Traditional Chinese Medicine Tongue Diagnosis Index of Early-Stage Breast Cancer

Lun-chien Lo [1,2] , Tsung-Lin Cheng [2] , Yi-Jing Chen [3] , Sainbuyan Natsagdorj [4] , and John Y. Chiang [3]

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

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

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

[4] Department of Mechanical Engineering, Cheng Shiu University, Kaohsiung, Taiwan.

Abstract—**This paper investigates discriminating tongue features to distinguish between early stage BC patients and normal persons via non-invaded methods, expecting to detect BC in the early stage and give treatment in time to increase the recovery rate and lower relapse rate. The tongue features for 67 breast cancer patients of 0 and 1 stages, and 70 normal persons are extracted by the Automatic Tongue Diagnosis System (ATDS) [4-6, 28-31]. A total of nine tongue features, namely, tongue color, tongue quality, tongue fissure, tongue fur, red dot, ecchymosis, tooth mark, saliva, and tongue shape are identified for each tongue. Features extracted are further sub-divided according to the areas located, i.e., spleen-stomach, liver-gall-left, liver-gall-right, kidney, and heart-lung area. The purpose focuses on inducing significant tongue features ($p<0.05$) to discriminate early-stage breast cancer patients from normal persons. The Mann-Whitney test shows that the amount of tongue fur ($p=0.024$), maximum covering area of tongue fur ($p=0.009$), thin tongue fur ($p=0.009$), the average area of red dot ($p=0.049$), the maximum area of red dot ($p=0.009$), red dot in the spleen-stomach area ($p=0.000$), and red dot in the heart-lung area ($p=0.000$) demonstrate significant differences. Next, the data collected are further classified into two groups. The training group consists of 57 early-stage breast cancer patients and 60 normal persons, while the testing group is composed of 10 early-stage breast cancer patients and 10 normal persons. The logistic regression by utilizing these 7 tongue features with significant differences in Mann-Whitney test as factors is performed. In order to reduce the number of tongue features employed in prediction, we remove one of the 7 tongue features with lesser significant difference ($p>0.05$) and perform logistic regression twice. In the first time, we remove the maximum area of red dot ($p=0.266$), and perform logistic regression. Among them, the amount of tongue fur ($p=0.000$), the maximum covering area of tongue fur ($p=0.000$), thin tongue fur ($p=0.008$), the average area of red dot ($p=0.056$), red dot in the spleen-stomach area ($p=0.005$), red dot in the heart-lung area ($p=0.011$) reveal independently significant meaning. In the second round, the average area of red dot ($p=0.056$) is removed. The logistic regression shows that the amount of tongue fur ($p=0.001$), the maximum covering area of tongue fur ($p=0.000$), thin tongue fur ($p=0.007$), red dot in the spleen-stomach area ($p=0.006$), red dot in the heart-lung area ($p=0.003$) reveal independently significant meaning. The tongue features of the testing group are employed in the aforementioned three models to test the power of significant tongue features identified in predicting early-stage breast cancer. An accuracy of 80%, 80% and 90% is reached on normal peoples by applying the 7, 6 and 5 significant tongue features obtained through Mann-Whitney test, respectively, while 60%, 60% and 50% is reached on the corresponding early-stage breast cancer patients.**

Keywords—breast cancer; Automatic Tongue diagnosis System (ATDS); Mann-Whitney test; logistic regression

I. INTRODUCTION

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. BC that started off in the lobules is known as lobular carcinoma, while one that developed from the ducts is called ductal carcinoma [7]. Mammogram, ultrasound, MRI, blood testing, and fine needle aspiration biopsy are usually applied to discriminate BC patients from normal persons [8, 9]. However, other than being invasive or radioactive, some disadvantages are associated with these traditional BC diagnoses, e.g., false negative of mammogram, un-detectability of breast calcifications with ultrasound, unclear image caused by 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. Even though the detection of breast cancer in early stage is utmost important, yet in light of the above pitfalls, the detection of early-stage breast cancer is very challenging.

Nicolao et al. (2010) reported that the TCM practices, such as acupuncture and phytotherapy, are considered as the most popular disciplines requested by both medical experts and students in Switzerland [10]. It has 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, i.e., observation, smelling/listening, inquiry, and palpation. Observation tops the four ways of TCM diagnosis and tongue corresponds to the major subject of focus during observation [11]. Therefore, tongue diagnosis plays an important role in TCM [12].

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

978-1-4799-4013-4/14 $31.00 © 2014 IEEE
DOI 10.1109/ICMB.2014.9

progression of disease are all reflected on the tongue [2, 3]. Organ conditions, properties and variations of pathogens can be revealed 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 have conducted on the issue regarding consistency of TCM diagnosis [13-23] as well as herbal prescription or treatment [17, 18, 20], indicating 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 [11] and manual for guiding treatments [24]. Recently, intra- and inter-observer agreements of the automatic tongue diagnosis system (ATDS) and TCM practitioners have been conducted in our laboratory [23]. 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 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 [23]. 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 early stage BC patients and normal persons via non-invaded methods, expecting to detect BC in the early stage and give treatment in time to increase the recovery rate and lower relapse rate. Nine tongue features, namely, tongue color, tongue quality, tongue fissure, tongue fur, red dot, ecchymosis, tooth mark, saliva, and tongue shape are extracted for breast cancer patients and normal persons by the Automatic Tongue Diagnosis System (ATDS) [4-6, 28-31]. Features identified are further sub-divided according to the areas located, i.e., spleen-stomach, liver-gall-left, liver-gall-right, kidney, and heart-lung area, as shown in Fig. 1. The Mann-Whitney test is performed based on the data collected to induce significant tongue features ($p<0.05$) to discriminate early-stage breast cancer patients from normal persons [25, 26]. Ten tongue features with significant differences, identified by the Mann-Whitney test, are further employed as factors in the logistic regression to derive features with independently significant meaning. Three logistic regression models are constructed to predict the probability of infecting early-stage breast cancer patients. An accuracy of 80% of normal persons and 60% of early-stage breast cancer patients, accuracy of 80% of normal persons and 60% of early-stage breast cancer patients, and accuracy of 90% of normal persons and 50% of early-stage breast cancer patients is achieved, respectively. To the best of our knowledge, this is the first attempt in applying TCM tongue diagnosis to the discrimination of early-stage breast cancer patients and normal persons.

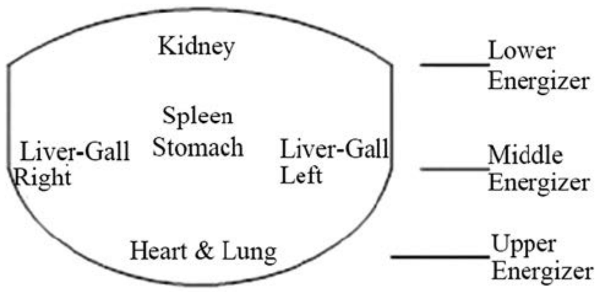


Fig. 1. The tongue is sub-divided into areas corresponding to different internal organs.

II. Materials and Methods

A. Participants

Two groups of the tongue images, i.e., experimental and control groups, are collected through the outpatients of Department of Traditional Chinese Medicine of Changhua Christian Hospital (CCH) in Taiwan (IRB no.: 120512). The tongue features for 67 early-stage breast cancer patients (include stage 0 and stage 1) in the experimental group and 70 normal persons in control group are extracted by the Automatic Tongue Diagnosis System (ATDS), respectively. The inclusion and exclusion criteria for the subjects of the experimental group are as follow:

- *Inclusion Criteria*
  - Diagnosed as breast cancer patient (ICD-9-CM 174-174.9) by a specialist.
  - Both male and female are included.
  - Patients with signed IRB agreement.
- *Exclusion Criteria*
  - Patients with no IRB agreement signed.
  - Pregnant woman.
  - Patients with acute infection.
  - Cognitive Impaired, e.g., imbecility dementia or delirium, caused by cancer metastasis to brain.
  - Patients unable to protrude the tongue stably.

Fig. 2(a) and (b) shows the exemplary tongue images of early stage breast cancer patients and normal persons recruited, respectively.

B. Automatic Tongue Diagnosis System (ATDS)

As shown in Fig. 3(a), the ATDS was developed to capture tongue images and extract features reliably to assist the

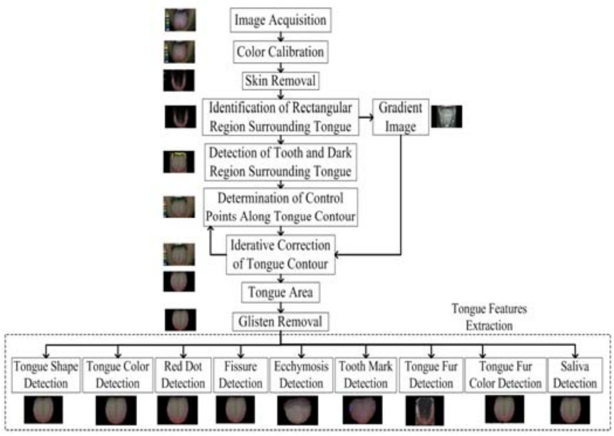
(a)

(b)

Fig. 2.(a)Tongue images of early stage breast cancer patients (ESBCP). (b) Tongue images of normal persons (NP).

diagnosis of TCM practitioners [23]. Fig. 3(b) demonstrates

(a)



(b)

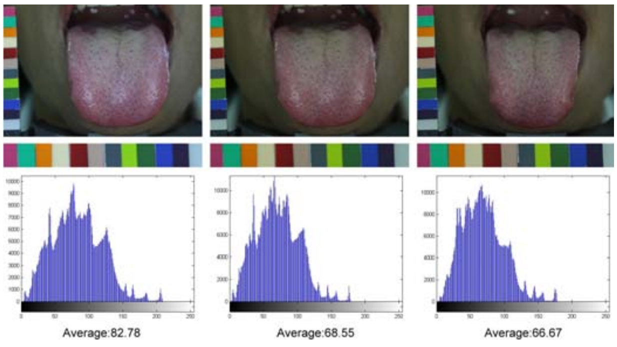
Fig. 3. (a) Components of the ATDS. (b)Illustrations of the processing steps of ATDS analysis.

the steps in three major functions, i.e., image capturing and color calibration, tongues area segmentation, and tongue feature extraction, included in the ATDS [5,23].

Variations in background lighting may change the color and brightness of the acquired images, greatly affecting consistency and stability of the extracted tongue features. The consistency and stability of tongue images captured and features extracted are achieved by calibrating brightness and color to compensate variations in intensity and color temperature of the light source and 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 taken at different circumstances. Fig. 4(a) and (b) display the images taken at T_1 before and after color calibration, respectively, whereas (c) 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 compensate the color deviation of the original image (Fig. 4(a) with a mean gray level 82.78) to allow colors in images taken at different time intervals consistent with each other (Fig. 4(b) and (c), with a mean gray level 68.55 and 66.67, respectively). Tongue images are 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 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 are extracted to generate detailed information regarding length, area, moisture, and number of fissures, tooth marks, and red dots to be employed in tongue diagnosis, as depicted in Fig. 5[21]. Features identified are further sub-divided according to the areas located, i.e., spleen-stomach, liver-gall-left, liver-gall-right, kidney, and heart-lung area, as shown in Fig. 1.

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

1. tongue color: including slightly white, slightly red, red, dark red, and dark purple.
2. tongue shape: including shape (thin and small, moderate, fat and large) and tongue body (normal, tilted to the left, tilted to the right).
3. Saliva: including total area and the amount of saliva (none, little, normal, excessive).
4. tongue fur: including 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: include amount, average covering area, shortest length, longest length.
7. Ecchymosis: amount, average coving 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: include number, average covering area,



(a) (b) (c)

Fig. 4. Calibration of image color using the color bar accompanied with the subject to make image quality consistent for images taken at different circumstances.

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: include 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).

C. *Statistical Analysis*

The tongue features of subjects participate are extracted by ATDS. The Mann-Whitney test is performed on the data sets acquired in the experimental and control groups to identify 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 predicting 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 $f(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k$ and $\beta_i, i \in Z, i \geq 0$ are the corresponding coefficients. The statistical analysis of this study is two folded: first, to conduct Mann-Whitney test, with respect to the tongue features, for the BC group and the normal group to select significant variables into the following logistic regression; 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 breast cancer.

III. Experimental Results

The tongue features for 67 of early-stage breast cancer patients (include 20 of stage 0 breast cancer patients and 47 of stage 1 breast cancer patients) and 70 normal persons are

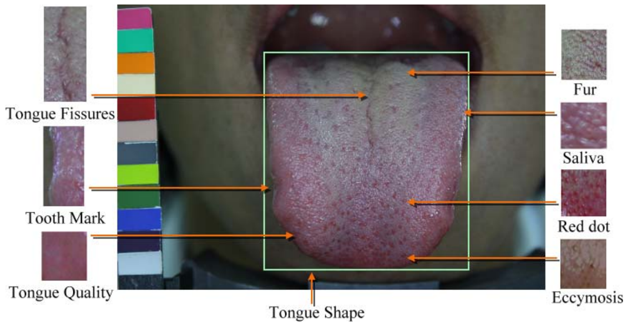


Fig. 5. Illustration of nine tongue features extracted by ATDS for tongue diagnosis.

extracted by the Automatic Tongue Diagnosis System (ATDS) [4-6]. The purpose focuses on inducing significant tongue features ($p<0.05$) to discriminate breast cancer patients from normal persons. The results obtained by applying Mann-Whitney test are listed in Table 1, with ten features, namely, the amount of tongue fur ($p=0.024$), maximum covering area of tongue fur ($p=0.009$), thin tongue fur ($p=0.009$), the average area of red dot ($p=0.049$), the maximum area of red dot ($p=0.009$), red dot in the spleen-stomach area ($p=0.000$) and red dot in the heart-lung area ($p=0.000$) demonstrating significant differences.

Next, the data collected are classified into two groups. The training group consists of 57 early-stage breast cancer patients and 60 normal persons, while the testing group is composed of 10 early-stage breast cancer patients and 10 normal persons. The logistic regression by utilizing these 7 tongue features with significant differences in Mann-Whitney test as factors is performed. Table 2 lists the result obtained.

A. *Model I*

The tongue features of the testing group are employed in two logistic regression models to test the power of significant tongue features identified in predicting breast cancer. In deriving the first model, the seven tongue features with significant difference, identified by Mann-Whitney test, are employed as factors $x_i, 1 \leq i \leq 7$, in deciding the subject with or without breast cancer. Let *p* represent the predicted

TABLE 1. Significant tongue features identified by applying Mann-Whitney to the data sets acquired from the group of normal persons (NP) and the group of early-stage breast cancer patients (ESBCP)

Significant tongue features	Group	sample size	Median	P-Value
The amount of tongue fur	NP	70	43.5	0.024
	ESBCP	67	52.00	
Maximum covering area of tongue fur	NP	70	10.62	0.009
	ESBCP	67	18.6	
Thin tongue fur	NP	70	68.0	0.009
	ESBCP	67	53.0	
The average area of red dot	NP	70	0.001	0.049
	ESBCP	67	0.001	
The maximum area of red dot	NP	70	0.003	0.009
	ESBCP	67	0.002	
Red dot in the spleen-stomach area	NP	70	52.0	0.000
	ESBCP	67	17.0	
Red dot in the heart-lung area	NP	70	60.0	0.000
	ESBCP	67	26.5	

TABLE 2.The results of applying the logistic regression by utilizing 7 tongue features with significant differences identified in Mann-Whitney test as factors

Predictor	Coeff.	SE Coeff	Z	P	Odds Ratio	Odds Ratio 95% CI	
						Lower	Upper
Constant	4.429	1.448	3.06	0.002			
The amount of tongue fur	0.08761	0.02500	-3.51	0.000	0.92	0.87	0.96
Maximum covering area of tongue fur	0.16424	0.04245	3.87	0.000	1.18	1.08	1.28
Thin tongue fur	0.02624	0.01025	-2.56	0.010	0.97	0.95	0.99
The average area of red dot	-1556.6	724.4	-2.15	0.032	0.00	0.00	0.00
The maximum area of red dot	474.4	426.1	1.11	0.266	0.00	0.00	0.00
Red dot in the spleen-stomach area	0.016158	0.005600	-2.89	0.004	0.98	0.97	0.99
Red dot in the heart-lung area	0.019463	0.007221	-2.70	0.007	0.98	0.97	0.99

TABLE 3. The probability of infecting breast cancer by employing Model I to the tongue features of 10 early-stage breast cancer patients (ESBCP) and 10 normal persons (NP) in the testing group

Group	the amount of tongue fur	the maximum covering area of tongue fur	thin tongue fur	the average area of red dot	the maximum area of red dot	red dot in the spleen-stomach area	red dot in the heart-lung area.	Model I	Probability of infecting breast cancer
ESBCP1	71	23.73	19	0	0.002	0	43	1.41	0.80
ESBCP2	97	57.74	79	0	0	0	0	3.58	0.97
ESBCP3	44	15.78	44	0	0.001	0	11	2.41	0.92
ESBCP4	17	5.48	30	0.001	0.003	0	51	1.88	0.87
ESBCP5	64	20.74	62	0	0.003	0	39	0.46	0.61
ESBCP6	81	20.66	43	0.001	0.003	73	52	-2.87	0.05
ESBCP7	48	12.14	9	0	0.003	185	156	-2.59	0.07
ESBCP8	8	1.45	98	0	0.002	17	33	1.30	0.79
ESBCP9	57	7.83	25	0.001	0.003	86	0	-1.63	0.16
ESBCP10	25	1.72	14	0.001	0.004	260	4	-2.01	0.12
NP 1	51	35.65	73	0.001	0.003	59	159	-0.28	0.43
NP2	15	0.44	15	0.001	0.004	409	27	-3.86	0.03
NP3	1	0	90	0.001	0.003	286	43	-3.17	0.04
NP4	14	3.18	51	0	0.03	227	195	-3.33	0.03
NP5	3	0	72	0.001	0.003	280	0	-2.09	0.11
NP6	50	10.91	70	0	0.001	15	6	0.25	0.56
NP7	39	11.83	65	0.001	0.003	114	60	-1.82	0.14
NP8	30	10.5	48	0.001	0.003	21	55	0.66	0.66
NP9	29	8.1	88	0.001	0.003	101	56	-1.88	0.13
NP10	23	9.26	41	0.001	0.004	183	238	-3.99	0.02

probability of infecting breast cancer using logistic regression. The coefficients in (1) can be determined as:

$$f(x) = 4.429 - 0.08761x_1 + 0.16424x_2 - 0.02624x_3 - 1556.6x_4 + 474.4x_5 - 0.016158x_6 - 0.019463x_7 \quad (2)$$

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

The accuracy of Model I in predicting breast cancer can be tested by substituting the values of tongue features extracted from the testing group into (2). According to (1), the probability of infecting breast cancer *p* is etermined once the value of $f(x)$ is given. A predicted probability larger than 0.5 corresponds to tongue features of a early-stage breast cancer patient, while that smaller than 0.5 represents those of a normal person. Table 3 lists the probability of infecting breast cancer for 10 early-stage breast cancer patients (ESBCP) and 10 normal persons (NP) in the testing group according to Model I. Among them, the probability of infection breast cancer is larger than 0.5 for ESBCPs 1, 2, 3, 4, 5, 8 and smaller than 0.5 for NPs 1, 2, 3, 4, 5, 7, 9, and 10. Correct diagnoses are reached for these 14 cases out of a total of 20 ones. An accuracy of 80% is achieved for normal persons and 60% accuracy is achieved for early-stage breast cancer patients by employing Model I in predicting breast cancer through tongue diagnosis.

B. Model II

Table 2 lists the results by employing logistic regression to analyze the relationship between dependent and independent variables, under the hypothesis of infecting breast cancer or not. We remove one of the 7 tongue features which is not the most significant differences (p>0.05) and perform logistic regression. Table 4 lists the result obtained. Among them, the amount of tongue fur (p=0.000), maximum covering area of tongue fur (p=0.000), thin tongue fur (p=0.008), the average area of red dot (p=0.056), red dot in the spleen-stomach area (p=0.005), and red dot in the heart-lung area (p=0.011) exhibit significant difference. Let *p* represent the probability of infecting breast cancer. Consider the significant variables in the above logistic regression, we consider the following model:

$$f(x) = 5.212 - 0.09014x_1 + 0.16016x_2 - 0.02701x_3 - 1044.2x_4 - 0.014443x_5 - 0.016043x_6 \quad (3)$$

TABLE 4. The first results of removing the most insignificant differences of 7 tongue features in Table 3 and applying the logistic regression

Predictor	Coeff.	SE Coeff	Z	P	Odds Ratio	Odds Ratio 95% CI	
						Lower	Upper
Constant	5.212	1.290	4.04	0.000			
the amount of tongue fur	-0.09014	0.02479	-3.64	0.000	0.91	0.87	0.96
maximum covering area of tongue fur	0.16016	0.04184	3.83	0.000	1.17	1.08	1.27
thin tongue fur	-0.02701	0.01014	-2.66	0.008	0.97	0.95	0.99
the average area of red dot	-1044.2	546.2	-1.91	0.056	0.00	0.00	2.67E+11
red dot in the spleen-stomach area	-0.014443	0.005123	-2.82	0.005	0.99	0.98	1.00
red dot in the heart-lung area	-0.016043	0.006340	-2.53	0.011	0.98	0.97	1.00

TABLE 5. The probability of infecting breast cancer by employing Model II to the tongue features of 10 early-stage breast cancer patients (ESBCP) and 10 normal persons (NP) in the testing group

Group	the amount of tongue fur	the maximum covering area of tongue fur	thin tongue fur	the average area of red dot	red dot in the spleen-stomach area	red dot in the heart-lung area.	Model II	Probability of infecting breast cancer
ESBCP1	71	23.73	19	0	0	43	1.41	0.80
ESBCP2	97	57.74	79	0	0	0	3.58	0.97
ESBCP3	44	15.78	44	0	0	11	2.41	0.92
ESBCP4	17	5.48	30	0.001	0	51	1.88	0.87
ESBCP5	64	20.74	62	0	0	39	0.46	0.61
ESBCP6	81	20.66	43	0.001	73	52	-2.87	0.05
ESBCP7	48	12.14	9	0	185	156	-2.59	0.07
ESBCP8	8	1.45	98	0	17	33	1.30	0.79
ESBCP9	57	7.83	25	0.001	86	0	-1.63	0.16
ESBCP10	25	1.72	14	0.001	260	4	-2.01	0.12
NP 1	51	35.65	73	0.001	59	159	-0.09	0.48
NP2	15	0.44	15	0.001	409	27	-3.86	0.02
NP3	1	0	90	0.001	286	43	-3.17	0.04
NP4	14	3.18	51	0	227	195	-3.33	0.03
NP5	3	0	72	0.001	280	0	-2.09	0.11
NP6	50	10.91	70	0	15	6	0.25	0.56
NP7	39	11.83	65	0.001	114	60	-1.82	0.14
NP8	30	10.5	48	0.001	21	55	0.66	0.66
NP9	29	8.1	88	0.001	101	56	-1.88	0.13
NP10	23	9.26	41	0.001	183	238	-3.99	0.02

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

The accuracy of using Model II in predicting breast cancer can be tested by substituting the values of tongue features extracted from the testing group into (3). According to (1), the probability of infecting early-stage breast cancer patients p is determined once the value of $f(x)$ is given. A predicted probability larger than 0.5 corresponds to tongue features of a breast cancer patient, while that smaller than 0.5 represents those of a normal person. Table 5 lists the probability of infecting breast cancer for 10 early-stage breast cancer patients (ESBCP) and 10 normal persons (NP) in the testing group according to Model II. Among them, the probability of infection breast cancer is larger than 0.5 for ESBCPs 1, 2, 3, 4, 5 and 8 and smaller than 0.5 for NPs 1, 2, 3, 4, 5, 7, 9, and 10. Correct diagnoses are reached for these 14 cases out of a total of 20 ones. An accuracy of 80% is achieved on normal persons and 60% accuracy is achieved on early-stage breast cancer patients by employing Model II in predicting breast cancer through tongue diagnosis.

C. Model III

Table 4 lists the results by employing logistic regression to analyze the relationship between dependent and independent variables, under the hypothesis of infecting breast cancer or not. We remove one of the tongue features which is not the most significant differences (p>0.05), the average area of red dot (p=0.056) of 6 tongue features and perform logistic regression. Table 6 lists the final result obtained. Among them, the amount of tongue fur (p=0.001), the maximum covering area of tongue fur (p=0.000), thin tongue fur (p=0.007), red dot in the spleen-stomach area (p=0.006), red dot in the heart-lung area (p=0.003) exhibit significant difference (p=0.05). Let p represent the probability of infecting breast cancer. Consider the significant variables in the above logistic regression, we consider the following model:

$$f(x) = 4.194 - 0.07131x_1 + 0.13398x_2 - 0.02740x_3 - 0.012447x_4 - 0.018591x_5 \quad (4)$$

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

The accuracy of using Model III in predicting breast cancer can be tested by substituting the values of tongue features extracted from the testing group into (4). According to (1), the probability of infecting early-stage breast cancer

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

Predictor	Coeff.	SE Coeff	Z	P	Odds Ratio	Odds Ratio 95% CI	
						Lower	Upper
Constant	4.194	1.117	3.75	0.000			
the amount of tongue fur	-0.07131	0.02135	-3.34	0.001	0.93	0.89	0.97
maximum covering area of tongue fur	0.13398	0.03684	3.643	0.000	1.14	1.06	1.23
thin tongue fur	-0.02740	0.01008	-2.72	0.007	0.97	0.95	0.99
red dot in the spleen-stomach area	-0.012447	0.004501	-2.77	0.006	0.99	0.98	1.00
red dot in the heart-lung area	-0.018591	0.006300	-2.95	0.003	0.98	0.97	0.99

TABLE 7. The probability of infecting breast cancer by employing Model III to the tongue features of 10 early-stage breast cancer patients (ESBCP) and 10 normal persons (NP) in the testing group

Group	the amount of tongue fur	the maximum covering area of tongue fur	thin tongue fur	red dot in the spleen-stomach area	red dot in the heart-lung area.	Model III	Probability of infecting breast cancer
ESBCP1	71	23.73	19	0	43	0.99	0.73
ESBCP2	97	57.74	79	0	0	2.85	0.95
ESBCP3	44	15.78	44	0	11	1.76	0.85
ESBCP4	17	5.48	30	0	51	1.95	0.87
ESBCP5	64	20.74	62	0	39	-0.01	0.49
ESBCP6	81	20.66	43	73	52	-1.87	0.13
ESBCP7	48	12.14	9	185	156	-3.05	0.05
ESBCP8	8	1.45	98	17	33	0.31	0.58
ESBCP9	57	7.83	25	86	0	-0.58	0.36
ESBCP10	25	1.72	14	260	4	-1.05	0.26
NP 1	51	35.65	73	59	159	-0.36	0.41
NP2	15	0.44	15	409	27	-2.82	0.06
NP3	1	0	90	286	43	-2.70	0.06
NP4	14	3.18	51	227	195	-4.23	0.01
NP5	3	0	72	280	0	-1.48	0.19
NP6	50	10.91	70	15	6	-0.13	0.47
NP7	39	11.83	65	114	60	-1.32	0.21
NP8	30	10.5	48	21	55	0.86	0.70
NP9	29	8.1	88	101	56	-1.50	0.18
NP10	23	9.26	41	183	238	-4.03	0.02

patients p is determined once the value of $f(x)$ is given. A predicted probability larger than 0.5 corresponds to tongue features of a breast cancer patient, while that smaller than 0.5 represents those of a normal person. Table 7 lists the probability of infecting breast cancer for 10 early-stage breast cancer patients (ESBCP) and 10 normal persons (NP) in the testing group according to Model III. Among them, the probability of infection breast cancer is larger than 0.5 for ESBCPs 1, 2, 3, 4, and 8 and smaller than 0.5 for NPs 1, 2, 3, 4, 5, 6, 7, 9, and 10. Correct diagnoses are reached for these 14 cases out of a total of 20 ones. An accuracy of 90% is achieved on normal persons and 50% accuracy is achieved on early-stage breast cancer patients by employing Model III in predicting breast cancer through tongue diagnosis.

IV. Conclusions

This paper investigates discriminating tongue features to distinguish between BC patients of earlier stage and normal people via non-invaded methods, expecting to find BC earlier and give treatment in time to rise the recovery rate and lower relapse rate. The Mann-Whitney test shows that the amount of tongue fur (p=0.018), maximum covering area of tongue fur (p=0.003), thin tongue fur (p=0.007), the average area of red dot (p=0.000), the maximum area of red dot (p=0.024), red dot in the spleen-stomach area (p=0.000), and red dot in the heart-lung area (p=0.000) demonstrate significant differences. Next, the logistic regression by utilizing these 7 tongue features with significant differences in Mann-Whitney test as factors is performed. We remove one of the 7 tongue features which is not the most significant differences (p>0.05) and perform logistic regression two times. In the first time, we remove the maximum area of red dot (p=0.266), and perform logistic regression. Among them, the amount of tongue fur (p=0.000), the maximum covering area of tongue fur (p=0.000), thin tongue fur (p=0.008), the average area of red dot (p=0.056), red dot in the spleen-stomach area (p=0.005), red dot in the heart-lung area (p=0.011) reveal independently significant meaning. In the second time, we remove the average area of red dot (p=0.056), and perform logistic regression. Among them, the amount of tongue fur (p=0.001), the maximum covering area of tongue fur (p=0.000), thin tongue fur (p=0.007), red dot in the spleen-stomach area (p=0.006), red dot in the heart-lung area (p=0.003) reveal independently significant meaning. The tongue features of the testing group are employed in three logistic regression models to test the predicting accuracy of significant tongue features identified in predicting breast cancer. An accuracy of 80% is reached on normal persons through the first logistic regression model by applying the 7 significant tongue features obtained through Mann-Whitney test and 60% is reached on early-stage breast cancer patients. For the second model employing 6 tongue features induced by logistic regression with independently significant meaning, 80% accuracy is achieved for normal persons and 60% accuracy is achieved for early-stage breast cancer patients. The third model employing 5 tongue features induced by logistic regression with independently significant meaning, 90% accuracy is achieved for normal persons and 50% accuracy is achieved for early-stage breast cancer patients. The TCM tongue diagnosis can serve as a preliminary screening procedure in early detection of breast cancer 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 early-stage breast cancer patients and normal persons.

REFERENCES

[1] Margaret McCartney, "How useful are lifetime risks of disease?" *British Medical Journal*, Vol. 342, 2011.

[2] N. Wiseman and A. Ellis, "Fundamentals of Chinese Medicine: Zhong Yi XueJi Chu," Brookline, MA: Paradigm Publications, 1996.

[3] B. Kirschbaum, "Atlas of Chinese Tongue Diagnosis," Seattle: Eastland Press, 2000.

[4] Lun-Chien Lo, Chia-Yun Chen, John Y. Chiang, Tsung-Lin Cheng, Hong-Jen Lin and Hen-Hong Chang, "Tongue Diagnosis of Traditional Chinese Medicine for Rheumatoid Arthritis," *Afr. J. Tradit. Complement Altern. Med.*, Vol. 10, Issue 5, pp. 360-369, 2013.

[5] Yu Cheng Hsu, Ying-ching Chen, Lunchien Lo, John Y. Chiang, "Automatic Tongue Feature Extraction," *Proc. Int'l Computer Symp. (ICS2010)*, pp. 936-941, Tainan, Taiwan, Dec. 16-18 2010. [EI]

[6] Lun-chien Lo, Mark Chun-cheng Hou, Ying-ling Chen, John Y. Chiang and Jeng-jiun Shyu, "Automatic Tongue Diagnosis System," *Proc. 2nd Int'l Conf. on BioMedical Engineering and Informatics (BMEI'09)*, pp. 140-144, Tianjin, China, 17-19 October, 2009.

[7] Wasif N, Maggard MA, Ko CY, Giuliano AE., "Invasive lobular vs. ductal breast cancer: a stage-matched comparison of outcomes." Ann SurgOncol., Vol. 17, pp. 1862–1869, 2010.

[8] Katrina Armstrong, Elizabeth A. Handorf, Jinbo Chen, Mirar N. Bristol Demeter, "Breast Cancer Risk Prediction and Mammography Biopsy Decisions : A Model-Based Study." *American Journal of Preventive Medicine*, Vol. 44, pp. 15–22, 2013.

[9] Elizabeth Meade, Maura Dowling, "Early breast cancer: diagnosis, treatment and survivorship." *British Journal of Nursing*, Vol. 21, pp. S4 - S7, 2012.

[10] M. Nicolao, M. G. Täuber& P. Heusser, "How should complementary and alternative medicine be taught to medical students in Switzerland? A survey of medical experts and students," *Medical Teacher*, 32: 50–55, 2010.

[11] G. G. Zhang, W. Lee, B, Bausell, L. Lao, B. Handwerger and B. Berman, "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.*, 11: 415-421, 2005.

[12] T. Kaptchuk, "Chinese Medicine: The Web That Has No Weaver," 2nd ed., London: Rider, 2000.

[13] K. A. O'Brien, E. Abbas, J. Zhang, Z. X. Guo, R. Luo, A. Bensoussan, and P. A. Komesaroff, "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.*, vol. 15, no. 5, pp. 259-266, 2009.

[14] K. A. O'Brien, E. Abbas, P. Movsessian, M. Hook, P. A. Komesaroff , and S. Birch, "Investigating the Reliability of Japanese Toyohari Meridian Therapy Diagnosis," *J. Altern. Complement. Med.,* vol. 15, no. 10, pp. 1099-1105, 2009.

[15] S. Mist, C. Ritenbaugh and M. Aickin, "Effects of Questionnaire-Based Diagnosis and Training on Inter-Rater Reliability among Practitioners of Traditional Chinese Medicine," *J. Altern. Complement. Med.,* vol. 15, no. 7, pp. 703-709, 2009.

[16] K. A. O'Brien and S. Birch, "A review of the reliability of Traditional East Asian Medicine diagnoses," *J. Altern. Complement. Med.,* vol. 15, no. 4, pp. 353-366, 2009.

[17] R. R. Coeytaux, W. Chen, C. E. Lindemuth, Y. Tan, and A. C. Reilly, "Variability in the diagnosis and point selection for persons with frequent headache by traditional Chinese acupuncture," *J. Altern. Complement. Med.*,vol. 12, no. 9, pp. 863-872, 2006.

[18] J. J. Sung, W. K. Leung, J. Y. Ching et al., "Agreements among traditional Chinese medicine practitioners in the diagnosis and treatment of irritable bowel syndrome," *J. Aliment. Pharmacol. Ther*., vol. 20, no. 10, pp. 1205-1210, 2004.

[19] B. Zell, J. Hirata, A. Marcus et al., "Diagnosis of symptomatic postmenopausal women by traditional Chinese medicine practitioners," *J. Menopause*, vol. 7, no. 2, pp. 129-134, 2000.

[20] G. G. Zhang, B. Bausell, L. Lao et al., "The variability of TCM pattern diagnosis and herbal prescription on rheumatoid arthritis patients," *J. Altern. Ther. Health Med.*, vol. 10, no. 1, pp. 58-63, 2004.

[21] G. G. Grant, B. Singh, W. Lee, H. Barry, L. Lao, and B. Brian, "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.*, vol. 14, no. 4, pp. 381-386, 2008.

[22] K. Minah, D. Cobbin, and C. Zaslawski, "Traditional Chinese Medicine tongue inspection: An examination of inter- and intra-practitioner reliability for specific tongue characteristics," *J. Altern. Complement. Med.*, vol. 14, no. 5, pp. 527-536, 2008.

[23] L. C. Lo, Y. F. Chen, W. J. Chen, T. L. Cheng, and J. Y. Chiang, "The Study on the Agreement between Automatic Tongue Diagnosis System and Traditional Chinese Medicine Practitioners," *Evidence-Based Complementary and Alternative Medicine*, vol. 2012, Article ID 505063, 9 pages.

[24] R. N. Schnyer, J. J. Allen, "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.*, 8:623-634, 2002.

[25] Michael P. Fay, Michael A. Proschan, "Wilcoxon-Mann-Whitney or t-test? On assumptions for hypothesis tests and multiple interpretations of decision rules." *PMC*, Vol. 4, pp. 1-39, 2010.

[26] Rebecca Siegel MPH, DeepaNaishadham MA, MS, AhmedinJemal DVM, PhD, "Cancer statistics, 2012," *A Cancer Journal for Clinicians*, Vol. 62, Iss. 1, pp. 10–29, January/February 2012.

[27] Hans van der Heijden,"Decision support for selecting optimal logistic regression models," *Expert Systems with Applications*, Vol. 39, Iss. 10, pp. 8573-8583, 2012.

[28] Lun-Chien Lo, Chia-Yun Chen, John Y. Chiang, Tsung-Lin Cheng, Hong-Jen Lin and Hen-Hong Chang, "Tongue Diagnosis of Traditional Chinese Medicine for Rheumatoid Arthritis," *Afr. J. Tradit. Complement Altern*. Med., Vol. 10, Issue 5, pp. 360-369, 2013.

[29] Lun-chien Lo, Tsung-Lin Cheng, John Y. Chiang and Natsagdorj Damdinsuren, "Breast Cancer Index - A Perspective on Tongue Diagnosis in Traditional Chinese Medicine," Journal of Traditional and Complementary Medicine, Vol. 3, No. 3, pp. 194-203, 2013.

[30] Lun-Chien Lo, Tsung-Lin Cheng, Wen-Jiuan Chen, Yung-Fu Chen and John Y. Chiang, "The Study on the Agreement between Automatic Tongue Diagnosis System and Traditional Chinese Medicine Practitioners," Evidence-Based Complementary and Alternative Medicine, Vol. 2012.

[31] Lun-Chien Lo, John Y. Chiang, Tsung-Lin Cheng and Pei-Shuan Shieh, "Visual Agreement Analyses of Traditional Chinese Medicine -A Multiple Dimensional Scaling Approach," Evidence-Based Complementary and Alternative Medicine, Vol. 2012.