BioMedical Engineering and Informatics (BMEI'09)*, Tianjin, China, 17-19 October, 2009. p. 140-4.
25. Fay MP, Proschan MA. Wilcoxon-Mann-Whitney or t-test? On assumptions for hypothesis tests and multiple interpretations of decision rules. *Stat Surv* 2010;4:1-39.
26. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2012. *CA Cancer J Clin* 2012;62:10-29.
27. van der Heijden H. Decision support for selecting optimal logistic regression models. *Expert Syst Appl* 2012;39:8573-83.