TCM tongue diagnosis index of early-stage breast cancer

Lun-chien Lo a, b, Tsung-Lin Cheng b, Yi-Jing Chen c, Sainbuyan Natsagdorj d, John Y. Chiang c,e,*

a Department of Traditional Chinese Medicine, Changhua Christian Hospital, Changhua, Taiwan
b Graduate Institute of Statistics and Information Science, National Changhua University of Education, Changhua, Taiwan
c Department of Computer Science and Engineering, National Sun Yat-sen University, Kaohsiung, Taiwan
d Department of Mechanical Engineering, Cheng Shiu University, Kaohsiung, Taiwan
e Department of Healthcare Administration and Medical Informatics, Kaohsiung Medical University. 80708 Kaohsiung, Taiwan

A R T I C L E   I N F O
Article history:
Received 12 September 2014
Received in revised form 25 February 2015
Accepted 5 July 2015
Available online 10 July 2015

Keywords:
Breast cancer
Automatic tongue diagnosis System (ATDS)
Mann–Whitney test
Logistic regression

A B S T R A C T
Objectives: This paper investigates discriminating tongue features to distinguish between early stage breast cancer (BC) patients and non-breast cancer individuals through non-invasive traditional Chinese medicine (TCM) tongue diagnosis.

Design: The tongue features for 67 patients with 0 and 1 stages of BC, and 70 non-breast cancer individuals are extracted by the automatic tongue diagnosis system (ATDS). A total of nine tongue features, namely, tongue color, tongue quality, tongue fissure, tongue fur, red dot, ecchymosis, tongue 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 areas. This study focuses on deriving significant tongue features (p < 0.05) to discriminate early-stage BC patients from non-breast cancer individuals.

Results: 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. The data collected are further classified into two groups. The training group consists of 57 early-stage BC patients and 60 non-breast cancer individuals, while the testing group is composed of 10 early-stage BC patients and 10 non-breast cancer individuals. 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, tongue features with the least amount of significant difference, namely, maximum area of red dot and average area of red dot, are removed progressively. 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 BC. An accuracy of 80%, 80% and 90% is reached on non-breast cancer individuals 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 BC patients.

Conclusion: The TCM tongue diagnosis can serve as a preliminary screening procedure in the 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 early-stage BC patients and non-breast cancer individuals.

© 2015 Elsevier Ltd. All rights reserved.

1. 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. BC that started off in the lobules is known as lobular carcinoma, while one that developed from the ducts is called ductal carcinoma. 1 Mammogram, ultrasound, MRI, blood testing, and fine needle aspiration biopsy are usually applied to discriminate BC patients from non-breast cancer individuals. 2,3 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, blurred 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 BC in early stage is utmost important, yet in light of the above pitfalls, the detection of early-stage BC is very challenging.

Nicola et al. 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. 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. Therefore, tongue diagnosis plays an important role in TCM.

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. 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 been conducted on the issue regarding consistency of TCM diagnosis as well as herbal prescription or treatment, 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 and manual for guiding treatments. Recently, intra- and inter-observer agreements of the automatic tongue diagnosis system (ATDS) and TCM practitioners have been conducted in our laboratory. 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. 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 non-breast cancer individuals via non-invasive 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 both BC patients and non-breast cancer individuals by the ATDS. 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 derive significant tongue features (p < 0.05) to discriminate early-stage BC patients from non-breast cancer individuals.

Ten tongue features with significant differences, identified by the Mann–Whitney test, are further employed as factors in the logistic regression to induce features with independently significant meaning. Three logistic regression models are constructed to predict the probability of infecting early-stage BC patients. An accuracy of 80% of non-breast cancer individuals and 60% of early-stage BC patients, accuracy of 80% of non-breast cancer individuals and 60% of early-stage BC patients, and accuracy of 90% of non-breast cancer individuals and 50% of early-stage BC 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 BC patients and non-breast cancer individuals.

2. Materials and methods

2.1. Participants

The tongue images of two groups, 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 BC patients (include stages 0 and 1) in the experimental group and 70 non-breast cancer individuals in control group are extracted by the ATDS. The inclusion and exclusion criteria for the subjects of the experimental group are as follow:

- **Inclusion criteria**
  1. The current study comprised subjects aged >18 years who were diagnosed with primary breast cancer (ICD-9-CM code 174) between 2010 and 2011.
  2. We identified the early stage (stage 0 and 1) of disease of patients with breast cancer according to chart review and pathologic findings proved.

- **Exclusion criteria**
  1. Distant metastases were defined as metastases to the lungs, liver, brain, bones, and other organs (ICD-9-CM codes 197.x, 198.0, 198.1, and 198.3–198.7).
  2. Pregnant woman
  3. Patients with acute illness, e.g., infection, hemorrhage, unstable vital signs, and etc.
  4. Patients unable to protrude the tongue stably.

Participants were introduced of all interests and benefits as well as risks first, and signed a written consent form. Fig. 2(a) and (b)
show the exemplary tongue images of the early stage BC patients and non-breast cancer individuals recruited, respectively.

2.2. Automatic tongue diagnosis system (ATDS)

As shown in Fig. 3(a), the ATDS was developed to capture tongue images and automatically extract features reliably to assist the diagnosis of TCM practitioners. The chassis of ATDS is custom-made, while the remaining parts, e.g., camera, circular LED light source, etc., can be purchased off-the-shelf. The value of ATDS hinges on its ability to segment the tongue region and extract tongue features automatically. Fig. 3(b) demonstrates the processing steps in three major functions, i.e., image capturing and color calibration, tongue area segmentation, and tongue feature extraction, embedded in the ATDS.

Variations in background lighting may change the color and brightness of the acquired images, greatly affecting the consistency and stability of the tongue features extracted. 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 surrounding environment with a color bar placed next to 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 T1 before and after color calibration, respectively, whereas (c) demonstrates the image taken at T2 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 (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, namely, 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. 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 areas, 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 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: 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).
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).

2.3. 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 with no assumption 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. 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. The sample size is 67 and 70 for breast cancer and non-breast cancer group, respectively. Concerning the sample size, if one would like to conduct a hypothesis testing that requires a power of testing achieving 0.8 at the significance level \( \alpha = 0.05 \), the following formula is used

\[
0.8 = P \left( N(0, 1) > 1.65 - \frac{\sqrt{n} \times 0.3}{1} \right),
\]

\[
1.65 - \frac{\sqrt{n} \times 0.3}{1} = -0.8416
\]

\( \Rightarrow n \approx 69 \)

The above equation utilizes asymptotic normality for observations with/without normal assumption. Therefore, the sample size is 67 and 70 for breast cancer and non-breast cancer group, respectively, is appropriate.

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.\textsuperscript{37} 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.,

\[
p = \frac{e^{f(x)}}{1 + e^{f(x)}}
\]

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: firstly, to conduct Mann–Whitney test, with respect to the tongue features, for both the BC and non-breast cancer groups to identify significant variables into the following logistic regression; secondly, conducting logistic regression based on the selection of the variables from the aforementioned Mann–Whitney test to obtain a predicting equation for the probability of with/without BC.

3. Experimental results

The tongue features for 67 early-stage BC patients (include 20 of stage 0 BC patients and 47 of stage 1 BC patients) and 70 non-breast cancer individuals are extracted by the ATDS. The purpose focuses on inducing significant tongue features \((p < 0.05)\) to discriminate BC patients from non-breast cancer individuals. 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 BC patients and 60 non-breast cancer individuals, while the testing group is composed of 10 early-stage BC patients and 10 non-breast cancer individuals. 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.

3.1. 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 BC. 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 BC. Let \( p \) represent the predicted probability of infecting BC using logistic regression. The coefficients in Eq. (1) can be determined as:

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

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, \( x_5 \) 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 BC can be tested by substituting the values of tongue features extracted from the testing group into Eq. (2). According to Eq. (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 an early-stage BC patient, while that smaller than 0.5 represents those of a non-breast cancer individual. Table 3 lists the probability of infecting BC for 10 early-stage BC patients (ESBCP) and 10 non-breast cancer individuals (NBCI) in the testing group according to Model I. Among them, the probability of infecting BC is larger than 0.5 for ESBCPs 1–5, 8 and smaller than 0.5 for NBCIs 1–5, 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 non-breast cancer individuals and 60% accuracy is achieved for early-stage BC patients by employing Model I in predicting BC through tongue diagnosis.

3.2. 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 BC 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 BC. Consider the significant variables in the above logistic regression, we consider the following model:

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

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 BC can be tested by substituting the values of tongue features extracted from the testing group into Eq. (3). According to Eq. (1), the probability of infecting early-stage BC 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 BC patient, while that smaller than 0.5 represents those of a non-breast cancer individual. Table 5 lists the probability of infecting BC for 10 early-stage BC patients (ESBCP) and 10 non-breast cancer individuals (NBCI) in the testing group according to Model II. Among them, the probability of infection BC is larger than 0.5 for ESBCPs 1–5 and 8 and smaller than 0.5 for NBCIs 1–9, and 10. Correct diagnoses are reached for these 14 cases out of a total of 20 ones. An accuracy of 80% is achieved on non-breast cancer individuals and 60% accuracy is achieved on early-stage BC patients by employing Model II in predicting BC through tongue diagnosis.

3.3. 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 BC 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 BC. Consider the significant variables in the above logistic regression, we consider the following model:

\[
f(x) = 4.194 - 0.07131x_1 + 0.13398x_2 - 0.02740x_3 - 0.012447x_4 - 0.018591x_5
\]

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.
Table 1  
Significant tongue features identified by applying Mann–Whitney to the data sets acquired from the group of non-breast cancer individuals (NBCI) and the group of early-stage BC patients (ESBCP).

<table>
<thead>
<tr>
<th>Significant tongue features</th>
<th>Group</th>
<th>Sample size</th>
<th>Median</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>The amount of tongue fur</td>
<td>NBCI</td>
<td>70</td>
<td>43.5</td>
<td>0.024</td>
</tr>
<tr>
<td></td>
<td>ESBCP</td>
<td>67</td>
<td>52.00</td>
<td></td>
</tr>
<tr>
<td>Maximum covering area of tongue fur</td>
<td>NBCI</td>
<td>70</td>
<td>10.62</td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td>ESBCP</td>
<td>67</td>
<td>18.6</td>
<td></td>
</tr>
<tr>
<td>Thin tongue fur</td>
<td>NBCI</td>
<td>70</td>
<td>68.0</td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td>ESBCP</td>
<td>67</td>
<td>53.0</td>
<td></td>
</tr>
<tr>
<td>The average area of red dot</td>
<td>NBCI</td>
<td>70</td>
<td>0.001</td>
<td>0.049</td>
</tr>
<tr>
<td></td>
<td>ESBCP</td>
<td>67</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td>The maximum area of red dot</td>
<td>NBCI</td>
<td>70</td>
<td>0.003</td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td>ESBCP</td>
<td>67</td>
<td>0.002</td>
<td></td>
</tr>
<tr>
<td>Red dot in the spleen–stomach area</td>
<td>NBCI</td>
<td>70</td>
<td>52.0</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>ESBCP</td>
<td>67</td>
<td>17.0</td>
<td></td>
</tr>
<tr>
<td>Red dot in the heart–lung area</td>
<td>NBCI</td>
<td>70</td>
<td>60.0</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>ESBCP</td>
<td>67</td>
<td>26.5</td>
<td></td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Coeff.</th>
<th>SE Coeff</th>
<th>Z</th>
<th>P</th>
<th>Odds ratio</th>
<th>Odds ratio 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>4.429</td>
<td>1.448</td>
<td>3.06</td>
<td>0.002</td>
<td>0.92</td>
<td>0.97</td>
</tr>
<tr>
<td>The amount of tongue fur</td>
<td>0.08761</td>
<td>0.02500</td>
<td>-3.51</td>
<td>0.000</td>
<td>0.97</td>
<td>0.97</td>
</tr>
<tr>
<td>Maximum covering area of tongue fur</td>
<td>0.16424</td>
<td>0.04245</td>
<td>3.87</td>
<td>0.000</td>
<td>1.18</td>
<td>1.08</td>
</tr>
<tr>
<td>Thin tongue fur</td>
<td>0.02624</td>
<td>0.01025</td>
<td>-2.56</td>
<td>0.010</td>
<td>0.97</td>
<td>0.95</td>
</tr>
<tr>
<td>The average area of red dot</td>
<td>-1.5566</td>
<td>724.4</td>
<td>-2.15</td>
<td>0.032</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>The maximum area of red dot</td>
<td>474.4</td>
<td>426.1</td>
<td>1.11</td>
<td>0.266</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>Red dot in the spleen–stomach area</td>
<td>0.016158</td>
<td>0.005600</td>
<td>-2.89</td>
<td>0.004</td>
<td>0.99</td>
<td>0.97</td>
</tr>
<tr>
<td>Red dot in the heart–lung area</td>
<td>0.019463</td>
<td>0.007221</td>
<td>-2.70</td>
<td>0.007</td>
<td>0.99</td>
<td>0.99</td>
</tr>
</tbody>
</table>

Table 3  
The probability of infecting BC by employing Model I to the tongue features of 10 early-stage BC patients (ESBCP) and 10 non-breast cancer individuals (NBCI) in the testing group.

<table>
<thead>
<tr>
<th>Group</th>
<th>The amount of tongue fur</th>
<th>The maximum covering area of tongue fur</th>
<th>Thin tongue fur</th>
<th>The average area of red dot</th>
<th>The maximum area of red dot</th>
<th>Red dot in the spleen–stomach area</th>
<th>Red dot in the heart–lung area</th>
<th>Model1</th>
<th>Probability of infecting breast cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td>ESBCP1</td>
<td>71</td>
<td>23.73</td>
<td>19</td>
<td>0</td>
<td>0.002</td>
<td>0</td>
<td>43</td>
<td>1.41</td>
<td>0.80</td>
</tr>
<tr>
<td>ESBCP2</td>
<td>97</td>
<td>57.74</td>
<td>79</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3.58</td>
<td>0.97</td>
</tr>
<tr>
<td>ESBCP3</td>
<td>44</td>
<td>15.78</td>
<td>44</td>
<td>0</td>
<td>0.001</td>
<td>0</td>
<td>11</td>
<td>2.41</td>
<td>0.92</td>
</tr>
<tr>
<td>ESBCP4</td>
<td>17</td>
<td>5.48</td>
<td>30</td>
<td>0</td>
<td>0.003</td>
<td>0</td>
<td>51</td>
<td>1.88</td>
<td>0.87</td>
</tr>
<tr>
<td>ESBCP5</td>
<td>64</td>
<td>20.74</td>
<td>62</td>
<td>0</td>
<td>0.003</td>
<td>0</td>
<td>39</td>
<td>0.46</td>
<td>0.61</td>
</tr>
<tr>
<td>ESBCP6</td>
<td>81</td>
<td>20.66</td>
<td>43</td>
<td>0</td>
<td>0.003</td>
<td>0</td>
<td>73</td>
<td>52</td>
<td>2.87</td>
</tr>
<tr>
<td>ESBCP7</td>
<td>48</td>
<td>12.14</td>
<td>9</td>
<td>0</td>
<td>0.003</td>
<td>0</td>
<td>185</td>
<td>156</td>
<td>-2.59</td>
</tr>
<tr>
<td>ESBCP8</td>
<td>8</td>
<td>1.45</td>
<td>98</td>
<td>0</td>
<td>0.002</td>
<td>0</td>
<td>17</td>
<td>33</td>
<td>1.30</td>
</tr>
<tr>
<td>ESBCP9</td>
<td>57</td>
<td>7.83</td>
<td>25</td>
<td>0.001</td>
<td>0.003</td>
<td>86</td>
<td>0</td>
<td>-1.63</td>
<td>0.16</td>
</tr>
<tr>
<td>ESBCP10</td>
<td>25</td>
<td>1.72</td>
<td>14</td>
<td>0</td>
<td>0.004</td>
<td>260</td>
<td>4</td>
<td>-2.91</td>
<td>0.12</td>
</tr>
<tr>
<td>NBCI 1</td>
<td>51</td>
<td>35.65</td>
<td>73</td>
<td>0</td>
<td>0.003</td>
<td>59</td>
<td>159</td>
<td>-0.28</td>
<td>0.43</td>
</tr>
<tr>
<td>NBCI 2</td>
<td>15</td>
<td>0.44</td>
<td>15</td>
<td>0</td>
<td>0.001</td>
<td>409</td>
<td>27</td>
<td>-3.86</td>
<td>0.03</td>
</tr>
<tr>
<td>NBCI 3</td>
<td>1</td>
<td>0.0</td>
<td>90</td>
<td>0</td>
<td>0.003</td>
<td>286</td>
<td>43</td>
<td>-3.17</td>
<td>0.04</td>
</tr>
<tr>
<td>NBCI 4</td>
<td>14</td>
<td>3.18</td>
<td>51</td>
<td>0</td>
<td>0.03</td>
<td>227</td>
<td>195</td>
<td>-3.33</td>
<td>0.03</td>
</tr>
<tr>
<td>NBCI 5</td>
<td>3</td>
<td>0.0</td>
<td>72</td>
<td>0</td>
<td>0.001</td>
<td>280</td>
<td>0</td>
<td>-2.09</td>
<td>0.11</td>
</tr>
<tr>
<td>NBCI 6</td>
<td>50</td>
<td>10.91</td>
<td>70</td>
<td>0</td>
<td>0.001</td>
<td>15</td>
<td>6</td>
<td>0.25</td>
<td>0.56</td>
</tr>
<tr>
<td>NBCI 7</td>
<td>39</td>
<td>11.83</td>
<td>65</td>
<td>0</td>
<td>0.003</td>
<td>114</td>
<td>60</td>
<td>-1.82</td>
<td>0.14</td>
</tr>
<tr>
<td>NBCI 8</td>
<td>30</td>
<td>10.5</td>
<td>48</td>
<td>0</td>
<td>0.003</td>
<td>21</td>
<td>55</td>
<td>0.66</td>
<td>0.66</td>
</tr>
<tr>
<td>NBCI 9</td>
<td>29</td>
<td>8.1</td>
<td>88</td>
<td>0</td>
<td>0.003</td>
<td>101</td>
<td>56</td>
<td>-1.88</td>
<td>0.13</td>
</tr>
<tr>
<td>NBCI10</td>
<td>23</td>
<td>9.26</td>
<td>41</td>
<td>0</td>
<td>0.004</td>
<td>183</td>
<td>238</td>
<td>-3.99</td>
<td>0.02</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Coeff.</th>
<th>SE coeff.</th>
<th>Z</th>
<th>P</th>
<th>Odds ratio</th>
<th>Odds ratio 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>5.212</td>
<td>1.290</td>
<td>4.04</td>
<td>0.000</td>
<td>0.91</td>
<td>0.87</td>
</tr>
<tr>
<td>The amount of tongue fur</td>
<td>-0.09014</td>
<td>0.02479</td>
<td>-3.64</td>
<td>0.000</td>
<td>0.89</td>
<td>0.96</td>
</tr>
<tr>
<td>Maximum covering area of tongue fur</td>
<td>0.16016</td>
<td>0.04184</td>
<td>3.83</td>
<td>0.000</td>
<td>1.17</td>
<td>1.08</td>
</tr>
<tr>
<td>Thin tongue fur</td>
<td>-0.02701</td>
<td>0.01014</td>
<td>-2.66</td>
<td>0.008</td>
<td>0.97</td>
<td>0.95</td>
</tr>
<tr>
<td>The average area of red dot</td>
<td>-1.0442</td>
<td>546.2</td>
<td>-1.91</td>
<td>0.056</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Red dot in the spleen–stomach area</td>
<td>-0.014443</td>
<td>0.005123</td>
<td>-2.82</td>
<td>0.005</td>
<td>0.99</td>
<td>0.98</td>
</tr>
<tr>
<td>Red dot in the heart–lung area</td>
<td>-0.016043</td>
<td>0.006340</td>
<td>-2.53</td>
<td>0.011</td>
<td>0.98</td>
<td>0.97</td>
</tr>
</tbody>
</table>
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 Eq. (4). According to Eq. (1), the probability of infecting early-stage BC 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 BC patient, while that smaller than 0.5 represents those of a non-breast cancer individual. Table 7 lists the probability of infecting BC for 10 early-stage BC patients (ESBCP) and 10 non-breast cancer individuals (NBCI) in the testing group.
14 cases out of a total of 20 ones. An accuracy of 90% is achieved on non-breast cancer individuals and 50% accuracy is achieved on early-stage BC patients by employing Model III in predicting BC through tongue diagnosis.

4. Conclusions

According to the concepts of TCM, cancer is a systemic disease formed by imbalances between endogenous physical conditions of the body and exogenous pathogenic factors, included accumulated toxins, heat and blood stasis. TCM is one of the most widely used integrative medicine therapies used by BC patients worldwide. If TCM syndrome differentiation is applied, the standard for judgment should be clearly stated. TCM syndrome differentiation depends mainly on four ways of diagnosis. Among them, observation diagnosis ranks first, and tongue is the major subject of observation. Clinical studies of TCM in patients with breast cancer remain scant. And 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 non-breast cancer individuals.

This paper investigates discriminating tongue features to distinguish between earlier stage BC patients and non-breast cancer individuals via non-invasive means, expecting to find BC earlier and give treatment in time to raise 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 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 red dot in the spleen–stomach area \( p = 0.000 \), and 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, and logistic regression is subsequently performed once again. 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 BC. An accuracy of 80% is reached on non-breast cancer individuals 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 BC patients. For the second model employing 6 tongue features induced by logistic regression with independently significant meaning, 80% accuracy is achieved for non-breast cancer individuals and 60% accuracy is achieved for early-stage BC patients. The third model employing 5 tongue features induced by logistic regression with independently significant meaning, 90% accuracy is achieved for non-breast cancer individuals and 50% accuracy is achieved for early-stage BC patients. 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 early-stage BC patients and non-breast cancer individuals. This successful application of tongue diagnosis is mainly attributed to the high intra- and inter-observer agreements of the ATDS developed. By following the same methodology, early detection of other types of diseases can be explored. We expect a series of tongue diagnosis findings based on ATDS will be released in the near future. Also, a large tongue database composed of persons with different patterns satisfying pre-set criteria, can be established first to serve as a control group in subsequent studies to save both time and cost.

Conflict of interest

No authors have any conflict of interest to declare.

Acknowledgement

The authors are grateful to the anonymous reviewers for their valuable suggestions.

References

19. Lo LC, Chen YF, Chen WJ, Cheng TL, Chang JY. The study on the agreement between automatic tongue diagnosis system and traditional Chinese


