Objective tongue color and gastroesophageal reflux disease: Cross-sectional study

Mosaburo Kainuma,1,2* Norihiro Furusyo,2 Yoshihisa Urita,3 Hiroaki Ikezaki,2 Kazuya Ura,2 Toshiya Nakaguchi4 & Takao Namiki5

1 Community Medicine Education Unit, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan
2 Department of General Internal Medicine, Kyushu University Hospital, Fukuoka, Japan
3 Department of General Medicine and Emergency Care, Toho University School of Medicine, Omori Hospital, Tokyo, Japan
4 Center for Frontier Medical Engineering, Chiba University, Chiba, Japan
5 Department of Japanese Oriental (Kampo) Medicine, Graduate School of Medicine, Chiba University, Chiba, Japan

ABSTRACT

Aim: Gastroesophageal reflux disease (GERD) is a heterogeneous, multi-symptom disease. The Frequency Scale for the Symptoms of GERD (FSSG) is used to evaluate the disease severity and therapeutic response. We reported that objective tongue diagnosis would be a useful, non-invasive screening tool for erosive gastritis, which makes it a useful complement to Kampo diagnosis. This study was done to investigate the utility of tongue color (TC) and FSSG in the diagnosis of GERD.

Methods: A total of 886 healthy Japanese residents aged 28–86 years had a photograph of the tongue taken, completed the FSSG, and underwent esophagogastroduodenoscopy. FSSG ≥8 was defined as GERD, mucosal break on endoscopy as reflux esophagitis (RE), and no erosive esophagitis as non-erosive reflux disease (NERD). TC was measured using the device-independent International Commission on Illumination (CIE) 1976 L*a*b* color space standards at four points.

Results: A total of 152 (17.2%) of the subjects had FSSG ≥8. On endoscopy 36 (4.1%) had RE and 116 (13.1%) had NERD. Age, sex, and smoking were significant in the multivariate analysis of GERD (OR, 0.64, P = 0.043; OR, 0.65, P = 0.036; and OR, 2.06, P = 0.014, respectively). Sex and TC (3b*/1b*) were significant in the multivariate analysis of RE (OR, 4.82, P < 0.001; OR, 3.59, P = 0.018, respectively). On receiver operating characteristics analysis, a combination of sex (male) and TC (3b*/1b*) ≥1.24 was more predictive of RE than sex (male) alone (area under the curve, 0.75 vs 0.70).

Conclusion: Combining FSSG ≥8, sex, and a positive tongue color coating would be a useful, accurate, non-invasive screening method for GERD, especially RE.

KEY WORDS: Frequency Scale for the Symptoms of GERD, gastroesophageal reflux disease, reflux esophagitis, tongue color, tongue image analyzing system

INTRODUCTION

Gastroesophageal reflux disease (GERD) is defined as the condition that develops when the reflux of stomach contents causes troublesome symptoms and/or complications, and is classified based on the appearance of esophageal mucosa on upper endoscopy as reflux esophagitis (RE), or as non-erosive reflux disease (NERD) [1]. GERD is estimated to affect approximately 20–30% of the population worldwide. In Japan, the prevalence of GERD is lower than in the USA and Europe. In a recent Japanese study, acid reflux-related symptoms were reported at least on a monthly bias by 6.6–9.8% of the population, with an increasing incidence in GERD [2]. GERD is one of the most prevalent gastrointestinal diseases worldwide, and GERD affects the quality of life and creates an important economic burden on the healthcare system due to over-the-counter and prescribed medications, diagnostic procedures, and work absenteeism [3].

Gastroesophageal reflux disease is a heterogeneous, multi-symptom disease with no single gold standard for testing. Furthermore, the tools currently available for diagnosis all have limitations. Usually, a combination of several diagnostic tools, including questionnaires and endoscopy, are used to make the diagnosis [4]. The decision about which test will be used, when it will be used, and what will be the benefit of diagnosis depends on several factors related to cost and availability at the primary care level.

Several strategies have been developed for the diagnosis of GERD that do not rely on invasive testing. Questionnaires...
allow an objective assessment of symptoms [5]. The advantages of questionnaires include that they can be self-administered, used as screening tools, are of low cost, and that they can be applied to anyone with GERD. The Frequency Scale for the Symptoms of GERD (FSSG) questionnaire was developed in Japan and is widely used to screen for GERD [6–10].

Tongue diagnosis is one of the most important diagnostic methods of kampo medical practice, in which doctors observe tongue color (TC), gloss, shape, and tongue coating in their diagnosis of a patient’s health status [11]. Many studies have reported a correlation between the shape and color of the tongue and an individual’s health [12,13]. In contrast, many problems related to a lack of objectivity have been reported for kampo medicine. Lo et al. noted low rates of agreement in tongue diagnoses, particularly differences in judgment about TC, and reported the usefulness of an automated tongue diagnosis system [14]. A focus of this research project was to develop a system that would provide objective data to support traditional kampo diagnostic techniques. Computer tongue diagnostic support systems have been developed that use image processing [15–18]. Furthermore, there have been attempts to match conventional diagnosis of cardiovascular disease, gastric cancer, and rheumatoid arthritis to tongue images that seem to correspond to these diseases [19–22]. We constructed a tongue image analyzing system (TIAS) that can be used for computer-aided tongue diagnosis based on TC [23,24]. The key characteristic of the tongue imaging method in TIAS is the exclusion of the influence of external light using an integrating sphere to achieve an evenly distributed light intensity with a halogen light source. Further, TIAS can remove the gloss of the tongue surface from the images in order to stabilize the color of the tongue surface and the coating of the tongue.

We recently reported that objective tongue diagnosis with this TIAS, would be a useful, non-invasive screening tool for erosive gastritis [25].

In the present study, the aim was not only to objectively investigate the utility of TC and FSSG in the diagnosis of GERD, especially RE and NERD, but also to assess the predictive value of TC for GERD and discriminative values for RE and NERD as a way to complement traditional kampo diagnosis.

**METHODS**

**Subjects and study protocol**

The current study began in 2007 as a survey of the incidence of vascular events associated with lifestyle-related disease in residents of Kyushu and Okinawa as part of the Kyushu and Okinawa Population Study (KOPS) [26–29]. We follows the methods of Kainuma et al. 2015 [25]. Briefly, the participants were residents of Ishigaki City, Okinawa Prefecture, who took part in a program of esophagogastroduodenoscopy screening for gastric cancer between October 2012 and January 2013. In the study protocol, after giving informed consent, each participant filled in the FSSG questionnaire. The FSSG has 12 questions: five regarding acid reflux-related (ARD) symptoms and seven regarding dyspeptic symptoms (DS). Each question is scored to indicate the frequency of the symptoms as never, 0; occasionally, 1; sometimes, 2; often, 3; and always, 4. Subjects with FSSG score ≥8 were considered positive for GERD, as previously validated [6].

Second, the tongue was photographed with the TIAS, which was validated in previous studies [23,24]. Photography was conducted in a shady controlled room with the subject’s face fixed using chin and forehead rests. Because the mouth cannot be opened after both the chin and forehead are fixed, first the chin is placed on the chin rest and, after swallowing saliva, the mouth is opened and the tongue extended, following which the forehead is placed against the forehead rest. Each tongue extension is for 20 s, and images are taken every 100 ms, for a total of 200 images. The operator then visually confirms the TC. After taking the photographs, esophagogastroduodenoscopy was done. Each participant underwent esophagogastroduodenoscopy at the Ishigaki City Health Center, carried out by highly experienced endoscopists who performed each examination blinded to the serology. Furthermore, endoscopists did not perform TIAS prior to esophagogastroduodenoscopy. RE was defined according to the Los Angeles classification as A–D [30]. We then estimated the associations between TC, FSSG, and the endoscopy findings. FSSG ≥8 and the presence of mucosal break on endoscopy was defined as RE, and FSSG ≥8 with no erosive esophagitis as NERD. TC was measured using the device-independent International Commission on Illumination (CIE) 1976 L*a*b* color space standards at four points: 1, tongue edge; 2, tongue posterior; 3, tongue middle; and 4, tongue apex (Fig. 1). The coating does not grow on the edge of the tongue, and the color of the edge can be considered the color of the tongue body. In contrast, the color at the other three points is a mixture of the coating and the body of the tongue.

From the point of view of kampo medicine, the posterior of the tongue reflects kidney function (traditional medicine); the middle of the tongue reflects the stomach and pancreas (gastrointestinal (GI) tract traditional medicine); and the apex of the tongue reflects the heart (traditional medicine). Therefore, we used these areas to assess the association of TC with the endoscopy findings.

In addition, we calculated the ratio of L*a*b* of the tongue edge to that of the other three points to examine the association between the coating of the tongue and FSSG and the endoscopy findings. By calculating the ratio to the tongue edge (<1.0), we can confirm that TC reflects the body of the tongue, not the coating.

Esophagogastroduodenoscopy was done for 919 residents, and the data for 886 (age range, 28–86 years; mean age,
57.8 years; 384 men and 502 women) from whom consent for the FSSG and TC were obtained, were available for analysis.

To ensure the validity of the data, all physicians participating in the study were staff members of the Department of General Internal Medicine, Kyushu University Hospital or the Department of General Medicine and Emergency Care, Toho University School of Medicine Omori Hospital. All were trained with regard to the study protocol and the medical procedures necessary for the study. The study protocol was approved by the Ishigaki City Health Center and the Kyushu University Hospital Ethics Committee. Written informed consent was obtained from all participants prior to the examination. The study was conducted in accordance with the principles of the Helsinki Declaration of 1975, as revised in 2000.

**Statistical analysis**

Data are expressed as n (%), mean ± SD, or median (IQR). Participant characteristics and TC calculated from tongue photography were compared between participants with and without FSSG ≥8 points, RE, and NERD using Fisher’s exact test for categorical variables and unpaired t-test or Mann–Whitney test for continuous variables. Variables with $P < 0.1$ on univariate analysis were used in multivariate analysis to determine independent, significant predictors. Odds ratio (OR) and 95%CI were calculated from the multiple logistic regression after adjustment with each variable. Receiver operating characteristic (ROC) analysis was used to calculate sensitivity, specificity, and area under the curve (AUC). All analyses were conducted using JMP ver. 11 (SAS Institute, Cary, NC, USA) and $P < 0.05$ was considered significant.

**RESULTS**

**FSSG score classification**

Of the 886 subjects, 152 (17.2%) had FSSG ≥8 and 734 (82.8%) had FSSG <8 (Fig. 2). Furthermore, endoscopy indicated RE for 36 patients (4.1%) with a mucosal break in the distal esophagus. Patients with reflux symptoms but no mucosal break were diagnosed as having NERD (116, 13.1%). In both groups, there was no significant difference in the DS or ARD FSSG scores.

**Factors associated with GERD, RE, and NERD**

Gastroesophageal reflux disease was significantly associated with age, sex and smoking ($P < 0.001$, 0.033, 0.040, respectively; Table 1). Compared with NERD, RE was significantly associated with sex, body mass index (BMI), smoking, and hypertension ($P < 0.001$, 0.005, 0.009, 0.010, respectively; Table 1).

**GERD, RE, NERD and TC**

Gastroesophageal reflux disease was significantly associated with TC (4 L; $P = 0.016$; Table 2). Compared with NERD, RE was significantly associated with TC (2b*, 3a*, 2b*/1b*, 3b*/1b*, and 4b*/1b* (P = 0.039, 0.018, 0.001, 0.008, 0.048; Table 2).

**Positive predictors of GERD and RE**

Univariate analysis extracted age, sex, and smoking as significantly associated with a differential diagnosis of GERD (Table 3). Age, sex and smoking were significant in the multivariate analysis (OR 0.64 $P = 0.043$, OR 0.65 $P = 0.036$, OR 2.06, $P = 0.014$), but there was no correlation with TC (4 L ≥ 67.3).

Univariate analysis extracted sex, BMI, smoking and TC (3b*/1b*) as significantly associated with a differential diagnosis of RE. Sex and TC (3b*/1b*) were significant in the multivariate analysis (OR, 4.82, $P < 0.001$; OR, 3.59, $P = 0.018$).
FSSG, Frequency Scale for the Symptoms of Gastroesophageal reflux disease; RE, reflux esophagitis.

**Discrimination of RE on ROC curve analysis**

On ROC analysis, sex had sensitivity 67%, specificity 74%; and AUC, 0.70 (Fig. 3a). The combination of sex and TC(3b*1b*≥ 1.24) had a sensitivity of 78%, specificity of 64%; and AUC, 0.75 (Fig. 3b).

**DISCUSSION**

This is the first report to demonstrate an objective criterion by which TC can be used for the screening of GERD. A Japanese population study involving 10,837 patients used only endoscopy to diagnose RE on the presence of mucosal breaks, and NERD as the presence of heartburn and/or acid regurgitation in RE-free subjects. They found RE and NERD prevalences of 6.8% and 15.9%, respectively [19]. In a meta-analysis the prevalence of RE was 1.4–52.1% in Japan, with a mean prevalence of 9.1% for persons who underwent regular health checkups [2]. In the present study, 201 participants (22.7%) were diagnosed with RE, similar to or higher than that of other studies. Furthermore, on the present

### Table 1 | Factors associated with FSSG, RE and NERD

|                        | FSSG ≥ 8 (n = 152) | FSSG < 8 (n = 734) | P-value | RE (n = 36) | NERD (n = 116) | P-value |
|------------------------|---------------------|---------------------|---------|-------------|---------------|---------|
| Age (years)            | 53.9 ± 12.3         | 58.6 ± 11.2         | <0.001  | 56.8 ± 12.9 | 53.0 ± 12.0   | 0.107   |
| Male (%)               | 54 (35.5)           | 330 (45.0)          | 0.033   | 23 (63.9)   | 29 (25.0)     | <0.001  |
| BMI (kg/m²)            | 24.5 ± 4.5          | 243 ± 3.8           | 0.537   | 26.4 ± 4.6  | 23.9 ± 4.4    | 0.005   |
| Smoking (%)            | 22 (14.5)           | 66 (9.0)            | 0.040   | 10 (27.8)   | 12 (10.3)     | 0.009   |
| Hypertension (%)       | 53 (34.9)           | 265 (36.1)          | 0.773   | 19 (52.8)   | 34 (29.3)     | 0.010   |

BMI, body mass index; FSSG, Frequency Scale for the Symptoms of Gastroesophageal reflux disease; NERD, non-erosive reflux disease; RE, reflux esophagitis.

### Table 2 | FSSG, RE, and NERD: Association with tongue color

| Tongue color† | FSSG ≥ 8 (n = 152) | FSSG < 8 (n = 734) | P-value | RE (n = 36) | NERD (n = 116) | P-value |
|---------------|---------------------|---------------------|---------|-------------|---------------|---------|
| 1 L*          | 68.2 ± 3.8          | 67.7 ± 3.9          | 0.123   | 67.8 ± 5.0  | 68.4 ± 3.4    | 0.425   |
| 1a*           | 24.6 ± 4.2          | 24.9 ± 4.7          | 0.544   | 23.7 ± 4.2  | 24.9 ± 4.1    | 0.134   |
| 1b*           | 48 ± 2.9            | 5.0 ± 2.8           | 0.388   | 4.4 ± 2.6   | 4.9 ± 3.0     | 0.374   |
| 2 L*          | 66.2 ± 7.0          | 65.2 ± 9.7          | 0.230   | 65.8 ± 7.9  | 66.4 ± 6.7    | 0.697   |
| 2a*           | 146 ± 3.9           | 154 ± 6.0           | 0.146   | 136 ± 3.4   | 15.0 ± 4.0    | 0.066   |
| 2b*           | 5.6 ± 3.6           | 6.0 ± 3.8           | 0.239   | 6.7 ± 5.5   | 5.3 ± 2.7     | 0.039   |
| 3 L*          | 70.1 ± 3.3          | 69.6 ± 3.6          | 0.134   | 70.5 ± 3.9  | 69.9 ± 3.1    | 0.328   |
| 3a*           | 23.3 ± 4.1          | 23.5 ± 4.6          | 0.456   | 21.9 ± 3.1  | 23.7 ± 4.3    | 0.018   |
| 3b*           | 45 ± 0.2            | 4.8 ± 2.7           | 0.232   | 4.5 ± 3.1   | 4.5 ± 2.5     | 0.988   |
| 4 L*          | 67.3 ± 3.6          | 66.5 ± 3.8          | 0.016   | 67.5 ± 4.4  | 67.2 ± 3.3    | 0.703   |
| 4a*           | 29.9 ± 6.0          | 29.4 ± 5.6          | 0.352   | 29.1 ± 6.3  | 30.1 ± 5.9    | 0.360   |
| 4b*           | 5.5 ± 2.8           | 5.4 ± 2.9           | 0.763   | 5.2 ± 2.4   | 5.6 ± 2.9     | 0.477   |
| 2 L*/1 L*      | 0.97 ± 0.11         | 0.97 ± 0.14         | 0.573   | 0.97 ± 0.11 | 0.96 ± 0.10   | 0.958   |
| 2a*/1a*        | 0.60 ± 0.15         | 0.62 ± 0.22         | 0.235   | 0.58 ± 0.15 | 0.61 ± 0.15   | 0.385   |
| 2b*/1b*        | 1.33 ± 1.13         | 0.81 ± 10.65        | 0.547   | 1.87 ± 13.6 | 1.16 ± 100    | 0.001   |
| 3 L*/1 L*      | 1.03 ± 0.06         | 1.03 ± 0.05         | 0.861   | 1.04 ± 0.01 | 1.02 ± 0.01   | 0.062   |
| 3a*/1a*        | 0.96 ± 0.16         | 0.96 ± 0.17         | 0.819   | 0.96 ± 0.18 | 0.96 ± 0.15   | 0.549   |
| 3b*/1b*        | 0.92 ± 0.85         | 0.82 ± 2.85         | 0.690   | 1.24 ± 1.22 | 0.81 ± 0.67   | 0.008   |
| 4 L*/1 L*      | 0.99 ± 0.05         | 0.98 ± 0.05         | 0.374   | 1.00 ± 0.07 | 0.98 ± 0.05   | 0.180   |
| 4a*/1a*        | 1.23 ± 0.25         | 1.20 ± 0.22         | 0.152   | 1.25 ± 0.30 | 1.22 ± 0.23   | 0.589   |
| 4b*/1b*        | 1.23 ± 1.20         | 1.10 ± 1.59         | 0.316   | 1.57 ± 1.91 | 1.12 ± 0.84   | 0.048   |

†1, tongue edge; 2, tongue posterior; 3, tongue middle; 4, tongue apex. Tongue color was measured using the device-independent International Commission on Illumination (CIE) 1976 L*a*b* color space standards.

FSSG, Frequency Scale for the Symptoms of Gastroesophageal reflux disease; NERD, non-erosive reflux disease; RE, reflux esophagitis.
multivariate analysis, men were more likely to develop RE (men, 6.0%; women, 2.6%), whereas women were more likely to develop NERD (men, 7.6%; women, 17.3%), which supports the results of previous studies [31]. Studies supporting the objectivity of the FSSG questionnaire have shown that it has a stronger correlation with the endoscopic grade of esophagitis than does the Questionnaire for the Diagnosis of Reflux Esophagitis (QUEST) [32]. The FSSG is the standard questionnaire used in the diagnosis of GERD in Japan [6].

In this study, FSSG score ≥8 points was recorded for 17.2% of the participants. Other studies based on health screening in the Japanese population have reported FSSG score ≥8 points for 9.3% (112/1203), 26.0% (230/886), and 28.7% (574/2002) of their participants [6–8]. Furthermore, two of these studies reported RE at 5.3% (47/886) and at 4.5% (91/2002). NERD was reported at 20.7% (183/886) and at 24.1% (483/2002) [8,9]. The prevalence of GERD in the present study is similar to that of other areas of Japan. According to the present results, a combination of FSSG score and endoscopic esophagitis would be useful for making a diagnosis of RE or NERD, supporting previous studies [8,9].

In the present study GERD was also related to the color of the tongue coating because the coat color of the middle (3b*/1b*) had a significantly higher association with RE than with NERD; TC of the middle (3b*), however, was not significantly different. TC (3b*/1b*) and TC (2b*/1b*) were correlated with each other (r = 0.63), and TC (3b*/1b*) was selected as a potential predictor in the multivariate model. This large-scale study has shown that the analysis of TC by TIAS can be useful in the diagnosis of gastroesophageal disease. The tongue coating is reported to become more yellowish as gastric erosion becomes more severe [33,34]. In contrast, the color of the coating was more useful for predicting RE than for predicting erosive gastritis on a large scale. We feel that a TC near yellow is predictive of RE. In the tongue diagnosis of kampo medicine, the middle of the tongue reflects the function of the GI tract, which was supported by the present data: the coat color of the middle (3b*/1b*) had a significantly higher association with RE.
than with NERD. Furthermore, we previously reported that multivariate analysis extracted TC (3a*,3b*) as an independent factor associated with a differential diagnosis of erosive gastritis. In contrast, in the present study TC (3b*/1b*) was extracted in the multivariate analysis of RE. It is interesting that the differences in the pathophysiology of erosive gastritis and GERD are reflected by differences in TC.

In the current clinical practice of modern medicine, it is recommended that patients with high scores on the GERD Q, another GERD questionnaire, be treated without further investigation for GERD and that patients with a low probability of GERD be evaluated further using methods that include endoscopy and pH test. Patients with high scores are treated with a proton pump inhibitor (PPI) testing period of 4–8 weeks, and patients with low scores undergo further diagnostic testing [35,36]. The utility of this PPI testing period, however, is limited for patients diagnosed with the current standard tests [37]. In contrast, low scores on FSSG related to GERD indicate poor responsiveness to PPI treatment, and high scores indicate good responsiveness [38]. From this and the present results, we think that the combination of FSSG score ≥8, sex, and positive TC will be useful in the diagnosis of GERD and that this combination will lead to increased efficacy in the prescription of treatment with a PPI for 4–8 weeks. Further prospective studies are needed for confirmation.

The present study had some limitations. The first is that all of the subjects were Japanese people participating in a routine medical check-up, thus the results cannot be generalized to other populations. We were not able to determine if the various GERD symptoms were of sufficient severity to require medical treatment. A second limitation was the use of a questionnaire to determine the GERD symptoms. No esophageal impedance-pH monitoring or PPI testing was done in this study. Therefore, true NERD was not rigorously distinguished from acid hypersensitive esophagus or functional heartburn. The third limitation was the non-prospective design, which carries all the limitations inherent to non-prospective studies.

We have shown that combining FSSG ≥8, sex, and a positive tongue color coating would be a useful, accurate, non-invasive screening method for GERD.

ACKNOWLEDGMENTS

We greatly thank Drs Takeshi Ihara, Yuji Harada, and Kunimitsu Eiraku of the Department of General Internal Medicine, Kyushu University Hospital, for their assistance in this research. Also, we greatly thank Drs Hitoshi Nakajima and Toshiyasu Watanabe from the Department of General Medicine and Emergency Care, Toho University School of Medicine Omori Hospital, Tokyo, Japan, for their technical assistance. This study was supported, in part, by JSPS KAKENH Grant Number 12103377, by the Foundation for Total Health Promotion and a research grant (18lk0310051h0001) from Japan Agency for Medical Research and Development.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

REFERENCES

1. Vakil N, van Zanten SV, Kahrilas P, Dent J, Jones R. Global Consensus Group. The Montreal definition and classification of gastroesophageal reflux disease: A global evidence-based consensus. Am. J. Gastroenterol. 2006; 101: 1900–1920.

2. Fujiwara Y, Arakawa T. Epidemiology and clinical characteristics of GERD in the Japanese population. J. Gastroenterol. 2009; 44: 518–534.

3. Yadlapati R, Dakhoul L, Pandolfino JE, Keswani RN. The quality of care for gastroesophageal reflux disease. Dig. Dis. Sci. 2017; 62: 569–576.

4. Katz PO, Gerson LB, Vela MF. Guidelines for the diagnosis and management of gastroesophageal reflux disease. Am. J. Gastroenterol. 2013; 108: 308–328.

5. Mouli VP, Ahuja V. Questionnaire based gastroesophageal reflux disease (GERD) assessment scales. Indian J. Gastroenterol. 2011; 30: 108–117.

6. Kusano M, Shimoyama Y, Sugimoto S et al. Development and evaluation of FSSG: Frequency scale for the symptoms of GERD. J. Gastroenterol. 2004; 39: 888–891.

7. Maekita T, Kato J, Enomoto S et al. Japanese apricot improves symptoms of gastrointestinal dysmotility associated with gastroesophageal reflux disease. World J. Gastroenterol. 2015; 21: 8170–8177.

8. Matsu K, Fujita T, Watanabe N et al. Lifestyle factors associated with gastroesophageal reflux disease in the Japanese population. J. Gastroenterol. 2013; 48: 340–349.

9. Okuyama M, Takai O, Nakahara K et al. Associations among gastroesophageal reflux disease, psychological stress, and sleep disturbances in Japanese adults. Scand. J. Gastroenterol. 2016; 30: 1–6.

10. Watanabe T, Urita Y, Sugimoto M, Miki K. Gastroesophageal reflux disease symptoms are more common in general practice in Japan. World J. Gastroenterol. 2007; 21: 4219–4223.

11. Japan Society for Oriental Medicine. Introduction to Kampo. Tokyo: Elsevier Japan, 2005.

12. Hanne K, Ingelise T, Linda C, Ulrich PP. Oral status and the need for oral health care among patients hospitalised with acute medical conditions. J. Clin. Nurs. 2012; 21: 2851–2859.

13. Kudo T, Shimazu Y, Yagishita H et al. Three-dimensional reconstruction of oral tongue squamous cell carcinoma at invasion front. Int. J. Dent. 2013; 2013: 482765.
14. 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.

15. Chiu CC. A novel approach based on computerized image analysis for traditional Chinese medical diagnosis of the tongue. *Comput. Methods Programs Biomed.* 2000; **61**: 77–89.

16. Kim J, Han GJ, Choi BH et al. Development of differential criteria on tongue coating thickness in tongue diagnosis. *Complement. Ther. Med.* 2012; **20**: 16–22.

17. Wang X, Zhang B, Yang Z, Wang H, Zhang D. Statistical analysis of tongue images for feature extraction and diagnostics. *IEEE Trans. Image Process.* 2013; **22**: 5336–5346.

18. Zhang B, Kumar BV, Zhang D. Detecting diabetes mellitus and nonproliferative diabetic retinopathy using tongue color, texture, and geometry features. *IEEE Trans. Biomed. Eng.* 2014; **61**: 491–501.

19. Sun DZ, Liu L, Jiao JP, Wei PK, Xu L. Syndrome characteristics of traditional Chinese medicine: Summary of a clinical survey in 767 patients with gastric cancer. Zhong Xi Yi Jie He Xue Bao 2010; **8**: 332–340.

20. Su W, Xu ZY, Wang ZQ, Xu JT. Objectified study on tongue images of patients with lung cancer of different syndromes. *Chin. J. Integr. Med.* 2011; **17**: 272–276.

21. Jiang M, Zha Q, Lu C, He Y, Lu A. Association between tongue appearance in traditional Chinese medicine and effective response in treatment of rheumatoid arthritis. *Complement. Ther. Med.* 2011; **19**: 115–121.

22. Lo LC, Cheng TL, Chiang JY, Damdinsuren N. Breast cancer index: A perspective on tongue diagnosis in traditional Chinese medicine. *J. Tradit. Complement. Med.* 2013; **3**: 194–203.

23. Yamamoto S, Ishikawa Y, Nakaguchi T et al. Temporal changes in tongue color as criterion for tongue diagnosis in Kampo medicine. *Forsch. Komplementmed.* 2012; **19**: 80–85.

24. Nakaguchi T, Takeda K, Ishikawa Y et al. Proposal for a new noncontact method for measuring tongue moisture to assist in tongue diagnosis and development of the tongue image analyzing system, which can separately record the gloss components of the tongue. *Biomed. Res. Int.* 2015; **2015**: 249609.

25. Kainuma M, Furusyo N, Urita Y et al. The association between objective tongue color and endoscopic findings: Results from the Kyushu and Okinawa population study (KOPS). *BMC Complement. Altern. Med.* 2015; **15**: 372.

26. Furusyo N, Koga T, Ai M et al. Utility of glycated albumin for the diagnosis of diabetes mellitus in a Japanese population study: Results from the Kyushu and Okinawa Population Study (KOPS). *Diabetologia* 2011; **54**: 3028–3036.

27. Furusyo N, Ihara T, Hayashi T et al. The serum undercarboxylated osteocalcin level and the diet of a Japanese population: Results from the Kyushu and Okinawa Population Study (KOPS). *Endocrine* 2013; **43**: 635–642.

28. Furusyo N, Koga T, Ai M et al. Plasma glycated albumin level and atherosclerosis: Results from the Kyushu and Okinawa Population Study (KOPS). *Int. J. Cardiol.* 2013; **167**: 2066–2072.

29. Furusyo N, Ai M, Okazaki M et al. Serum cholesterol and triglyceride reference ranges of twenty lipoprotein subclasses for healthy Japanese men and women. *Atherosclerosis* 2013; **231**: 238–245.

30. Armstrong D, Bennett JR, Blum AL et al. The endoscopic assessment of esophagitis: A progress report on observer agreement. *Gastroenterology* 1996; **111**: 85–92.

31. Minatsuiki C, Yamamichi N, Shimamoto T et al. Background factors of reflux esophagitis and non-erosive reflux disease: A cross-sectional study of 10,837 subjects in Japan. *PLoS One* 2013; **8**: e69861.

32. Danjo A, Yamaguchi K, Fujimoto K et al. Comparison of endoscopic findings with symptom assessment systems (FSSG and QUEST) for gastroesophageal reflux disease in Japanese centres. *J. Gastroenterol. Hepatol.* 2009; **24**: 633–638.

33. Tosa H, Shimada Y, Terasawa K, Mitsuma T, Tanaka Y. Study on relationship between tongue coating and gastric lesion (in Japanese). *Gastroenterol. Endosc.* 1988; **30**: 303–313.

34. Ishizaki N, Yamamura Y, Egawa M et al. Relation between tongue mucosal findings and gastric mucosal lesion (in Japanese). *J. Jpn. Soc. Acupunc. Moxibust.* 1990; **40**: 299–305.

35. Vakil N. The initial diagnosis of GERD. *Best Pract. Clin. Gastroenterol.* 2013; **27**: 365–371.

36. Bergquist H, Agréus L, Tillander L, Johnsson F, Sörnård H, Sjöstedt S, Hellström PM. Structured diagnostic and treatment approach versus the usual primary care approach in patients with gastroesophageal reflux disease: A cluster-randomized multicenter study. *J. Clin. Gastroenterol.* 2013; **47**: e65–e73.

37. Bytzer P, Jones R, Vakil N, Unghard O, Lind T, Wernersson B, Dent J. Limited ability of the proton-pump inhibitor test to identify patients with gastroesophageal reflux disease. *Clin. Gastroenterol. Hepatol.* 2012; **10**: 1360–1366.

38. Komatsu-Tanaka M, Iwakiri R, Fujimoto K et al. Clinical symptoms of FSSG in gastroesophageal reflux disease are critical for PPI treatment: Japanese multi-centers with 185 patients. *Dig. Endosc.* 2012; **24**: 407–411.