

Literature Research

Reliability, Accuracy, and Use Frequency of Evaluation Methods for Amount of Tongue Coating*

Su-Ryun Kim and Dong-Hyun Nam

ABSTRACT **Objective:** To classify the evaluation methods for amount of tongue coating (TC) and investigate their reliability, accuracy, and frequency of use. **Methods:** Articles published from 1985 to 2015 were searched for evaluation methods for the amount of TC in PubMed and the Cochrane Library. Only clinical researches were included except protocol articles. The methods were classified according to their characteristics. **Results:** Finally, 113 articles were selected. The evaluation method for the amount of TC from the articles was classified into 4 types: intuitive, specificative, computerized, and weighing TC. The reliability in the intuitive and specificative methods ($\kappa = 0.33\text{--}0.92$) showed varying levels among the studies. In general, the amount of TC calculated by the specificative method (Spearman's $r = 0.68\text{--}0.80$) was more strongly related to the directly measured value than to the value estimated by the computerized method (Pearson's $r = 0.442$). The number of articles published on this topic has increased consistently, and the specificative method was the most frequently used. Despite the higher reliability of the computerized method, it has not been widely used. **Conclusions:** The high prevalence of the specificative method would continue in clinical practice because of its convenience and accuracy. However, to establish higher reliability, the limitation of the subjectivity of the assessors should be overcome through calibration training. In the computerized method, novel algorithms are needed to obtain a higher accuracy so that it can help the practitioners confidently estimate the amount of TC.

KEYWORDS tongue coating, evaluation methods, reliability, accuracy, frequency

Tongue diagnosis is used as an important diagnostic method in traditional Korean medicine (TKM) and Chinese medicine (CM) because the tongue rapidly reflects physiological and pathological changes in the human body.^(1,2) The tongue is classified into tongue body and tongue coating when it is examined. Tongue coating (TC) is the fur-like matter attached to the dorsum of the tongue; it consists of bacteria, metabolites, exfoliated epithelium, leukocytes, and food waste.⁽³⁾ When the practitioner examines the TC, its amount, color, humidity, and exfoliation are considered. Among these factors, the amount of tongue coating (ATC) is one of the primary diagnostic points. In TKM and CM, TC is related to the internal organs, especially digestive organs, and the ATC reflects the severity of disease.^(1,4) In Western medicine, the main reason for evaluating TC is to assess the sanitary conditions in the oral cavity; the ATC influences the severity of halitosis⁽⁵⁾ and might be a risk factor for diseases such as pneumonia.⁽⁶⁾

For these clinical uses, various methods have been developed for evaluating ATC. Until now, researchers have used different methods for evaluating ATC; however,

their reliability and accuracy have not been consistent in terms of study results. This has hindered researchers from discussing their findings with a high degree of accuracy.

There are some previous reviews^(7–9) dealing with the evaluation of TC; however, they did not focus on evaluating TC, especially the ATC. Therefore, the purpose of this study was to classify the methods of evaluating ATC, and to investigate their reliability and accuracy, and the frequency of their use. Additionally, we want to present ideas to improve the evaluation methods for ATC on the basis of the information from our investigation.

©The Chinese Journal of Integrated Traditional and Western Medicine Press and Springer-Verlag GmbH Germany, part of Springer Nature 2018

*Supported by the New Researcher Support Program through the National Research Foundation of Korea Funded by the Ministry of Science, ICT & Future Planning, Republic of Korea (No. NRF-2013R1A1A1059950)

Department of Biofunctional Medicine and Diagnosis, College of Korean Medicine, Sangji University, Wonju 26339, Republic of Korea

Correspondence to: Dr. Dong-Hyun Nam, Tel: 82-33-738-7504, Fax: 82-33-730-0653, E-mail: omdnam@sangji.ac.kr

DOI: <https://doi.org/10.1007/s11655-018-2552-y>

METHODS

In January 2016, we searched for articles involving the evaluation methods for ATC, published during the past 30 years, from two online databases, PubMed and the Cochrane Library. The search words were 'tongue coating' or 'tongue coat' or 'tongue fur', accompanied by words such as 'thickness', 'amount', 'evaluation', 'assessment', 'score', and 'index'. Only articles written in English were included. We finally selected a total of 113 articles. The earliest of the selected articles was published in 1992; we could find no article matching the criteria before 1992. In some articles in the final selection, the authors did not describe the details of the method they used, but cited other papers using the evaluation methods for ATC; in those cases, we reviewed the cited papers to ascertain the details of the methods, although these articles were not included in the selection criteria. The search process is outlined in Figure 1. The methods were classified according to their characteristics, and then their reliability, accuracy, and frequency of use were investigated.

RESULTS

Classifications of Evaluation Methods for ATC

The evaluation method for ATC from the articles was classified into 4 types: (1) the intuitive method; (2) the specificative method; (3) the computerized method; and (4) the weighing TC method, i.e., scraping the TC and then

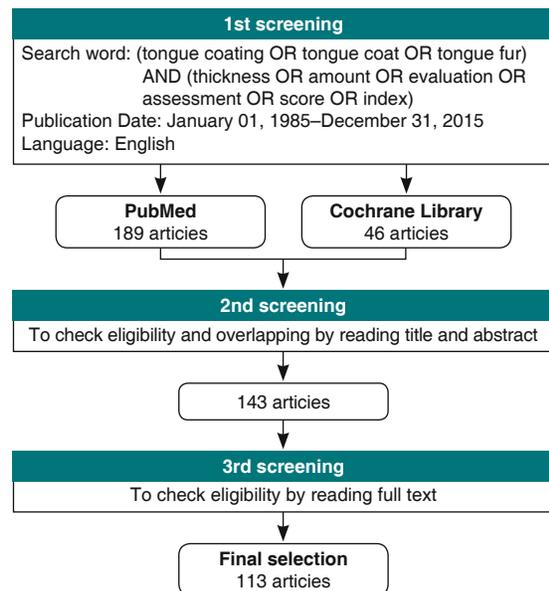


Figure 1. Process for Selecting Articles about Tongue Coating

measuring its physical weight. The criteria for more detailed classification were the following: ways and numbers of division; range of the score; and evaluation indices (thickness/distribution/both thickness and distribution). The summary of each method for evaluating ATC is illustrated as Figure 2.

Intuitive Method

This method involved practitioners observing the

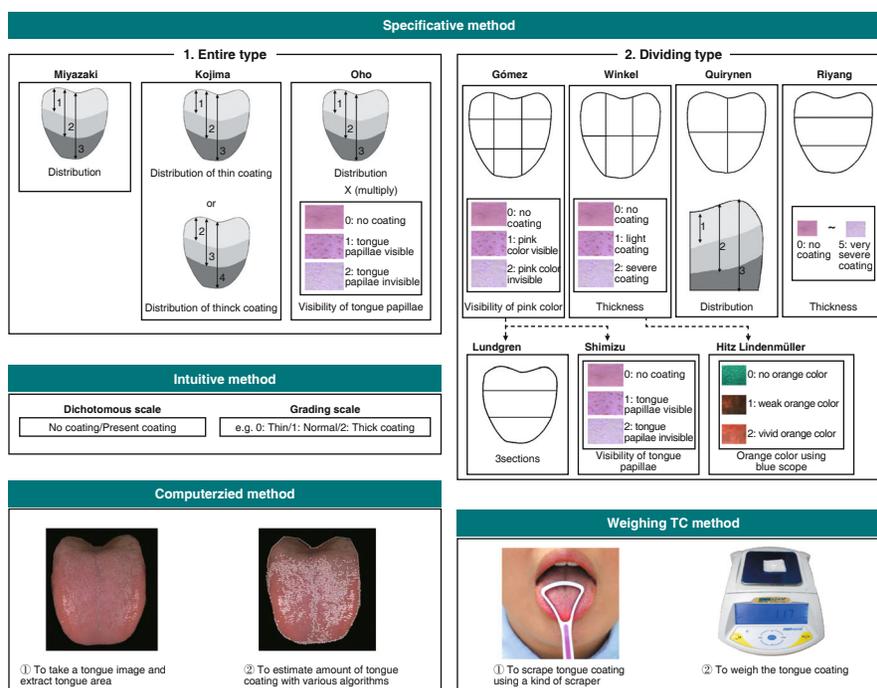


Figure 2. Summary of the Evaluation Methods for the Amount of Tongue Coating

Notes: The evaluation methods for ATC from the articles were classified into the intuitive, the specificative, the computerized, and the weighing method.

tongue with their naked eyes and assessing ATC by using their intuition and previous experience. The simplest of the intuitive method was to determine on a dichotomous scale whether TC is present.^(6,10-14) Another method proposed by Gross, et al⁽¹⁵⁾ in the 1970s was to give scores according to the severity of ATC, and Çiçek, et al⁽¹⁶⁾ applied it. In this method, they gave a score using 4 grades, from score 0 (no coating) to score 3 (severe coating). Moreover, there were modified methods for the range of scores,^(12,17-25) such as 3 grades (thin/normal/thick).

Specificative Method

The specificative method involved evaluating ATC with naked eyes according to predetermined criteria. This method was classified into two subtypes: entire type and dividing type. In the entire type, ATC of the whole tongue was assessed simultaneously, accordance with the criteria such as the distribution and the thickness of TC. This type was sub-classified as the first author's name of the studies which initially proposing the criteria: Miyazaki,⁽²⁶⁻⁵⁰⁾ Kojima⁽⁵¹⁻⁶²⁾ and Oho.⁽⁶³⁻⁷⁴⁾ Whereas, the dividing type was to divide the dorsum of the tongue into several sections and then gave a score for each section according to defined criteria as the number of divisions and the way in which the score was given. This type was also sub-classified: Gómez,⁽⁷⁵⁻⁷⁸⁾ Lundgren,^(8,41,42,79) Shimizu,⁽⁸⁰⁻⁸⁷⁾ Winkel,^(3,41,42,85,82-104) Hitz Lindenmüller,⁽¹⁰⁰⁾ Quirynen,^(29,30,105) and Riyang.⁽¹⁰⁶⁾ The details of the specificative method were described in Appendix 1.

Computerized Method

This method was the most objective for TC evaluation. A tongue image was taken under a constant light source and the tongue area was extracted from the background, such as face, teeth, and lips. The ATC was then assessed by a software program using various algorithms. Various kinds of the equipment dealing with this process, called as the tongue image analysis system (TIAS), were developed and used.^(3,25,94,107-109) The details were also explained as Appendix 2.

Weighing TC Method

This was the most direct method, measuring the weight of TC to evaluate ATC. There were two subtypes of the measurement: wet weight and dry weight. Wet weight type was used initially by Yaegaki and Sanada⁽⁵⁾ in 1992. The moisture on the tongue surface was cleaned with cotton rolls or pulp tissue papers. The TC was scraped from the sulcus terminalis to the apex of the tongue. Then, the wet weight of TC was weighed for evaluating ATC.^(8,25,41,42,79,101,110-115) Dry weight type, initially proposed by Monteiro-Amado, et al⁽¹¹²⁾

was that the scraped wet TC should be dried for 24 h at room temperature and then the dried TC be weighed.^(42,115)

Reliability of the Evaluation Methods for ATC

A description of the reliabilities of the evaluation methods for ATC was found in a total of 17 articles. The reliability was presented as various indices, such as kappa (κ) value, the percentage (%) of agreement, and correlation coefficients (r). The reliability levels of each method are summarized in Appendix 3.

For the intuitive method, Quirynen, et al⁽¹⁰⁾ presented the reliability of the dichotomous scale; the inter-rater reliability was 0.74, and intra-rater reliability was 0.85. Kim, et al⁽¹²⁾ also reported that the inter- and intra-rater reliability of the dichotomous scale was high when 30 raters assessed 10 tongue images twice (inter-rater agreement level was 80%–100%; 96.7% of the raters achieved over 80% intra-rater agreement level). The reliability of the three-grade scale in the intuitive method was studied by Kim, et al.⁽⁹⁴⁾ Sixty raters assessed 50 images of tongues and the TC thickness of each image was determined by their majority decision. While the match rate between an individual assessor's decision and the majority decision varied from 58%–88%, inter-rater reliability between 24 assessors over 80% was fair to good (Fleiss' $\kappa = 0.558$) using Fleiss's guidelines.⁽¹¹⁶⁾ In the other study of Kim, et al⁽²⁵⁾ the inter-rater reliability of five well-trained assessors was also fair to good (Fleiss' $\kappa = 0.563$) and intra-rater reliability of each assessor varied from 0.777 to 0.923 in weighted κ value.

In the specificative method, the reliability of the entire type had a quite wide range. According to guidelines characterized by Landis and Koch,⁽¹¹⁷⁾ the inter-rater reliability of Miyazaki's method was almost perfect ($\kappa = 0.92$,⁽³⁴⁾ $\kappa \geq 0.8$ ⁽⁴⁷⁾). However, another study⁽⁸⁰⁾ reported that inter- and intra-rater reliability were fair ($\kappa = 0.33$) and moderate ($\kappa = 0.42$) respectively. The reliability of Kojima's method was moderate ($\kappa = 0.56$ – 0.58).⁽⁵⁶⁾ For Oho's method, inter- and intra-rater reliability were both almost perfect ($\kappa = 0.82$,⁽⁶⁶⁾ $\kappa = 0.9$ ⁽⁷⁴⁾).

The reliability of the dividing type in the specificative method also showed various levels. In Gómez's method, the agreement rate between the raters was 58%, and within the raters was 70% on average.⁽⁷⁵⁾ For Shimizu's method, which was a clarified version of Gómez's, inter- and intra-rater reliability ($\kappa = 0.66$, $\kappa = 0.80$) were better than those of Gómez's.⁽⁸⁰⁾ The inter-rater reliability of Lundgren's method was studied in two different ways. In the study by Chérel, et

al⁽⁷⁹⁾ the relationship between scores by the raters was very strong (Pearson's $r=0.9$); however, in the study of Lundgren, et al⁽⁸⁾ the relationship was relatively weak (Pearson's $r=0.48$). The relationship was strong when score 1 was removed in Lundgren's method (Pearson's $r=0.93$).⁽⁶⁾ In the case of Winkel's method, intra-rater reliability was almost perfect ($\kappa=0.871$).⁽⁹⁹⁾ Inter-rater reliabilities were investigated in two studies. In the study of Hitz Lindenmüller, et al⁽¹⁰⁰⁾ it was substantial ($\kappa=0.654$). The other study performed by Kim, et al⁽³⁾ applied Winkel's method in two different ways: inspection of the real tongue and of the tongue images. The results of the two methods were similar: Pearson's r was 0.762 for real tongue and 0.773 for a tongue image. Ileri Keceli, et al⁽¹⁰²⁾ reported that the test-retest reliability of Winkel's method was high (Pearson's $r=0.988$). Moreover, the inter-rater reliability of the dichotomized Winkel's method was 0.696 in Pearson's r .⁽³⁾ In the case of the computerized method, the repeatability of the distribution ratio of TC using the TIAS was almost perfect (weighted $\kappa=0.851$) when analyzing tongue images acquired twice at 30-min intervals from 60 functional dyspepsia patients.⁽²⁵⁾

Correlations between Weighing TC Method and the Indirect Methods

In order to assess the accuracy of each method, the relationships between the weighing TC method, which is the most direct method, and the various indirect methods were investigated. Spearman's rank correlation coefficients between the wet weight of TC and scores in some kinds of specificative method ranged from 0.68 to 0.8 (Miyazaki's method 0.68; Winkel's method 0.8; dichotomized Winkel's method 0.71).⁽⁴¹⁾ In the relationship between wet weight of TC and scores from Lundgren's method, there was a considerable difference according to the raters (Pearson's r 0.2–0.6), but in the dichotomized Lundgren's method, the difference decreased (Pearson's r 0.69–0.75).⁽⁸⁾ In the computerized method, the correlation between wet weight of TC and its distribution ratio using the TIAS was not strong (Pearson's $r = 0.442$).⁽²⁵⁾ There was a relatively strong correlation between the wet weight of TC and its dry weight (Spearman's $r=0.81$).⁽¹¹²⁾

Frequency of Use

The numbers of the articles involving the evaluation methods for ATC and the frequency of use of the methods are presented in Figure 3 and Figure 4, respectively. The number of articles has increased, with no article in 1985-1991 to 3 articles in 1992-1997, 12 in 1998–2003, 26 in 2004–2009, and 72 in 2010–2015. The evaluation methods for ATC were used a total of 134 times in the 113

selected articles; 15 articles used 2–5 kinds of method for evaluating ATC, whereas 98 articles used only one method. Through this investigation, we found that the specificative method was used widely (90 times, 67.1%) and the frequency and proportion of studies reporting this method, especially for the dividing type, has increased constantly. On the other hand, the intuitive method (17 times, 12.7%) and the weighing TC method (16 times, 11.9%) were used relatively rarely, and their proportions tended to decrease. The computerized method was considered to be still in the tentative stage because it was used just 11 times (8.2%), but its frequency of use has increased rapidly since 2010.

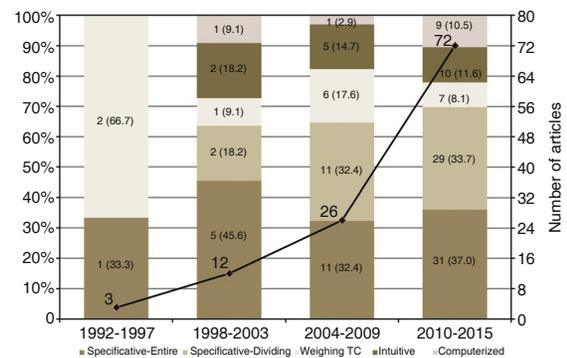


Figure 3. Number of Articles Involving the Evaluation Methods for the Amount of Tongue Coating (ATC) and the Proportion of Each Method as Time Series [Times (%)]

Notes: The line graph presents the change in the number of articles including the evaluation methods for ATC. The bar graph shows the proportion of each method as time progressed. The number of articles has increased constantly. The specificative method was used widely and the number of studies reporting this method has increased constantly. In contrast, the intuitive, computerized, and weighing TC methods were used relatively rarely.

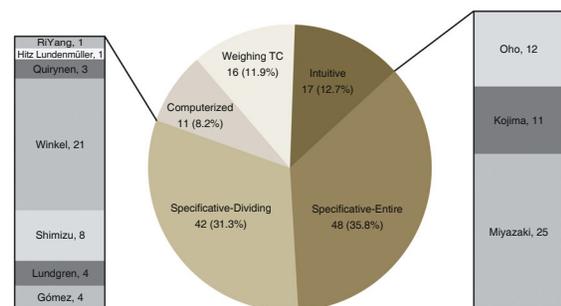


Figure 4. Total Proportion of Each Method and the Frequency of the Specificative Method [Times (%)]

The frequency of use of each subtype in the specificative method was also counted. The entire type of the specificative method was used 48 times (35.8%). Miyazaki's method was the most frequently used, followed by Oho's and Kojima's method. The dividing type was used a total of 42 times (31.3%). Winkel's method was most frequently used of the dividing type and

the other methods used frequently were in the following order: Shimizu's, Gómez's and Lundgren's, Quirynen's, and Hitz Lindenmüller's and Riyang's method (Figure 4).

DISCUSSION

The reliability in the intuitive and specificative methods showed varying levels, according to the study. In the intuitive method, assessors would judge the ATC depending solely on their intuition and experience, so a lack of consistency within raters is unavoidable. In the specificative method, as well as the intuitive method, there was still a problem with inter-rater reliability. In the case of Miyazaki's method, the inter-rater reliability was evaluated by three different teams: two teams^(34,47) reported that its inter-rater reliability was almost perfect, whereas the other team⁽⁸⁰⁾ showed that it was just fair. When the inter-rater reliability of Lundgren's method was also evaluated twice by almost the same team, it was revealed that there was a remarkable difference between the reliabilities found in the two evaluations.^(8,79) On the contrary, the reliability of the computerized method was relatively higher than that of the other methods. The higher reliability of TIAS was also reported by other studies.^(118,119) The results implied that the intuitive and specificative methods could be affected by the subjectivity of the assessors even if well-defined criteria were provided. To overcome the subjectivity, Han, et al⁽¹²⁰⁾ emphasized the importance of the calibration training for clinical practitioners. As our investigation, the inter-reliabilities were high when the assessors underwent calibration training prior to the experiment.^(8,47,75,79)

Nevertheless, the reliability tended to increase as the criteria were simplified in the specificative method. The reliability of Lundgren's method,⁽⁷⁹⁾ a simplified version of Gómez's with regard to the number of tongue sections, was better than that of Gómez's,⁽⁷⁵⁾ and the reliability of the dichotomized Lundgren's method was better than its original. This tendency was maintained also in the intuitive method; when we compared the reliability of the dichotomous scale type^(10,12) and the several grades type,^(25,94) the dichotomous method was more reliable. Considering that as the number of grades increases, the vague boundary areas where the raters are reluctant to judge would also increase, this tendency could be understood. However, the reduction in the number of grades could deplete the accuracy and could make researchers lose plenty of information for ATC. In particular, clinical data obtained through a simplistic method may not be sufficient, because the ATC needs to be minutely and elaborately evaluated for pattern identification in TKM and CM. Therefore, it is necessary to take into account both the reliability and the resolution of the method.

In order to find out the accuracy of each method, we investigated the relationships between weighing TC and the other methods. The accuracies of the specificative method were similar to one another in our investigation, whereas the ATCs evaluated by a TIAS were less accurate than those from the specificative method. Even though the criteria were given, the raters would assess the ATC in a comprehensive way and it could improve the accuracy. On the other hand, the computer used only the distribution rate of the TC; therefore, in order to improve the accuracy of computerized method, further various parameters for the ATC should be studied, such as the color and humidity of TC.

Frequency of use of the intuitive method was not high, although it is very simple to use. It seemed that its resolution is too low to evaluate ATC elaborately. Weighing TC method also showed low prevalence, although it is the most direct method. We suppose that it can make the patients feel discomfort such as nausea or vomiting, making it inconvenient to apply in the clinics. The specificative method, especially Miyazaki's method, was used most frequently. We thought that the high prevalence of the specificative method would be a result of its substantial accuracy and convenience, and its high prevalence would continue in clinical practice. The computerized method was rarely used despite its high reliability. This could result from the unclear validity, unfamiliarity, and low penetration rate of the commercialized TIAS.

This review could not cover all of the evaluation methods for ATC. We only reviewed the articles written in English. Moreover, we focused on ATC, so other important factors in tongue diagnosis, such as the color, shape, texture, crust and humidity of the tongue, were excluded.

In summary, the specificative method was widely used for evaluating ATC, but its reliability has not yet been fully demonstrated. We recommend that the calibration training for the assessors should be performed to improve its reliability. The computerized method was not widely used because of its unclear validity and unfamiliarity. However, if novel algorithms are developed to evaluate ATC more accurately, the computerized method can help the practitioners evaluate ATC with confidence.

Conflict of Interest

The authors declare that there is no conflict of interest regarding the publication of this manuscript.

Author Contributions

Nam DH designed the study, and searched the articles. Kim SR also searched the articles, and wrote the manuscript.

Electronic Supplementary Material: Supplementary materials (Appendixes) are available in the online version of this article at <https://doi.org/10.1007/s11655-018-2552-y>

REFERENCES

- Lee B, Park Y, Kim T. Diagnosis in Korean traditional medicine. Seoul: Seongbo Publishing Inc.; 2000:70-91.
- Qi W, Zhang M, Wang H, et al. Research on the relationship between thick greasy tongue fur formation and vascular endothelial cell permeability with the protein expression of zonula occludens-1. *Chin J Integr Med* 2011;17:510-516.
- Kim J, Jung Y, Park K, et al. A digital tongue imaging system for tongue coating evaluation in patients with oral malodour. *Oral Dis* 2009;15:565-569.
- Gao L, Liu P, Song J, et al. Study on the correlation of tongue manifestation with fibrinogen and neutrophil in acute cerebral infarction patients. *Chin J Integr Med* 2012;18:942-945.
- Yaegaki K, Sanada K. Volatile sulfur compounds in mouth air from clinically healthy subjects and patients with periodontal disease. *J Periodontol Res* 1992;27(4 Pt 1):233-238.
- Abe S, Ishihara K, Adachi M, et al. Tongue-coating as risk indicator for aspiration pneumonia in edentate elderly. *Arch Gerontol Geriatr* 2008;47:267-275.
- Danser MM, Gómez SM, Van der Weijden GA. Tongue coating and tongue brushing: a literature review. *Int J Dent Hyg* 2003;1:151-158.
- Lundgren T, Mobilia A, Hallström H, et al. Evaluation of tongue coating indices. *Oral Dis* 2007;13:177-180.
- Kuo YW, Yen M, Fetzer S, et al. Toothbrushing versus toothbrushing plus tongue cleaning in reducing halitosis and tongue coating: a systematic review and meta-analysis. *Nurs Res* 2013;62:422-429.
- Quirynen M, Mongardini C, Van Steenberghe D. The effect of a 1-stage full-mouth disinfection on oral malodor and microbial colonization of the tongue in periodontitis. A pilot study. *J Periodontol* 1998;69:374-382.
- Amir E, Shimonov R, Rosenberg M. Halitosis in children. *J Pediatr* 1999;134:338-343.
- Kim M, Cobbin D, Zaslawski C. Traditional Chinese medicine tongue inspection: An examination of the inter- and intrapractitioner reliability for specific tongue characteristics. *J Altern Complement Med* 2008;14:527-536.
- Matsuo K, Watanabe R, Kanamori D, et al. Associations between oral complications and days to death in palliative care patients. *Support Care Cancer* 2015;24:157-161.
- Morishita S, Watanabe Y, Ohara Y, et al. Factors associated with older adults' need for oral hygiene management by dental professionals. *Geriatr Gerontol Int* 2016;16:956-962.
- Gross A, Barnes GP, Lyon TC. Effects of tongue brushing on tongue coating and dental plaque scores. *J Dent Res* 1975;54:1236.
- Çiçek Y, Arabaci T, Canakçi CF. Evaluation of oral malodour in left- and right-handed individuals. *Laterality* 2010;15:317-326.
- Maeda M. Dermoscopic patterns of the filiform papillae of the tongue in patients with Sjögren's syndrome. *J Dermatol* 2006;33:96-102.
- Wu Z, Li M, Zhang Y, et al. Study on relationship between the thickness of tongue fur and the expressions of apoptosis-related genes of the tongue epithelial cells in patients with diseases of the digestive system. *J Tradit Chin Med* 2007;27:148-152.
- Takeshita T, Tomioka M, Shimazaki Y, et al. Microfloral characterization of the tongue coating and associated risk for pneumonia-related health problems in institutionalized older adults. *J Am Geriatr Soc* 2010;58:1050-1057.
- Su W, Xu ZY, Wang ZQ, et al. Objectified study on tongue images of patients with lung cancer of different syndromes. *Chin J Integr Med* 2011;17:272-276.
- Wang HH, Pan CH, Wu PP, et al. Alteration of the tongue manifestation reflects clinical outcomes of peptic ulcer disease. *J Altern Complement Med* 2012;18:1038-1044.
- Gao L, Liu P, Song JX, et al. A pilot study on the relationship between tongue manifestation and the degree of neurological impairment in patients with acute cerebral infarction. *Chin J Integr Med* 2013;19:149-152.
- Liu P, Gao L, Song JX, et al. A pilot study on the correlation of tongue manifestation with the site of cerebral infarction in patients with stroke. *Chin J Integr Med* 2014;20:823-828.
- Anastasi JK, Chang M, Quinn J, et al. Tongue inspection in TCM: observations in a study sample of patients living with HIV. *Med Acupunct* 2014;26:15-22.
- Kim J, Han GJ, Ko SJ, et al. Tongue diagnosis system for quantitative assessment of tongue coating in patients with functional dyspepsia: a clinical trial. *J Ethnopharmacol* 2014;155:709-713.
- Miyazaki H, Sakao S, Katoh Y, et al. Correlation between volatile sulphur compounds and certain oral health measurements in the general population. *J Periodontol* 1995;66:679-684.
- Morita M, Wang HL. Relationship between sulcular sulfide level and oral malodor in subjects with periodontal disease. *J Periodontol* 2001;72:79-84.
- Koshimune S, Awano S, Gohara K, et al. Low salivary flow and volatile sulfur compounds in mouth air. *Oral Surg Endodontol Oral Med Oral Pathol Oral Radiol* 2003;96:38-41.
- Quirynen M, Avontroodt P, Soers C, et al. Impact of tongue cleansers on microbial load and taste. *J Clin Periodontol* 2004;31:506-510.
- Quirynen M, Zhao H, Soers C, et al. The impact of periodontal therapy and the adjunctive effect of antiseptics on breath odor-related outcome variables: a double-blind randomized study. *J Periodontol* 2005;76:705-712.
- Liu XN, Shinada K, Chen XC, et al. Oral malodor-related parameters in the Chinese general population. *J Clin Periodontol* 2006;33:31-36.
- Nalcaci R, Sönmez IS. Evaluation of oral malodor in children. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 2008;106:384-388.
- Nalcaci R, Baran I. Factors associated with self-reported halitosis (SRH) and perceived taste disturbance (PTD) in elderly. *Arch Gerontol Geriatr* 2008;46:307-316.
- Callil C, Liberato F, Pereira A, et al. The relationship between volatile sulphur compounds, tongue coating and periodontal disease. *Int J Dent Hyg* 2009;7:251-255.
- Kikutani T, Tamura F, Nishiwaki K, et al. The degree of tongue-coating reflects lingual motor function in the elderly. *Gerodontology* 2009;26:291-296.
- Schaefer I, Braumann B. Halitosis, oral health and quality of life during treatment with Invisalign® and the effect of a low-dose chlorhexidine solution. *J Orofac Orthop* 2010;71:430-441.
- Kamaraj DR, Laxman VK, Bhushan KS, et al. Detection of odoriferous subgingival and tongue microbiota in diabetic and nondiabetic patients with oral malodor using polymerase chain reaction. *Indian J Dent Res* 2011;22:260-265.
- Tashiro K, Katoh T, Yoshinari N, et al. The short-term effects of various oral care methods in dependent elderly: comparison between toothbrushing, tongue cleaning with sponge brush and wiping on oral

- mucous membrane by chlorhexidine. *Gerodontology* 2012;29:e870-882.
39. Zurfluh MA, van Waes HJM, Filippi A. The influence of fixed orthodontic appliances on halitosis. *Schweiz Monatsschr Zahnmed* 2013;123:1064-1075.
 40. Evirgen Ş, Kamburoşlu K. Effects of tongue coating and oral health on halitosis among dental students. *Oral Health Prev Dent* 2013;11:169-173.
 41. Van Tornout M, Dadamio J, Coucke W, et al. Tongue coating: related factors. *J Clin Periodontol* 2013;40:180-185.
 42. Vancauwenberghe F, Dadamio J, Laleman I, et al. The role of *Solobacterium moorei* in oral malodour. *J Breath Res* 2013;7:046006.
 43. Ueno M, Takeuchi S, Takehara S, et al. Saliva viscosity as a potential risk factor for oral malodor. *Acta Odontol Scand* 2014;72:1005-1009.
 44. Kamaraj DR, Bhushan KS, Laxman VK. An evaluation of microbial profile in halitosis with tongue coating using PCR (polymerase chain reaction)- a clinical and microbiological study. *J Clin Diagnostic Res* 2014;8:263-267.
 45. Gulsahi A, Evirgen Ş, Öztaş B, et al. Volatile sulphur compound levels and related factors in patients with chronic renal failure. *J Clin Periodontol* 2014;41:814-819.
 46. Rani H, Ueno M, Zaitsu T, et al. Original article factors associated with clinical and perceived oral malodor among dental students. *J Med Dent Sci* 2015;62:33-41.
 47. Yokoi A, Maruyama T, Yamanaka R, et al. Relationship between acetaldehyde concentration in mouth air and tongue coating volume. *J Appl Oral Sci* 2015;23:64-70.
 48. Bornstein MM, Stocker BL, Seemann R, et al. Prevalence of halitosis in young male adults: a study in Swiss army recruits comparing self-reported and clinical data. *J Periodontol* 2009;80:24-31.
 49. Kislig K, Wilder-Smith CH, Bornstein MM, et al. Halitosis and tongue coating in patients with erosive gastroesophageal reflux disease versus nonerosive gastroesophageal reflux disease. *Clin Oral Investig* 2013;17:159-165.
 50. Ueno M, Shinada K, Zaitsu T, et al. Effects of an oral health education program targeting oral malodor prevention in Japanese senior high school students. *Acta Odontol Scand* 2012;70:426-431.
 51. Kojima K. Clinical studies on the coated tongue. *Japan J Oral Maxillofac Surg* 1985;1659-1679.
 52. Hinode D, Fukui M, Yokoyama N, et al. Relationship between tongue coating and secretory-immunoglobulin A level in saliva obtained from patients complaining of oral malodor. *J Clin Periodontol* 2003;30:1017-1023.
 53. Yonezawa H, Takasaki K, Teraoka K, et al. Effects of tongue and oral mucosa cleaning on oral *Candida* species and production of volatile sulfur compounds in the elderly in a nursing home. *J Med Dent Sci* 2003;50:1-8.
 54. Yasukawa T, Ohmori M, Sato S. The relationship between physiologic halitosis and periodontopathic bacteria of the tongue and gingival sulcus. *Odontology* 2010;98:44-51.
 55. Yoneda M, Masuo Y, Suzuki N, et al. Relationship between the β -galactosidase activity in saliva and parameters associated with oral malodor. *J Breath Res* 2010;4:017108.
 56. Takeuchi H, Machigashira M, Yamashita D, et al. The association of periodontal disease with oral malodour in a Japanese population. *Oral Dis* 2010;16:702-706.
 57. Masuo Y, Suzuki N, Yoneda M, et al. Salivary β -galactosidase activity affects physiological oral malodour. *Arch Oral Biol* 2012;57:87-93.
 58. Ramesh G, Nagarajappa R, Madhusudan A, et al. Estimation of salivary and tongue coating pH on chewing household herbal leaves: A randomized controlled trial. *Anc Sci Life* 2012;32:69-75.
 59. Takeshita T, Suzuki N, Nakano Y, et al. Relationship between oral malodor and the global composition of indigenous bacterial populations in saliva. *Appl Environ Microbiol* 2010;76:2806-2814.
 60. Petrini M, Costacurta M, Ferrante M, et al. Association between the organoleptic scores, oral condition and salivary β -galactosidases in children affected by halitosis. *Int J Dent Hyg* 2014;12:213-218.
 61. Kaygisiz E, Uzuner FD, Yuksel S, et al. Effects of self-ligating and conventional brackets on halitosis and periodontal conditions. *Angle Orthod* 2015;85:468-473.
 62. Migliario M, Rimondini L. Oral and non oral diseases and conditions associated with bad breath. *Minerva Stomatol* 2011;60:105-115.
 63. Oho T, Yoshida Y, Shimazaki Y, et al. Characteristics of patients complaining of halitosis and the usefulness of gas chromatography for diagnosing halitosis. *Oral Surg, Oral Med Oral Pathol Oral Radiol Endodontol* 2001;91:531-534.
 64. Nonaka A, Tanaka M, Anguri H, et al. Clinical assessment of oral malodor intensity expressed as absolute value using an electronic nose. *Oral Dis* 2005;11 Suppl 1:35-36.
 65. Romano F, Pigella E, Guzzi N, et al. Patients' self-assessment of oral malodour and its relationship with organoleptic scores and oral conditions. *Int J Dent Hyg* 2010;8:41-46.
 66. Tanaka M, Toe M, Nagata H, et al. Effect of eucalyptus-extract chewing gum on oral malodor: a double-masked, randomized trial. *J Periodontol* 2010;81:1564-1571.
 67. Dadamio J, Laleman I, De Geest S, et al. Usefulness of a new malodour-compound detection portable device in oral malodour diagnosis. *J Breath Res* 2013;7:046005.
 68. Aimetti M, Perotto S, Castiglione A, et al. Prevalence estimation of halitosis and its association with oral health-related parameters in an adult population of a city in North Italy. *J Clin Periodontol* 2015;42:1105-1114.
 69. Hakuta C, Mori C, Ueno M, et al. Evaluation of an oral function promotion programme for the independent elderly in Japan. *Gerodontology* 2009;26:250-258.
 70. Ohnuki M, Ueno M, Zaitsu T, et al. Taste hyposensitivity in Japanese schoolchildren. *BMC Oral Health* 2014;14:36.
 71. Da Cruz MK, De Moraes TMN, Trevisani DM. Clinical assessment of the oral cavity of patients hospitalized in an intensive care unit of an emergency hospital. *Rev Bras Ter Intensiva* 2014;26:379-383.
 72. Samnieng P, Ueno M, Shinada K, et al. Daily variation of oral malodour and related factors in community-dwelling elderly Thai. *Gerodontology* 2012;29:e964-971.
 73. Takeuchi S, Ueno M, Takehara S, et al. The relationship between turbidity of mouth-rinsed water and oral health status. *Acta Odontol Scand* 2013;71:183-188.
 74. Conti S, dos Santos SSF, Koga-Ito CY, et al. Enterobacteriaceae and pseudomonadaceae on the dorsum of the human tongue. *J Appl Oral Sci* 2009;17:375-380.
 75. Gómez SM, Danser MM, Sipoş PM, et al. Tongue coating and salivary bacterial counts in healthy/gingivitis subjects and periodontitis patients. *J Clin Periodontol* 2001;28:970-978.
 76. Shinada K, Ueno M, Konishi C, et al. Effects of a mouthwash with chlorine dioxide on oral malodor and salivary bacteria: a randomized placebo-controlled 7-day trial. *Trials* 2010;11:14.
 77. Pham TAV, Ueno M, Zaitsu T, et al. Clinical trial of oral malodor treatment in patients with periodontal diseases. *J Periodontol Res* 2011;46:722-729.
 78. Van der Sluijs E, Slot DE, Bakker E, et al. The effect of water on morning bad breath: a randomized clinical trial. *Int J Dent Hyg* 2016;14:124-134.
 79. Chérel F, Mobilia A, Lundgren T, et al. Rate of reformation of tongue coatings in young adults. *Int J Dent Hyg* 2008;6:371-375.
 80. Shimizu T, Ueda T, Sakurai K. New method for evaluation of tongue-coating status. *J Oral Rehabil* 2007;34:442-447.

81. Ryu M, Ueda T, Saito T, et al. Oral environmental factors affecting number of microbes in saliva of complete denture wearers. *J Oral Rehabil* 2010;37:194-201.
82. Dupim Souza AC, Franco CF, Pataro AL, et al. Halitosis in obese patients and those undergoing bariatric surgery. *Surg Obes Relat Dis* 2013;9:315-321.
83. Han DH, Lee SM, Lee JG, et al. Association between viral hepatitis B infection and halitosis. *Acta Odontol Scand* 2014;72:274-282.
84. Hammad MM, Darwazah AM, Al-Waeli H, et al. Prevalence and awareness of halitosis in a sample of Jordanian population. *J Int Soc Prev Community Dent* 2014;4(Suppl 3):S178-S186.
85. Cunha KS, Rozza-de-Menezes RE, Luna EB, et al. High prevalence of hyposalivation in individuals with neurofibromatosis 1: a case-control study. *Orphanet J Rare Dis* 2015;10:8-11.
86. Kobayashi K, Ryu M, Izumi S, et al. Effect of oral cleaning using mouthwash and a mouth moisturizing gel on bacterial number and moisture level of the tongue surface of older adults requiring nursing care. *Geriatr Gerontol Int* 2017;17:116-121.
87. Ayo-Yusuf OA, Postma TC, van Wyk C. Clinical correlates of oral malodour in a population of patients attending a preventive clinic in Pretoria, South Africa. *SADJ* 2011;66:326, 328-331.
88. Winkel EG, Roldan S, Van Winkelhoff AJ, et al. Clinical effects of a new mouthrinse containing chlorhexidine, cetylpyridinium chloride and zinc-lactate on oral halitosis. A dual-center, double-blind placebo-controlled study. *J Clin Periodontol* 2003;30:300-306.
89. Roldán S, Winkel EG, Herrera D, et al. The effects of a new mouthrinse containing chlorhexidine, cetylpyridinium chloride and zinc lactate on the microflora of oral halitosis patients: a dual-centre, double-blind placebo-controlled study. *J Clin Periodontol* 2003;30:427-434.
90. Roldán S, Herrera D, Santa-Cruz I, et al. Comparative effects of different chlorhexidine mouth-rinse formulations on volatile sulphur compounds and salivary bacterial counts. *J Clin Periodontol* 2004;31:1128-1134.
91. Roldán S, Herrera D, O'Connor A, et al. A combined therapeutic approach to manage oral halitosis: a 3-month prospective case series. *J Periodontol* 2005;76:1025-1033.
92. Saad S, Greenman J. Tongue biofilm areal density and tongue coating index. *J Breath Res* 2008;2:017008.
93. Erovic Ademovski S, Lingström P, Winkel E, et al. Comparison of different treatment modalities for oral halitosis. *Acta Odontol Scand* 2012;70:224-233.
94. Kim J, Han GJ, Choi BH, et al. Development of differential criteria on tongue coating thickness in tongue diagnosis. *Complement Ther Med* 2012;20:316-322.
95. Apatzidou AD, Bakirtzoglou E, Vouros I, et al. Association between oral malodour and periodontal disease-related parameters in the general population. *Acta Odontol Scand* 2013;71:189-195.
96. Matsui M, Chosa N, Shimoyama Y, et al. Effects of tongue cleaning on bacterial flora in tongue coating and dental plaque: a crossover study. *BMC Oral Health* 2014;14:4.
97. Guentsch A, Pfister W, Cachovan G, et al. Oral prophylaxis and its effects on halitosis-associated and inflammatory parameters in patients with chronic periodontitis. *Int J Dent Hyg* 2013;12:199-207.
98. Alqumber MA, Arafa KA. Site-specific mouth rinsing can improve oral odor by altering bacterial counts. *Blind crossover clinical study. Saudi Med J* 2014;35:1412-1416.
99. Soares LG, Castagna L, Weyne SC, et al. Effectiveness of full- and partial-mouth disinfection on halitosis in periodontal patients. *J Oral Sci* 2015;57:1-6.
100. Hitz Lindenmüller I, Weiss P, Volken M, et al. Diagnostics of tongue coating using autofluorescence. *Swiss Dent J* 2015;125:1074.
101. Yadav SR. Inhibition of tongue coat and dental plaque formation by stabilized chlorine dioxide vs. chlorhexidine mouthrinse: a randomized, triple blinded study. *J Clin Diagn Res* 2015;69-74.
102. Ileri Keceli T, Gulmez D, Dolgun A, et al. The relationship between tongue brushing and halitosis in children: a randomized controlled trial. *Oral Dis* 2015;21:66-73.
103. Marchetti E, Tecco S, Santonico M, et al. Multi-sensor approach for the monitoring of halitosis treatment via *Lactobacillus brevis* (CD2)—containing Lozenges—a randomized, double-blind placebo-controlled clinical trial. *Sensors* 2015;15:19583-19596.
104. Bolepalli A, Munireddy C, Peruka S, et al. Determining the association between oral malodor and periodontal disease: A case control study. *J Int Soc Prev Community Dent* 2015;5:413-418.
105. Tsai CC, Chou HH, Wu TL, et al. The levels of volatile sulfur compounds in mouth air from patients with chronic periodontitis. *J Periodontol Res* 2008;43:186-193.
106. Lin RY, Yu HY, Qin JY, et al. Association between tongue coating thickness and clinical characteristics among idiopathic membranous nephropathy patients. *J Ethnopharmacol* 2015;171:1-6.
107. 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.
108. Lee ES, Yim HK, Lee HS, et al. Clinical assessment of oral malodor using autofluorescence of tongue coating. *Photodiagnosis Photodyn Ther* September 2015:1-7.
109. Nohno K, Yamaga T, Kaneko N, et al. Tablets containing a cysteine protease, actinidine, reduce oral malodor: a crossover study. *J Breath Res* 2012;6:017107.
110. Yaegaki K, Sanada K. Biochemical and clinical factors influencing oral malodor in periodontal patients. *J Periodontol* 1992;63:783-789.
111. Lee CH, Kho HS, Chung SC, et al. The relationship between volatile sulfur compounds and major halitosis-inducing factors. *J Periodontol* 2003;74:32-37.
112. Monteiro-Amado F, Chinellato LEM, de Rezende MLR. Evaluation of oral and nasal halitosis parameters in patients with repaired cleft lip and/or palate. *Oral Surg, Oral Med Oral Pathol Oral Radiol Endodontology* 2005;100:682-687.
113. Peruzzo DC, Salvador SL, Sallum AW, et al. Flavoring agents present in a dentifrice can modify volatile sulphur compounds (VSCs) formation in morning bad breath. *Braz Oral Res* 2008;22:252-257.
114. Peruzzo DC, Salvador SL, Sallum AW, et al. Effects of sodium lauryl sulphate (SLS), present in dentifrice, on volatile sulphur compound (VSC) formation in morning bad breath. *J Int Acad Periodontol* 2008;10:130-136.
115. Albuquerque DF, Tolentino EDS, Amado FM, et al. Evaluation of halitosis and sialometry in patients submitted to head and neck radiotherapy. *Med Oral Patol Oral Cir Bucal* 2010;15:850-854.
116. Fleiss JL. *Statistical methods for rates and proportions*. 3rd ed. Hoboken: John Wiley and Sons; 2003:598-621.
117. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159-174.
118. Lo LC, Chen YF, Chen WJ, et al. The study on the agreement between automatic tongue diagnosis system and traditional Chinese medicine practitioners. *Evidence-Based Complement Altern Med* 2012;2012:1-9.
119. Jung CJ, Kim KH, Jeon YJ, et al. Improving color and shape repeatability of tongue images for diagnosis by using feedback gridlines. *Eur J Integr Med* 2014;6:328-336.
120. Han G, Park J, Ko S, et al. Study on tongue coating patterns of the xerostomia in the Elderly Patients. *J Soc Korean Med Diagnostics* 2013;17:189-202.