



Research Paper

Tongue diagnosis system for quantitative assessment of tongue coating in patients with functional dyspepsia: A clinical trial



Juyeon Kim^a, Gajin Han^b, Seok-Jae Ko^b, Dong-Hyun Nam^c, Jae-Woo Park^b, Bongha Ryu^b, Jinsung Kim^{b,*}

^a Department of Clinical Korean Medicine, Graduate School, Kyung Hee University, Kyung Hee dae-ro 26, Dongdaemun-gu, Seoul 130-701, Republic of Korea

^b Department of Gastroenterology, College of Korean Medicine, Kyung Hee University, Kyung Hee dae-ro 26, Dongdaemun-gu, Seoul 130-701, Republic of Korea

^c Department of Biofunctional Medicine and Diagnosis, College of Korean Medicine, Sangji University, Sangjidae-gil 83, Wonju 220-702, Republic of Korea

ARTICLE INFO

Article history:

Received 31 March 2014

Received in revised form

19 May 2014

Accepted 4 June 2014

Available online 13 June 2014

Keywords:

Tongue diagnosis

Tongue diagnosis system

Tongue coating

Diagnostic agreement

Functional dyspepsia

ABSTRACT

Ethnopharmacological relevance: Tongue diagnosis is a significant procedure to examine the physiological and pathological changes of the human body in oriental medicine. However, the conventional method of tongue diagnosis including direct observation of tongue has limitations because of various external factors and subjective factors. Therefore, the current study investigated the usefulness of the tongue diagnosis system (TDS) as a diagnostic tool for evaluating tongue coating thickness (TCT) by assessing the agreement between the TDS and a gold standard established by assessors using the conventional method.

Materials and methods: The present study was designed as a prospective clinical trial including 60 patients with functional dyspepsia. The TDS was used to capture tongue images twice within a 30-min interval to assess its reproducibility. Tongue coating percentage was measured by the TDS, and TCT was classified as either no coating, thin coating, or thick coating according to the existing diagnostic criteria. After both TDS examinations, the weight of tongue coating was quantitatively measured, and the correlation between the actual quantity of tongue coating and the percentage of the coating measured by the TDS was analyzed using Pearson's correlation. After collecting tongue images in all 60 patients, tongue coating was evaluated using a conventional method by 5 well-trained assessors to establish the gold standard for evaluating TCT, which allowed us to assess the diagnostic agreement between the TDS and the gold standard. After 2 weeks, TCT evaluation was repeated by the same assessors using the same images but in a random order.

Results: The agreement between the TDS and the gold standard for evaluating TCT was almost perfect (weighted kappa, 0.840), as was the reproducibility of the TDS (weighted kappa, 0.851). The percentage of tongue coating measured by the TDS was significantly correlated with the weight of tongue coating ($r=0.442$, $p < 0.001$). The levels of intra-rater reliability ranged from substantial to almost perfect (range of weighted kappa, 0.777–0.923). The inter-rater reliability of 5 assessors was moderate (weighted kappa, 0.563).

Conclusions: The present study demonstrated that the TDS can be used as a diagnostic tool for the objective and standardized evaluation of TCT in actual clinical practice.

© 2014 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

Tongue diagnosis is an important diagnostic method in traditional Korean medicine (TKM) and traditional Chinese medicine (TCM) (Ko et al., 2012; Jiang et al., 2012). It is conducted to

examine the physiological function and pathological changes of the internal organs (Fei and Gu, 2007). Tongue diagnosis involves examination of the tongue body (its color, shape, moisture, and movement) and coating (its color, thickness, and distribution) (Chen, 2007). Changes in tongue features, such as the tongue body and coating, provide significant information for TCM diagnosis in clinical practice (Kim et al., 2008). The tongue coating is a fur-like substance covering the surface of the human tongue, and it is considered to reflect not only the characteristics of the pathogenic factors but also the condition of the internal organs, especially the stomach (Barbara, 2000; Sun et al., 2013). In addition, tongue

* Corresponding author. Tel.: +82 2 958 8895; fax: +82 2 958 9136.

E-mail addresses: kjycom5@gmail.com (J. Kim), kmdhgj@naver.com (G. Han), kokokoko119@hanmail.net (S.-J. Ko), omdnam@sangji.ac.kr (D.-H. Nam), pjw2907@hanmail.net (J.-W. Park), rbongh@khu.ac.kr (B. Ryu), oridoc@khu.ac.kr (J. Kim).

coating thickness (TCT) is believed to reflect the progression and degree of diseases in TKM (Fei and Gu, 2007).

In the conventional method of tongue diagnosis, TCT is evaluated by a direct observation of a protruded tongue by a practitioner, and TCT is classified as no coating (absence of tongue coating), thin coating (tongue body barely visible), or thick coating (tongue body not visible at all (Fei and Gu, 2007). However, the result is affected not only by external factors such as brightness but also by subjective factors related to experience, knowledge, and diagnostic skills of individual practitioners (Kim et al., 2008; Ko et al., 2012).

To overcome the above limitations, various tongue diagnosis systems (TDSs) have been recently developed (Jung et al., 2012). Generally, the TDS consists of several processes including image capturing and storage, color correction, tongue segmentation, and image analysis (Kim et al., 2013). In particular, in the process of image analysis, the tongue is analyzed according to selected diagnostic parameters (Kim et al., 2009). In the present study, the percentage of tongue coating was measured by the TDS, and tongue coating was classified as no coating, thin coating, or thick coating according to the diagnostic criteria developed by Kim et al. (2012). They developed the differential criteria for the evaluation of TCT in tongue diagnosis. In their study, a total of 24 well-trained and reliable assessors evaluated TCT on 50 tongue images captured using the TDS, which measured the tongue coating percentage. Using the proportional-odds model, they reported that 29.06% was a cut-off point to differentiate between no coating and thin coating, and 63.51% was a cut-off point to differentiate between thin and thick coating.

Functional dyspepsia (FD) is a functional gastrointestinal disorder characterized by chronic or recurrent dyspeptic symptoms without any structural abnormality (Tack et al., 2006). In Western countries, 11.5% to 14.7% of the general population suffers from FD (El-Serag and Talley, 2004), while in South Korea, a recent epidemiological survey reported that the prevalence of FD is approximately 25% (Jee et al., 2011). Therefore, we decided to include patients with FD in the present study because, according to TKM, FD is one of the common stomach diseases that are well reflected by various features of tongue coating (Zhang et al., 2010; Han et al., 2011).

There have been various attempts to develop a TDS for the objective and standardized tongue diagnosis in TKM. However, there have been few clinical trials examining its efficacy and safety as a diagnostic tool. In addition, thus far, there have been only few studies investigating the usefulness of TDS by comparing it with the conventional method. It is necessary to assess the usefulness of TDS in actual clinical practice as well as to further develop the system itself. Therefore, the aim of the present study was to investigate the usefulness of TDS as a diagnostic tool for the evaluation of TCT by assessing the diagnostic agreement between TDS and conventional method.

2. Methods and materials

2.1. Ethics approval

The protocol of the trial was approved by the Institutional Review Board of the Korean Medicine Hospital of the Kyung Hee University in Seoul, South Korea (IRB no. KOMCIRB-2013-01). The protocol identification number at <http://www.clinicaltrials.gov> is NCT01864837. The current study was conducted in accordance with the standards of the International Committee on Harmonization of Good Clinical Practice and the revised version of the Declaration of Helsinki. Written informed consent was obtained from all patients before enrollment.

2.2. Study subjects

Patients aged 20 years and older who met the Rome III criteria for FD (Tack et al., 2006) were recruited from the Korean Medicine Hospital of the Kyung Hee University. The inclusion and exclusion criteria for the current study are shown in the protocol paper already published (Kim et al., 2013).

2.3. Sample size

Because there were few relevant previous studies available, we estimated an approximate sample size on the basis of the number of outpatients with FD visiting the Korean Medicine Hospital of the Kyung Hee University, as described previously (Kim et al., 2013). A total of 60 patients were required to conduct this clinical trial.

2.4. Study design

The present study was designed as a prospective clinical trial for investigating the usefulness of the TDS in clinical practice by examining its efficacy and safety as a diagnostic tool. All eligible patients were instructed to avoid mouth rinsing as well as food and liquid intake for 4 h before the examination. TDS examination was performed under constant conditions. The method of examination was previously described in detail (Kim et al., 2013). From the tongue coating percentage measured by the TDS, TCT could be classified as either no coating, thin coating, or thick coating according to the existing diagnostic criteria (Kim et al., 2012). To assess the reproducibility of TDS, the images were again captured by the same method after 30 min. After both TDS examinations, the weight of tongue coating was quantitatively measured with a method described by Yaegaki and Sanada (1992) to analyze the correlation between the actual quantity of tongue coating and the percentage of the coating measured by the TDS. The method of coating collection was previously described in detail (Kim et al., 2013). After tongue images of all 60 patients had been collected, the first TCT evaluations with the images were performed by well-trained assessors to establish the gold standard for evaluating TCT in clinical practice. After 2 weeks, the second TCT evaluations with the same images were conducted under the same conditions and by the same assessors, but the images were presented in a random order. A flow chart of the trial is presented in Fig. 1.

2.5. Tongue diagnosis system

The TDS consisted of several components including an image acquisition system, LED illuminator, case, and analysis software. In particular, the analysis software of TDS was developed to calculate the percentage of tongue coating from the captured tongue image. The components and image analysis of TDS were previously described in detail (Kim et al., 2013).

2.6. Gold standard for evaluating TCT

To establish the gold standard for evaluating TCT in clinical practice, 5 assessors with 5 or more years of clinical experience and working at the Korean Medicine Hospital of the Kyung Hee University were involved in the present study. To obtain a reliable gold standard, they had been instructed in the conventional method of evaluating TCT along with the differential criteria before the trial. All 60 images captured in the first TDS examination were presented to assessors using a picture-viewer software and an LCD monitor (DN-50PZ66, 50" HD, 1366 × 768 resolution, LG, Korea). Each image was displayed for 10 s (10 min in total for the assessment of 60 images). All TCT evaluations were

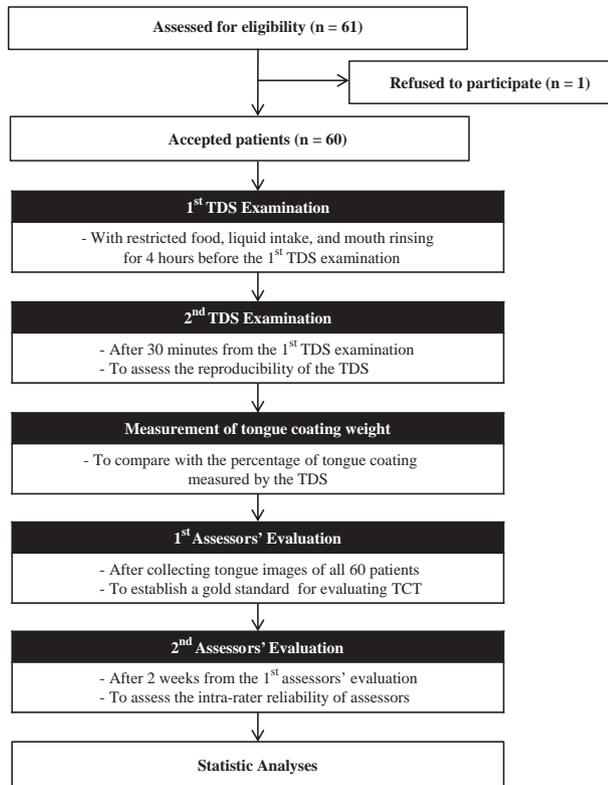


Fig. 1. Flow chart of the trial. TDS: tongue diagnosis system, TCT: tongue coating thickness.

simultaneously performed in the same room and under constant conditions such as brightness or distance between the assessor and the monitor. Assessors labeled each image as either “no coating,” “thin coating,” or “thick coating.” Each assessor was blinded to the evaluation done by other assessors as well as to the clinical data of the patients. The TCT evaluations of all 5 assessors were collected for each patient. If there was a disagreement, we chose the majority to establish the final diagnosis. The final diagnosis was considered as the gold standard for evaluating TCT in clinical practice.

2.7. Outcomes

2.7.1. Primary outcome

The percentage of tongue coating (%) was measured by the TDS, and TCT was classified as either no coating, thin coating, or thick coating according to the existing diagnostic criteria (Kim et al., 2012). After collecting tongue images captured by the TDS, TCT shown on the images was evaluated by 5 assessors. Thus, we could finally evaluate the diagnostic agreement between the results of the TDS and the gold standard established by assessors.

2.7.2. Secondary outcome

The reproducibility of the TDS was assessed by comparing the first and second TCT measurements in each patient. The weight of tongue coating (mg) was quantitatively measured to analyze the correlation with the percentage of tongue coating measured by the TDS. In addition, the intra- and inter-rater reliability were calculated to ensure the validity of TCT assessments by assessors.

2.8. Statistical analyses

All analyses in the current study were based on the intention-to-treat principle. The baseline characteristics of the subjects were

presented as numbers (%) for categorical variables and mean ± standard deviation for continuous variables. The weighted kappa (k) statistics was used to assess the diagnostic agreement in evaluating TCT between the TDS and the gold standard established by assessors as well as to assess the reproducibility of the TDS by comparing the two TCT measurements (Cohen, 1968). The correlation between the weight and percentage of the tongue coating measured by the TDS was analyzed by the Pearson's correlation coefficient. In addition, to ensure the validity of TCT evaluations by assessors, Fleiss' kappa (k) coefficient (Fleiss, 1971) was calculated to assess the inter-rater reliability and weighted kappa (k) statistics to assess the intra-rater reliability. The kappa values were interpreted according to Landis and Koch (1977) ($k < 0.1$ = poor; $0.1 \leq k \leq 0.2$ = slight; $0.2 < k \leq 0.4$ = fair; $0.4 < k \leq 0.6$ = moderate; $0.6 < k \leq 0.8$ = substantial; $0.8 < k \leq 1.0$ = almost perfect). All statistical analyses were performed using PASW Statistics 18 (SPSS Inc., Chicago, Illinois, United States) and SAS software, version 9.1.3 (SAS Institute Inc., Cary, North Carolina, United States). A p-value of < 0.05 was considered statistically significant.

3. Results

3.1. Study participants and baseline characteristics

Of 61 eligible subjects, 60 patients were included in the study between February and July of 2013. One subject refused to participate for personal reasons (Fig. 1). The baseline characteristics of the subjects are presented in Table 1. The classification of TCT by the TDS and assessors is presented in Table 2.

3.2. Adverse event

After the twice TDS examinations and measurement of tongue coating weight, we assessed the safety issues. There was no adverse event in the present study.

3.3. Outcome variables

3.3.1. Agreement between TDS and gold standard in evaluating TCT

We assessed the diagnostic agreement in evaluating TCT between the TDS and the gold standard established by assessors,

Table 1
Characteristics of study subjects (N=60).

Variables	
Sex (Male / Female), n (%)	12 (20)/48 (80)
Age (Mean ± SD)	48.18 ± 11.33
Duration of functional dyspepsia, n (%)	
< 5 years	22 (36.67)
5–10 years	11 (18.33)
10 years <	27 (45)

Table 2
TCT classifications by TDS and assessors.

TCT classification	No coating n (%)	Thin coating n (%)	Thick coating n (%)
TDS measurement	11(18.3)	43(71.7)	6(10.0)
Gold standard by assessors	12(20.0)	39(65.0)	9(15.0)

TCT: Tongue coating thickness, TDS: Tongue diagnosis system. The numbers mean the frequency of the TCT evaluations.

Table 3
Agreement between TDS and gold standard by assessors in evaluating TCT.

Variable	Agreement rate (%)	Weighted kappa	Standard error (95% CI)	Level of agreement
TCT	90	0.840	0.128(0.589–1.000)	Almost perfect

TDS: Tongue diagnosis system, TCT: Tongue coating thickness.

Table 4
Reproducibility of TDS between the twice TCT measurements.

Variable	Agreement rate (%)	Weighted kappa	Standard error (95% CI)	Level of agreement
TCT	91.67	0.851	0.128(0.600–1.000)	Almost perfect

TDS: Tongue diagnosis system, TCT: Tongue coating thickness.

Table 5
Intra-rater reliability of assessors in evaluating TCT.

Assessor	Weighted kappa	Standard error (95% CI)	Level of agreement
1	0.923	0.129(0.670–1.000)	Almost perfect
2	0.777	0.128(0.526–1.000)	Substantial
3	0.870	0.129(0.617–1.000)	Almost perfect
4	0.813	0.127(0.564–1.000)	Almost perfect
5	0.836	0.129(0.583–1.000)	Almost perfect

TCT: Tongue coating thickness.

Table 6
Inter-rater reliability of assessors in evaluating TCT.

Variable	Number of raters	Fleiss' kappa	Standard error (95% CI)	Level of agreement
TCT	5	0.563	0.022 (0.520–0.606)	Moderate

TCT: Tongue coating thickness.

and the results are presented in [Table 3](#). The level of the agreement was almost perfect (weighted kappa, 0.840).

3.3.2. Reproducibility of TDS

The kappa value for assessing the reproducibility of TDS is presented in [Table 4](#). In the evaluation of TCT by the TDS, the reproducibility was almost perfect (weighted kappa, 0.851).

3.3.3. Correlation between TDS measurement and the weight of tongue coating

The percentage of tongue coating measured by the TDS was significantly correlated with the weight of tongue coating, which indicated the actual quantity of tongue coating ($r=0.442$, $p < 0.001$).

3.3.4. Intra- and inter-rater reliability of assessors

The kappa values for intra-rater reliability of each assessor in the evaluation of TCT are presented in [Table 5](#). The levels of intra-rater reliability ranged from substantial to almost perfect (range of weighted kappa, 0.777–0.923). In addition, the inter-rater reliability of the assessors was moderate and the kappa value was 0.563 as presented in [Table 6](#).

4. Discussion

The objective of the present study was to investigate the usefulness of TDS in clinical practice by examining its efficacy and safety as a diagnostic tool. By comparing the TDS with the conventional method, we demonstrated that the diagnostic agreement between the TDS and the gold standard established by assessors is almost perfect. In addition, we observed that TDS had almost perfect reproducibility in the evaluation of TCT. Furthermore, no adverse event associated with the use of TDS was reported in this study.

The conventional method of tongue diagnosis in TKM involves a direct examination of the tongue by practitioners, so the results are affected by practitioners' subjectivity as well as external factors (Ko et al., 2012). Thus, it is necessary to ensure a high level of agreement between practitioners for objective tongue diagnosis.

There have been several studies investigating the agreement between practitioners in tongue diagnosis in TKM and TCM. In a study on 658 patients with stroke assessed by 2 TCM experts, Ko et al. (2012) reported different levels of the inter-observer agreement, ranging from fair to substantial. In particular, the inter-observer agreement in evaluating TCT was moderate (range of kappa, 0.49–0.60). Li et al. (2010) examined the agreement in the assessment of tongue and pulse signs between 2 TCM clinicians in 55 patients, and also reported a moderate agreement in the assessment of TCT (kappa, 0.525). The above findings were similar to our results in that the levels of the agreement were moderate. Kim et al. (2008) evaluated the reliability of tongue examination by 30 TCM practitioners assessing 10 tongue images taken with a digital camera. They analyzed not only the inter-reliability but also the intra-reliability of practitioners in evaluating TCT, however, both of the reliability failed to reach the agreement rate of 80%. Nonetheless, the analysis method of their study was consistent with that of our study.

The present study is distinct from the above mentioned studies in that we focused on the diagnostic agreement between the TDS and the gold standard established by assessors as well as on the intra- and inter-rater reliability of assessors. To our knowledge, this is the first clinical trial to examine the efficacy and safety of the TDS as a diagnostic tool by comparing it with the conventional method. In addition, we assessed the reproducibility of the TDS by comparing the repeated measurements. If the level of the diagnostic agreement between the TDS and the reliable gold standard established by assessors is shown to be high and the reproducibility of TDS is confirmed, and the safety for using TDS is ensured, TDS can be considered as a useful tool to facilitate an objective diagnosis of TCT in actual clinical practice.

First, the establishment of a reliable gold standard is essential. In addition, it is necessary to train assessors in the evaluation of TCT according to the existing diagnostic criteria. There have been several studies emphasizing the importance of training based on the diagnostic criteria in investigating diagnostic agreement. Zhang et al. (2008) demonstrated that training based on the established TCM diagnostic criteria improved the agreement of TCM diagnosis in patients with rheumatoid arthritis. Son et al. (2012) also investigated the importance of training in TCT evaluation for the standardization of tongue diagnosis in TKM. They found that the agreement in evaluating TCT among 15 clinicians after training was higher than that before training. Therefore, we included 5 assessors with 5 or more years of clinical experience and sufficiently trained them to differentiate TCT before participating in the trial. As a result, we obtained the intra-rater reliability that ranged from substantial to almost perfect and moderate inter-rater reliability. Therefore, we believe that we established a reliable gold standard for the objective evaluations of TCT.

Second, it is necessary to ensure the performance of TDS. In this study, we found that the TDS had almost perfect reproducibility for evaluating TCT. In addition, the tongue coating percentage measured by TDS was significantly correlated with the actual quantity of tongue coating, which indicates that TDS could quantitatively measure the percentage of tongue coating. Moreover, we did not observe any adverse events associated with the use of TDS.

The limitation of the present study is that the inter-rater reliability of the assessors was only moderate although they received sufficient training. For more standardized evaluations of TCT, the reliability among assessors needs to be improved by more consistent and longer training. In addition, further studies with a bigger sample size are needed to provide a rationale for a wider use of TDS in clinical practice.

In conclusion, the results of our study demonstrated that TDS can be used as an objective and standardized method of TCT evaluation in actual clinical practice. We believe that the present study may contribute to making the diagnosis system in TKM more objective and standardized.

Acknowledgments

This study was supported by a grant of the Korean Medicine Advanced Technology R&D Project, Ministry for Health, Welfare & Family Affairs, Republic of Korea (B100030).

References

- Barbara, K., 2000. Atlas of Chinese Tongue Diagnosis. Eastland Press, Seattle, Wash, USA.
- Chen, J., 2007. Chinese Medicine Study Guide: Diagnostics. People's medical publishing house, Beijing.
- Cohen, J., 1968. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin* 70, 213.
- El-Serag, H.B., Talley, N.J., 2004. Systemic review: the prevalence and clinical course of functional dyspepsia. *Alimentary Pharmacology & Therapeutics* 19, 643–654.
- Fei, Z., Gu, Y., 2007. Mirror of Health: Tongue Diagnosis in Chinese Medicine. People's Medical Publishing House, Beijing.
- Fleiss, J.L., 1971. Measuring nominal scale agreement among many raters. *Psychological Bulletin* 76, 378–382.
- Han, G., Kim, J., Park, J.W., Ryu, B., 2011. Pattern identification of 97 functional dyspepsia patients and the characteristics of each pattern type. *Journal of Korean Oriental Medicine* 32, 42–62.
- Jee, S.R., Jung, H.K., Min, B.H., Choi, K.D., Rhee, P.L., Kang, Y.W., Lee, S.I., Korean Society of Neurogastroenterology and Motility, 2011. Guidelines for the treatment of functional dyspepsia. *The Korean Journal of Gastroenterology* 57, 67–81.
- Jiang, B., Liang, X., Chen, Y., Ma, T., Liu, L., Li, J., Jiang, R., Chen, T., Zhang, X., Li, S., 2012. Integrating next-generation sequencing and traditional tongue diagnosis to determine tongue coating microbiome. *Scientific Reports* 2, 936.
- Jung, C.J., Jeon, Y.J., Kim, J.Y., Kim, K.H., 2012. Review on the current trends in tongue diagnosis systems. *Integrative Medicine Research* 1, 13–20.
- Kim, J., Son, J., Jang, S., Nam, D., Han, G., Yeo, I., Ko, S., Park, J., Ryu, B., Kim, J., 2013. Availability of tongue diagnosis system for assessing tongue coating thickness in patients with functional dyspepsia. *Evidence-Based Complementary and Alternative Medicine* 2013, 348272.
- Kim, J., Han, G.J., Choi, B.H., Park, J.W., Park, K., Yeo, I.K., Ryu, B.H., 2012. Development of differential criteria on tongue coating thickness in tongue diagnosis. *Complementary Therapies in Medicine* 20, 316–322.
- Kim, J., Jung, Y., Park, K., Park, J.W., 2009. A digital tongue imaging system for tongue coating evaluation in patients with oral malodour. *Oral Diseases* 15, 565–569.
- Kim, M., Cobbin, D., Zaslowski, C., 2008. Traditional Chinese Medicine tongue inspection: an examination of the inter- and intrapractitioner reliability for specific tongue characteristics. *Journal of Alternative and Complementary Medicine* 14, 527–536.
- Ko, M.M., Lee, J.A., Kang, B., Park, T., Lee, J., Myeong, S.L., 2012. Interobserver reliability of tongue diagnosis using traditional Korean medicine for stroke patients. *Evidence-Based Complementary and Alternative Medicine* 2012, 209345.
- Landis, J.R., Koch, G.G., 1977. The measurement of observer agreement for categorical data. *Biometrics* 33, 159–174.
- Li, X., Wen, Z., Liang, W., Xu, W., Li, S., Yu, X., Wang, H., 2010. Evaluation of consistency of tongue and pulse signs observed by traditional Chinese medicine clinicians. *Zhong Xi Yi Jie He Xue Bao* 8, 1153–1158.
- Son, J., Kim, J., Park, J.W., Ryu, B.H., 2012. A proposal for standardization of tongue diagnosis based on diagnostic criteria of tongue coating thickness. *The Journal of Korean Oriental Internal Medicine* 34, 1–13.
- Sun, Z.M., Zhao, J., Qian, P., Wang, Y.Q., Zhang, W.F., Guo, C.R., Pang, X.Y., Wang, S.C., Li, F.F., Li, Q., 2013. Metabolic markers and microecological characteristics of tongue coating in patients with chronic gastritis. (Research article). *BMC Complementary and Alternative Medicine* 13, 227.
- Tack, J., Talley, N.J., Camilleri, M., Holtmann, G., Hu, P., Malagelada, J.R., Stanghellini, V., 2006. Functional gastroduodenal disorders. *Gastroenterology* 130, 1466–1479.
- Yaegaki, K., Sanada, K., 1992. Volatile sulfur compounds in mouth air from clinically healthy subjects and patients with periodontal disease. *Journal of Periodontal Research* 27, 233–238.
- Zhang, G.G., Singh, B., Lee, W., Handwerger, B., Lao, L., Berman, B., 2008. Improvement of agreement in TCM diagnosis among TCM practitioners for persons with the conventional diagnosis of rheumatoid arthritis: effect of training. *Journal of Alternative and Complementary Medicine* 14, 381–386.
- Zhang, S.S., Wang, H.B., Li, Q.G., 2010. Clinical consensus on functional dyspepsia in traditional Chinese medicine. *Chinese Journal of Integrated Traditional and Western Medicine* 30, 533–537.