# A Randomized Parallel Study to Assess the Effect of Three Tongue Cleaning Modalities on Oral Malodor

**Yiming Li, DDS, MSD, PhD  Sean Lee, DDS  Joni Stephens, RDH, MEd  Wu Zhang, MD  Montry Suprono, DDS, MSD**

**Loma Linda University School of Dentistry**
**Center for Dental Research**
**Loma Linda, CA, USA**

**Anthony Mwatha, MS  Marilyn Ward, DDS  Farah Mirza, MS, BDS**

**Philips Oral Healthcare**
**Bothell, WA, USA**

**Abstract**

- **Objective:** The objective of this study was to compare the effects of three tongue hygiene regimens on oral malodor.
- **Methods:** This was a single-center, randomized, parallel design study with three treatment groups. Subjects were randomly assigned to perform tongue hygiene with either the Philips Sonicare TongueCare+ BreathRx regimen (STC), Listerine Cool Mint antiseptic rinse (LCM), or tongue brushing with an ADA reference manual toothbrush (MTB). Tooth brushing was standardized for all subjects during the study period, and no other oral or breath hygiene measures were allowed. Eligible subjects met the following criteria: aged 18–70 years, in good general and oral health, non-smoker, with an organoleptic score between 2.7 and 4.5 following a 12–18 hour oral hygiene abstention period. Subjects who had oral appliances or who had periodontal disease or excessive recession were not eligible. The primary endpoint analysis was to evaluate oral malodor based on an organoleptic (OL) score. Additional surrogate measures for oral malodor included quantification of oral hydrogen sulfide ($H_2S$) level and counts of oral bacteria in secondary analyses. At Day 1, all three malodor endpoints were assessed prior to product use, immediately after use, and four and eight hours after use. Subjects were then provided with instructions on product use at home. Subjects returned to the clinic on Day 8 and the assessments for malodor were repeated for each of the three endpoints, *i.e.*, prior to in-clinic use of the products, immediately after use, and four and eight hours after use.
- **Results:** One hundred sixty-eight (168) subjects were randomized to three groups, with 56 per treatment group. Of these, 165 completed all study visits. Randomized subjects were comparable for baseline characteristics (OL score, age, race, and ethnicity). Overall, oral malodor based on the organoleptic score decreased for all treatment groups at all timepoints. For the primary endpoint, reduction of OL score eight hours following a single product use, the STC regimen reduced malodor per OL score by 46.67% (SE = 2.28%), the LCM value was 22.83% (SE = 2.29%), and MTB was 26.19% (SE = 2.29%). The pair-wise comparisons between STC and each of the treatment groups were statistically significant (p-values < 0.0001). Statistically significant differences were also observed between STC and both LCM and MTB groups in pair-wise comparisons at Day 8 (p-values < 0.0001).
- **Conclusion:** Reductions in malodor were evident following a single use of each product, and also following a seven-day repeat use period. The STC regimen, however, was statistically significantly superior to both LCM and MTB at improving malodor eight hours following the first use. Statistically significant differences in OL scores were sustained between STC and LCM, and STC and MTB at each efficacy timepoint following the seven-day home use period.

**(J Clin Dent 2019;30(Spec Iss A)A30–38)**

## Introduction

Oral malodor is attributed to gaseous metabolites from bacteria in the oral cavity that stream into exhaled breath.[1,2] The tongue, in particular the posterior dorsum of the tongue, is noted as one of the major sites with a high concentration of microbes coating the mucosal surface. This site is implicated as the dominant site of malodor production.[3,4] The metabolites of tongue bacteria can produce volatile sulfur compounds (VSCs), including hydrogen sulfide ($H_2S$), dimethylsulfide [$(CH_3)_2S$], and methylmercaptan ($CH_3SH$), which are the main odiferous culprits contributing to halitosis. Oral hygiene, inflammation, and infection[5-7] can affect the character of breath.

The management strategies for reducing oral malodor span a wide range of available medicaments and tools. Oral rinses are commonly used,[8-10] and generally include an antimicrobial ingredient. Mechanical tongue-brushing or tongue-scraping devices[11,12] are also employed to mechanically reduce the overall quantity of bacteria coating the tongue, much like a toothbrush is used to eliminate the surface plaque that coats teeth. In either case, the treatment is targeted to reduce the causative bacteria residing on the tongue, thus reducing the resultant concentration of VSCs as a means to help reduce and control oral malodor.

This study was a randomized and controlled clinical trial initiated to explore whether a two-pronged approach to malodor management (medicament plus tongue cleaning) exhibited any advantages over either rinsing with a medicament alone or to mechanical tongue cleaning alone. In particular, the study evaluated effects on the organoleptic character of breath up to eight hours following a single

use of the assigned product. In this study, only subjects with an existing level of malodor were included.

Surrogate measures for malodor were also included for exploratory purposes. These included assessments of hydrogen sulfide ($H_2S$), as well as the quantification of both aerobic and anaerobic bacteria following tongue biofilm sampling and culture.

## Materials and Methods

### *Study Design and Objectives*

This study was reviewed and approved by the Institutional Review Board of Loma Linda University. All screened and enrolled subjects provided informed consent. The study was conducted according to the ICH Guideline for Good Clinical Practices (GCPs) and the standards of ISO 14155. There were two clinical evaluation days (Day 1 and Day 8), each with four evaluations (prior to in-clinic use of the products, immediately after use, and four and eight hours after use). Table I provides an outline of study visits and the procedures that were performed at each visit. This was a three-arm, single-blind, repeat-measure, parallel-design clinical trial.

**Table I**
Study Visit Timeline and Procedures

| Day 1 | | | | | Day 8 | | | | |
|---|---|---|---|---|---|---|---|---|---|
| • Informed Consent<br>• Medical Dental History<br>• Oral Exam<br>• Randomization<br>• Dispense, Instruct, Use Assigned Product<br>• Provide Compliance Instructions and Diary | | | | | • Medical Dental History Update<br>• Collect Compliance Diary<br>• Oral Exam<br>• Product Use<br>• Collect Products<br>• Dismiss | | | | |
| | Pre | Post | 4 Hours | 8 Hours | | Pre | Post | 4 Hours | 8 Hours |
| OL | X | X | X | X | OL | X | X | X | X |
| $H_2S$ | X | X | X | X | $H_2S$ | X | X | X | X |
| Micro | X | X | X | X | Micro | X | X | X | X |

The primary objective was to compare the reduction in organoleptic scores between three oral malodor treatments eight hours after a single use. Secondary objectives included organoleptic score comparisons at the following timepoints: immediately and four hours following a single use, and then following a one-week period of daily home use, after which organoleptic (OL) scores were taken again in the clinic before, immediately, four, and eight hours following product use.

Similar timepoint comparisons were made between products for the other study surrogate efficacy measures, $H_2S$, and tongue microbial count (total bacterial load, CFU/cm$^2$). Safety was also assessed via intraoral examination and per subject report.

### *Study Subjects*

Eligible subjects were male and female adults 18–70 years of age, able to provide informed consent, available to attend study visits, and comply with study procedures. Subjects were non-smokers (defined as use of < 100 cigarettes in their lifetime) with an organoleptic score of 2.7–4.5 following a 12–18 hour oral hygiene abstention period. (Note: the OL score was an average based on the assessment of the three independent judges.) Subjects were not eligible in the event of pregnancy or nursing, a diagnosis of xerostomia, periodontal disease or a dental condition requiring care, Type II diabetes, a gagging reflex that precluded tongue-cleaning, usage of medications known to alter oral flora within one month of study, or the presence of orthodontic brackets or other intra-oral hardware or piercing.

Subjects were required to abstain from the use of any other oral and breath-hygiene products or devices, other than those dispensed for the study. The use of antibiotics or antimicrobials (other than tongue spray or rinse, if assigned) was also prohibited. In the event that a subject required dental care outside the scope of the study, she/he was discontinued.

Prior to each clinical evaluation day (Day 1 and Day 8), subjects were to observe a 12-hour abstention period from alcohol consumption. Subjects also abstained from the application of scented cosmetics, and withheld food and fluid consumption, other than clear liquids, the midnight prior. On study visit days, subjects were provided a standardized meal that did not include foods known to exacerbate oral malodor.

### *Treatment Groups*

Study subjects were randomized to one of the following three tongue-cleaning regimen:

- Tongue brushing with Philips Sonicare TongueCare+ tongue brush used on the Philips Sonicare EasyClean toothbrush handle in Clean mode, with TongueCare+ antimicrobial tongue spray (STC), 20 seconds x 3 (Philips, Bothell, WA, USA);
- Full-mouth rinse with 20 ml Listerine Cool Mint Antiseptic Rinse (Johnson & Johnson, New Brunswick, NJ, USA) for 30 seconds (LCM); or
- Tongue brushing with an ADA reference manual toothbrush (MTB).

After the Day 1 visit procedures were complete, all subjects were provided a standardized at-home tooth brushing regimen. This consisted of the use of a Philips Sonicare EasyClean power toothbrush handle and ProResults brush head in Clean mode, twice daily. Dentifrice was also standardized, with all subjects using Crest® Cool Mint Gel (Procter & Gamble, Cincinnati, OH, USA) for each brushing encounter. Tongue cleaning was performed once daily, in the morning, following tooth brushing.

### *Randomization, Controls to Minimize Bias, and Data Capture*

Randomization was performed by a designated member of the study staff who did not perform any efficacy assessments. Subjects were allocated to a treatment group according to a randomization schedule that was provided to the study site by the sponsor. Approximately equal numbers of subjects, of each gender, were randomized to each treatment group.

In order to minimize bias, the judges performing organoleptic evaluations completed a calibration exercise. This session was conducted with 12 subjects; the intraclass correlation (coefficient [ICC]) for the three judges was 0.901 (95% confidence interval: 0.736, 0.969). Each OL evaluator was blinded to the treatment assignment of each subject, and to the assessments of his/her OL peers. The laboratory personnel performing the microbial counts were also blinded to each subject's treatment assignment.

Study data were collected on a web-based electronic data capture (EDC) system. The system utilized programmed logic and edit-check functions. Access to the EDC system was based on the role of the user (to maintain the study blind), and was protected by log-in identification and password. Source document forms were used by the study site, where necessary. Study staff performed data-quality checks to ensure accuracy of reporting.

*Efficacy and Safety Measures*

**Organoleptic Assessment.** Three experienced organoleptic judges were assigned to perform each OL assessment for the duration of the study. All subjects underwent OL assessment, at each timepoint, by each of the three judges. All organoleptic assessments were performed in an examination operatory that preserved the study-blind. A small glass tube was inserted into an aperture in a wall that separated each subject and judge, as well as between the OL judges.

Following product use, and at the assigned time interval, the subject was asked to close his/her mouth for two minutes. Thereafter, a signal prompted the subject to exhale gently through the glass tube. The judge then performed the organoleptic assessment according to the following scale:[13] 0 = odor cannot be detected; 1 = questionable malodor, barely detectable; 2 = slight malodor, exceeds the threshold of malodor recognition; 3 = malodor is definitely detected; 4 = strong malodor; and 5 = very strong malodor. Subjects repeated this procedure three times, once for each OL judge, and they were instructed to keep their mouth closed for two minutes before moving to the next OL judge. The three OL assessments were recorded per subject, and were then averaged.

**Hydrogen Sulfide Assessment.** Each $H_2S$ assessment was performed with the OralChroma Gas Chromatography device (Nissha Co., Ltd., Schaumburg, IL, USA). This device measures three volatile sulfur compounds. For this study, only $H_2S$ outcomes were collected and reported in parts per billion (ppb).

Oral gas samples were taken using a sterile single-use 1 mL syringe. Subjects inserted the syringe into their mouth with lips closed tightly around the syringe. Subjects were instructed to breathe through their nose for one minute, after which the subject pulled the syringe piston to the end of the syringe, filling the lumen of the syringe with a breath sample. This was released back into the oral cavity and the procedure was repeated, filling the syringe with a second breath sample. This sample was injected into the OralChroma device. The device displayed the results, which were recorded on the study Case Report Form.

**Tongue Bacterial Collection and Analysis.** The tongue sampling method was based on previously published methods.[14,15] A manual toothbrush (Shaha 5 toothbrush, abcOralCare, Cupertino, CA, USA) was used to collect tongue samples. Each toothbrush was immersed in 70% ethanol for 30 seconds and dried overnight in a sanitized laminar flow hood. At the time of sampling, the brush head was placed on the dorsum of the subject's tongue, 5 cm from the tip, with all bristles in contact with the tongue surface. The brush head was then moved in five gentle oscillations, without bristle movement across the tongue. The brush head was removed and then immediately soaked in a 15 mL sterile centrifuge tube containing ¼-strength 5 mL Ringer's solution. Each tube was labeled with the subject's assigned study ID number.

Each sample was processed within two hours of collection, and kept on ice, or in a 4°C refrigerator, until processing. Samples were vortexed for 30 seconds. For each sample, a 100 µL dilution was plated using an L-shaped rod to evenly spread the inoculum on the surface of an agar plate. Samples were plated in duplicates using a non-selective FAA agar plate (FAA + 7% [v/v] defibrinated horse blood) for aerobes, and a selective FAA agar plate for anaerobes (FAA + 7% [v/v] defibrinated horse blood + vancomycin 2.5 mg/L). Plates were incubated for three days at 37°C for aerobic culture, and seven days at 37°C for anaerobes. Colonies were counted by a blinded laboratory staffer, and recorded in CFU/cm$^2$.

*Safety*

Subject safety was assessed by intraoral examination at each study visit, and by subjects' diary report of any adverse experiences occurring at home during the study period.

*Statistical Methods*

**Sample Size Determination.** In a previous pilot study,[16] the mean reduction in organoleptic score at six hours post brushing, in subjects using the STC regimen compared to subjects using water only, was 1.6 (SE = 0.13). When compared to subjects who used BreathRx only, the mean reduction in organoleptic score was 1.1 (SE = 0.124).

For the current study, a sample size of 50 subjects per treatment group would provide approximately 80% power to detect a 1.0 difference in the mean OL score between the STC, LCM, and MTB treatment groups, assuming a common standard deviation of 1.5, with a two-sided independent sample t-test with a Dunnett's adjustment for multiple testing (*i.e.*, alpha equal to 0.027).

The remaining efficacy endpoints, $H_2S$ and bacterial counts, were included with no prior pilot study outcomes. As a result, statistical comparisons in the current study were exploratory in nature.

**Demographics.** Standard subject demographics and baseline characteristics were summarized for all randomized subjects, and for modified Intent-to-Treat (mITT) subject populations. For continuous characteristics, means were compared using one-way analysis of variance (ANOVA). The incidence of the categorical variables was compared using the Chi-square test.

**Primary Efficacy Analysis.** The primary efficacy measure for this study was the OL score after eight hours of product use, based on assessments provided by three independent, blinded judges. For each subject the organoleptic score was a value obtained by averaging the scores of each of the three independent judges. The primary analysis was performed on a mITT basis; that is, including all subjects with a baseline (prior to the single use of products) and an eight-hour efficacy evaluation. The following hypotheses were evaluated:

- Null Hypothesis $H_o$: No difference among the three treatment groups; and
- Alternative Hypothesis $H_a$: At least two of the treatment groups differ.

The analysis was implemented using ANOVA modelling with overall comparisons between the three treatment groups performed using an F-test. If the overall F-test was significant then pairwise differences between STC and each of the two-comparator groups (LCM and MTB) were performed using contrast statements (SAS PROC MIXED), with Dunnett's procedure used to control for multiple comparisons. For the eight-hour outcome, the ANOVA model included the randomized treatment group and baseline OL score as predictor variables. Similar models were constructed to evaluate the four-hour and immediately after treatment timepoints.

**Secondary Efficacy and Safety Analysis.** Secondary efficacy analyses evaluated the change in OL score at Day 8 following an in-clinic use of the three study products. Similar to Day 1 visits, Day 8 assessments were performed after a 12–18 hour oral hygiene abstention period, with OL assessment performed at the same timepoints as at Day 1, *i.e.*, immediately after in-clinic product use, and at four and eight hours after in-clinic use. However, the difference between Day 1 and Day 8 assessments was that the subjects had been using their assigned products at home for a seven-day period. Secondary efficacy

analysis also evaluated surrogate oral malodor measures (hydrogen sulfide level, aerobic and anaerobic bacteria counts in CFU/cm$^2$) at both Day 1 and Day 8. Similar ANOVA analyses were used for these outcomes. In these analyses, logarithmic transformations were performed on the hydrogen sulfide and bacterial outcomes.

Safety analyses included clinical oral examination findings (the presence of abnormalities in the oral cavity) and adverse events (AE) experienced by the subjects. Oral examination findings were analyzed as the number and percent of subjects with abnormal results, and AEs were listed.

**General Analysis Considerations.** For all outcome comparisons, the least squares (LS) mean, Dunnett's adjusted standard error (SE) of the mean, and the two-sided 95% confidence intervals (CI) were presented by treatment group.

Due to the short duration and low-risk nature of this study, there were no pre-defined stopping rules. There were also no considerations for an interim analysis for this study.

## Results

There were 214 subjects screened for this study. Of these, 168 subjects were enrolled and randomized, with 77 male and 91 female participants. There was no statistical difference in gender distribution between treatment groups. The mean (SD) age of randomized subjects was 38.9 (14.8) years. Table II provides a depiction of subject screening, enrollment, randomization, and completion.

**Table II**
Subject Enrollment

| Subjects Screened N= 214 | | | | | | | |
|---|---|---|---|---|---|---|---|
| Screen Failures N=46 | Enrolled N=168 | | | | | | |
| | Not Randomized N=0 | Randomized N=168 | | | | | |
| | | STC N=56 | | LCM N=56 | | MTB N=56 | |
| | | C[a] N=55 | D[b] N=1 | C N=55 | D N=1 | C N=55 | D N=1 |

a: completed
b: discontinued

### *Organoleptic Endpoint Results*

Table III provides the analysis results of all organoleptic outcomes.

The primary efficacy objective was OL score reduction at eight hours post first product use at Day 1. For this timepoint, the LS mean (95% CI) organoleptic scores were: 1.70 (1.56, 1.84) for STC; 2.42 (2.28, 2.56) for LCM; and 2.33 (2.19, 2.47) for MTB (overall F-test p-value < 0.0001). The differences between STC and MTB, and STC and LCM, were significant (p-value < 0.0001, for each pair-wise comparison). With OL score expressed as LS mean (95% CI) percent reduction from pre-treatment, the following reductions were estimated: 46.7% (42.18%, 51.57%) for STC; 22.8% (18.31, 27.35%) for LCM; and 26.2% (21.66%, 30.72%) for MTB.

At immediately post-treatment, malodor was lowest in the STC group. The LS mean (95% CI) values were: 1.47 (1.36, 1.59) for STC; 1.57 (1.46, 1.69) for LCM; and 1.67 (1.55, 1.78) for MTB. However, only the STC versus MTB comparison was statistically significant (p-value = 0.0330). With OL score expressed as LS mean (95% CI) percent reduction from pre-treatment, the following reductions were estimated: 53.0% (49.37%, 56.71%); 49.6% (45.90%, 53.28%); and 47.1% (43.40%, 50.79%), for STC, LCM, and MTB, respectively.

At the four-hour post-treatment at Day 1, the STC group continued to have the lowest OL score. The LS mean (95% CI) values were: 1.77 (1.65, 1.90) for STC; 2.05 (1.92, 2.18) for LCM; and 1.94 (1.81, 2.07) for MTB, with the difference between STC and LCM statistically significant (p-value = 0.0062). Expressed as percent reduction from pre-treatment, the estimated LS mean (95% CI) values were: 44.06% (39.89%, 48.23%); 34.22% (30.03%, 38.41%); and 38.37% (34.16%, 42.57%) for STC, LCM, and MTB, respectively.

For organoleptic outcomes at Day 8, the STC treatment group exhibited the lowest OL value throughout the visit. Statistically significant differences (p-value < 0.05) were observed between STC and LCM, and STC and MTB, at each timepoint, immediately post-treatment, as well as four hours and eight hours following product use.

### *Surrogate Efficacy Endpoint Results*

**Hydrogen Sulfide.** Table IV provides the statistical analysis results of all $H_2S$ outcomes. Each of the treatments exhibited reductions in $H_2S$ compared to the pre-treatment value.

For between-group comparisons of $H_2S$ measurements following a single product use at Day 1, statistically significant differences were observed between STC and LCM at each timepoint (p-values of < 0.0001, = 0.0329, = 0.0073 at immediately, four-, and eight-hours following product use, respectively). A statistically significant difference was observed between STC and MTB only at the immediately post-treatment timepoint (p-value < 0.0001).

At Day 8, statistically significant differences were observed between STC and LCM, as well as STC and MTB, only at the immediately post-treatment timepoint (p-values < 0.0001).

**Microbial Counts.** Tables V and VI provide the statistical analysis results for the aerobic and anaerobic endpoints. Each of the three treatments exhibited reductions in aerobic and anaerobic counts compared to pre-treatment at Day 1 and Day 8.

For between-group comparisons in the quantification of aerobes following a single product use at Day 1, statistically significant differences were observed between STC and LCM at the immediately post timepoint (p-value = 0.0139), and between STC and MTB at the immediately-post and four-hour timepoints (p-value = 0.0136, and 0.0166, respectively). No differences were detected between treatment groups for anaerobic cultures following Day 1 product use.

At Day 8, for aerobes, statistically significant differences between STC and LCM were observed at each timepoint (p-value = 0.0016 at immediately-post, p-value = 0.0005 at four hours, and p-value = 0.0055 at eight hours). For STC and MTB, significant differences were observed at the immediately post and four-hour timepoints (p-value = 0.0064, and 0.0062, respectively). For anaerobes, the only statistically significant difference detected was between STC and LCM at the Day 8, immediately post timepoint (p-value = 0.0229).

### *Safety Results*

Three adverse events were reported during the study, including a bilateral linea alba, cheek biting, and chipped incisor edges on teeth numbers 8 and 9. The first two occurred in the STC treatment group and were judged as unlikely related to the study by the investigator, while the third occurred in the LCM group and was assessed as

unrelated to the study. All three adverse events were assessed as mild in severity.

## Discussion and Conclusions

Within the limits and controls of this study, each of the breath hygiene regimens tested are effective and safe for use. For the primary study objective, OL score eight hours following a single-use of the assigned product, the STC breath hygiene regimen (antimicrobial tongue spray plus powered tongue brushing) was superior to both LCM (rinse alone) and MTB (tongue brushing alone). This difference was not uniformly detected between products at the immediately or four hours post single-use at Day 1. However, statistically significant differences were sustained between STC and LCM, and STC and MTB following the seven-day product use period. Figure 1 illustrates the effect of each regimen on OL outcomes, by visit, at each timepoint. It is noted that the combined-regimen STC group exhibits the lowest OL value throughout.

For the surrogate endpoint, hydrogen sulfide gas chromatography, all three treatment groups exhibited reductions in the pre-treatment value at both Day 1 and Day 8. Evaluating the between-group comparisons, both tongue-brushing treatment groups appear to have a more pronounced effect overall, compared to use of rinse alone, at each study visit. The effect of the STC regimen is both immediate and sustained until eight hours, at both Day 1 and Day 8; whereas, the effect of tongue-brushing with MTB alone does not appear to have an immediate effect on $H_2S$, but it does show an effect by four and eight hours. This also appears to be the trend exhibited by the LCM treatment group, though to a more modest magnitude. At both Day 1 and Day 8, each of the breath hygiene regimens appear to exhibit the most impact on $H_2S$ at the four-hour timepoint. The depiction of $H_2S$ outcomes is provided in Figure 2.

The detection of an effect on breath as measured by a tongue microbial sample that was cultured under aerobic and anaerobic conditions indicates reductions from pre-treatment for all three treatment groups at both Day 1 and Day 8. For aerobes, intermittent differences between STC and either LCM or MTB were observed following a

**Table III**
Organoleptic Analysis

| | Statistic | STC | LCM | MTB | p-value[a] |
|---|---|---|---|---|---|
| Day 1 | | | | | |
| Pre-treatment | LS Mean (SE) | 3.09 (0.05) | 3.16 (0.05) | 3.20 (0.05) | 0.3602 |
| | 95% CI | (2.98, 3.20) | (3.05, 3.27) | (3.09, 3.31) | |
| | p-value[b] | | 0.5885 | 0.2682 | |
| Post-treatment | LS Mean (SE) | 1.47 (0.06) | 1.57 (0.06) | 1.67 (0.06) | 0.0592 |
| | 95% CI | (1.36, 1.59) | (1.46, 1.69) | (1.55, 1.78) | |
| | p-value[b] | | 0.3645 | 0.0330 | |
| | LS Mean (SE) PRFP[c] | 53.04 (1.86) | 49.59 (1.87) | 47.10 (1.87) | |
| | 95% CI | (49.37, 56.71) | (45.90, 53.28) | (43.40, 50.79) | |
| 4 hours | LS Mean (SE) | 1.77 (0.07) | 2.05 (0.07) | 1.94 (0.07) | 0.0123 |
| | 95% CI | (1.65, 1.90) | (1.92, 2.18) | (1.81, 2.07) | |
| | p-value[b] | | 0.0062 | 0.1393 | |
| | LS Mean (SE) PRFP[c] | 44.06 (2.11) | 34.22 (2.12) | 38.37 (2.13) | |
| | 95% CI | (39.89, 48.23) | (30.03, 38.41) | (34.16, 42.57) | |
| 8 hours | LS Mean (SE) | 1.70 (0.07) | 2.42 (0.07) | 2.33 (0.07) | <0.0001 |
| | 95% CI | (1.56, 1.84) | (2.28, 2.56) | (2.19, 2.47) | |
| | p-value[b] | | <0.0001 | <0.0001 | |
| | LS Mean (SE) PRFP[c] | 46.67 (2.28) | 22.83 (2.29) | 26.19 (2.29) | |
| | 95% CI | (42.18, 51.17) | (18.31, 27.35) | (21.66, 30.72) | |
| Day 8 | | | | | |
| Pre-treatment | LS Mean (SE) | 2.62 (0.08) | 2.95 (0.08) | 2.89 (0.08) | 0.0082 |
| | 95% CI | (2.46, 2.77) | (2.79, 3.10) | (2.74, 3.05) | |
| | p-value[b] | | 0.0074 | 0.0292 | |
| Post-treatment | LS Mean (SE) | 1.41 (0.05) | 1.95 (0.05) | 1.84 (0.05) | <0.0001 |
| | 95% CI | (1.31, 1.51) | (1.85, 2.05) | (1.74, 1.93) | |
| | p-value[b] | | <.0001 | <.0001 | |
| | LS Mean (SE) PRFP[c] | 49.96 (2.03) | 28.60 (2.00) | 34.82 (1.99) | |
| | 95% CI | (45.96, 53.96) | (24.64, 32.55) | (30.88, 38.75) | |
| 4 hours | LS Mean (SE) | 1.75 (0.07) | 2.26 (0.07 | 2.18 (0.06 | <0.0001 |
| | 95% CI | (34.32, 44.18) | (13.46, 23.21) | (16.20, 25.90) | |
| | p-value[b] | | <0.0001 | <0.0001 | |
| | LS Mean (SE) PRFP[c] | 39.25 (2.50) | 18.34 (2.47) | 21.05 (2.46) | |
| | 95% CI | (34.32, 44.18) | (13.46, 23.21) | (16.20, 25.90) | |
| 8 hours | LS Mean (SE) | 1.89 (0.07) | 2.41 (0.07 | 2.40 (0.07 | <0.0001 |
| | 95% CI | (1.75, 2.03 | (2.28, 2.55) | (2.26, 2.53 | |
| | p-value[b] | | <0.0001 | <0.0001 | |
| | LS Mean (SE) PRFP[c] | 34.54 (2.66) | 12.07 (2.63) | 13.57 (2.61) | |
| | 95% CI | (29.29, 39.79) | (6.88, 17.26) | (8.40, 18.73) | |

[a] p-value is based on an ANOVA model F-test (Ho: No differences between the 3 treatment groups)
[b] Dunnett's test p-values, for multiple comparisons, each treatment is compared to STC
[c] PRFP = Percent Reduction from Pre-treatment value

**Table IV**
Hydrogen Sulfide Analysis, Log10 in ppb

| | Statistic | STC | LCM | MTB | p-value[a] |
|---|---|---|---|---|---|
| Day 1 | | | | | |
| Pre-treatment | LS Mean (SE) | 2.04 (0.11) | 2.03 (0.12) | 2.00 (0.12) | 0.9575 |
| | 95% CI | (1.82, 2.27) | (1.80, 2.26) | (1.77, 2.23) | |
| | p-value[b] | | 0.9955 | 0.9423 | |
| Post-treatment | LS Mean (SE) | 0.89 (0.11) | 1.81 (0.11) | 1.55 (0.11) | <0.0001 |
| | 95% CI | (0.68, 1.10) | (1.60, 2.02) | (1.34, 1.76) | |
| | p-value[b] | | <0.0001 | <0.0001 | |
| 4 Hours | LS Mean (SE) | 0.82 (0.12) | 1.22 (0.12) | 0.87 (0.12) | 0.0359 |
| | 95% CI | (0.59, 1.05) | (0.99, 1.45) | (0.64, 1.11) | |
| | p-value[b] | | 0.0329 | 0.9345 | |
| 8 Hours | LS Mean (SE) | 0.92 (0.12) | 1.42 (0.12) | 1.18 (0.12) | 0.0148 |
| | 95% CI | (0.68, 1.16) | (1.18, 1.66) | (0.94, 1.42) | |
| | p-value[b] | | 0.0073 | 0.2222 | |
| Day 8 | | | | | |
| Pre-treatment | LS Mean (SE) | 1.83 (0.13) | 2.07 (0.13) | 1.75 (0.13) | 0.2172 |
| | 95% CI | (1.57, 2.09) | (1.81, 2.33) | (1.49, 2.01) | |
| | p-value[b] | | 0.3405 | 0.8850 | |
| Post-treatment | LS Mean (SE) | 0.88 (0.10) | 1.80 (0.11) | 1.56 (0.11) | <0.0001 |
| | 95% CI | (0.68, 1.09) | (1.59, 2.01) | (1.36, 1.77) | |
| | p-value[b] | | <0.0001 | <0.0001 | |
| 4 Hours | LS Mean (SE) | 0.80 (0.12) | 1.16 (0.12) | 0.93 (0.12) | 0.1053 |
| | 95% CI | (0.56, 1.03) | (0.92, 1.39) | (0.70, 1.17) | |
| | p-value[b] | | 0.0654 | 0.6304 | |
| 8 Hours | LS Mean (SE) | 1.08 (0.12) | 1.33 (0.12) | 1.08 (0.12) | 0.2447 |
| | 95% CI | (0.85, 1.31) | (1.09, 1.56) | (0.84, 1.31) | |
| | p-value[b] | | 0.2509 | 0.9999 | |

[a] p-value is based on an ANOVA model F-test (Ho: No differences between the 3 treatment groups)
[b] Dunnett's test p-values, for multiple comparisons, each treatment is compared to STC

**Table V**
Tongue Microbial Sample, Aerobes, Log10 CFU/mL

| | Statistic | STC | LCM | MTB | p-value[a] |
|---|---|---|---|---|---|
| Day 1 | | | | | |
| | LS Mean (SE) | 6.65 (0.06) | 6.65 (0.06) | 6.66 (0.06) | 0.9948 |
| | 95% CI | (6.54, 6.76) | (6.54, 6.76) | (6.54, 6.77) | |
| | p-value[b] | | 0.9999 | 0.9952 | |
| Post-treatment | LS Mean (SE) | 6.07 (0.07) | 6.32 (0.07) | 6.32 (0.07) | 0.0082 |
| | 95% CI | (5.94, 6.20) | (6.19, 6.45) | (6.19, 6.45) | |
| | p-value[b] | | 0.0139 | 0.0136 | |
| 4 Hours | LS Mean (SE) | 6.12 (0.06) | 6.29 (0.06) | 6.35 (0.06) | 0.0235 |
| | 95% CI | (6.00, 6.24) | (6.17, 6.42) | (6.23, 6.47) | |
| | p-value[b] | | 0.0843 | 0.0166 | |
| 8 Hours | LS Mean (SE) | 6.14 (0.06) | 6.25 (0.06) | 6.29 (0.06) | 0.2269 |
| | 95% CI | (6.02, 6.26) | (6.12, 6.37) | (6.16, 6.41) | |
| | p-value[b] | | 0.3637 | 0.1716 | |
| Day 8 | | | | | |
| Pre-treatment | LS Mean (SE) | 6.70 (0.05) | 6.67 (0.05) | 6.64 (0.05) | 0.7131 |
| | 95% CI | (6.60, 6.80) | (6.57, 6.77) | (6.55, 6.74) | |
| | p-value[b] | | 0.8781 | 0.6225 | |
| Post-treatment | LS Mean (SE) | 6.08 (0.06) | 6.36 (0.06) | 6.32 (0.06) | 0.0013 |
| | 95% CI | (5.96, 6.19) | (6.25, 6.47) | (6.21, 6.44) | |
| | p-value[b] | | 0.0016 | 0.0064 | |
| 4 Hours | LS Mean (SE) | 6.02 (0.06) | 6.33 (0.06) | 6.27 (0.06) | 0.0006 |
| | 95% CI | (5.90, 6.14) | (6.22, 6.45) | (6.15, 6.39) | |
| | p-value[b] | | 0.0005 | 0.0062 | |
| 8 Hours | LS Mean (SE) | 6.15 (0.06) | 6.40 (0.06) | 6.31 (0.06) | 0.0098 |
| | 95% CI | (6.03, 6.26) | (6.28, 6.51) | (6.20, 6.43) | |
| | p-value[b] | | 0.0055 | 0.0846 | |

[a] p-value is based on an ANOVA model F-test (Ho: No differences between the 3 treatment groups)
[b] Dunnett's test p-values, for multiple comparisons, each treatment is compared to STC

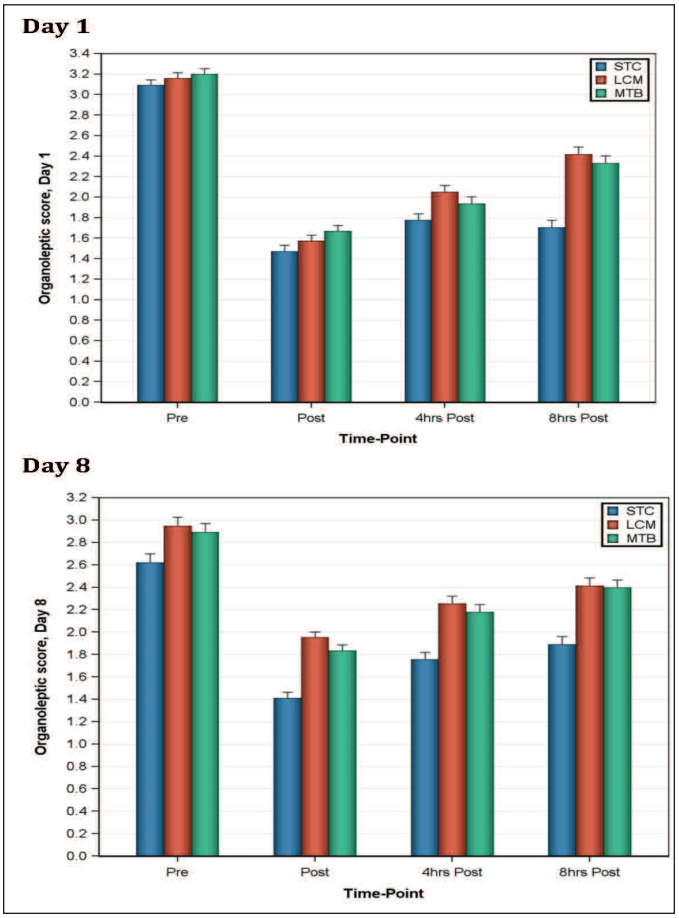


**Figure 1.** *Least squares mean, organoleptic score.*

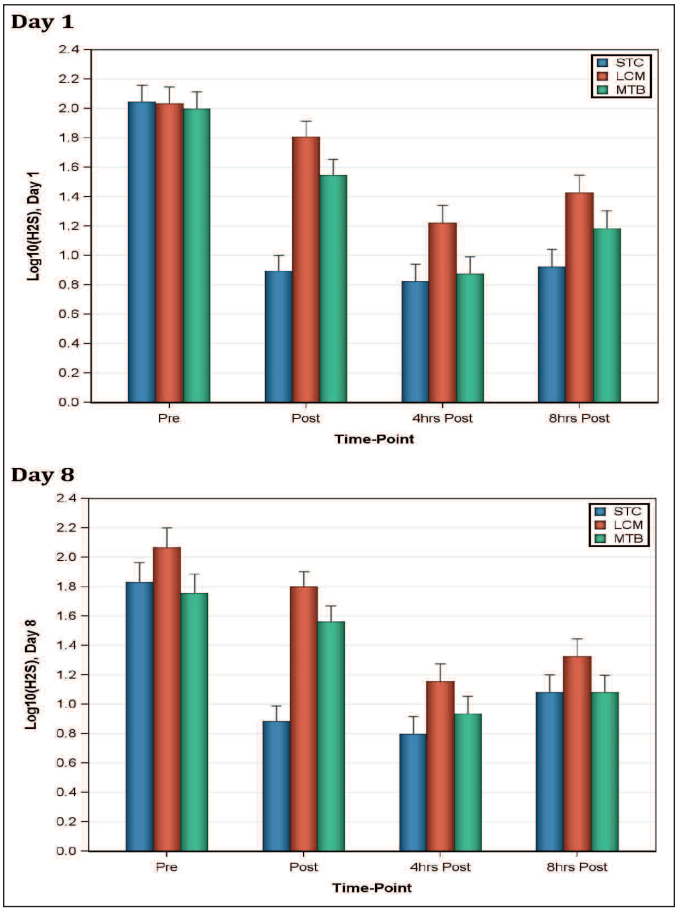


**Figure 2.** *Hydrogen sulfide, log10 reduction from pre-treatment.*

**Table VI**
Tongue Microbial Sample, Anaerobes, Log10 CFU/mL

| | Statistic | STC | LCM | MTB | p-value[a] |
|---|---|---|---|---|---|
| Day 1 | | | | | |
| Pre-treatment | LS Mean (SE) | 6.63 (0.06) | 6.49 (0.06) | 6.45 (0.06) | 0.1082 |
| | 95% CI | (6.51, 6.75) | (6.37, 6.61) | (6.33, 6.57) | |
| | p-value[b] | | 0.2104 | 0.0814 | |
| Post-treatment | LS Mean (SE) | 6.13 (0.05) | 6.24 (0.05) | 6.23 (0.06) | 0.2825 |
| | 95% CI | (6.02, 6.24) | (6.14, 6.35) | (6.12, 6.34) | |
| | p-value[b] | | 0.2389 | 0.3518 | |
| 4 hours | LS Mean (SE) | 5.93 (0.07) | 6.15 (0.07) | 6.00 (0.07) | 0.0922 |
| | 95% CI | (5.79, 6.07) | (6.01, 6.29) | (5.86, 6.15) | |
| | p-value[b] | | 0.0596 | 0.6967 | |
| 8 Hours | LS Mean (SE) | 5.85 (0.08) | 6.05 (0.08) | 5.97 (0.08) | 0.1709 |
| | 95% CI | (5.70, 6.00) | (5.90, 6.20) | (5.81, 6.12) | |
| | p-value[b] | | 0.1098 | 0.4631 | |
| Day 8 | | | | | |
| Pre-treatment | LS Mean (SE) | 6.34 (0.06) | 6.29 (0.06) | 6.25 (0.06) | 0.5584 |
| | 95% CI | (6.23, 6.46) | (6.18, 6.41) | (6.14, 6.37) | |
| | p-value[b] | | 0.7622 | 0.4539 | |
| Post-treatment | LS Mean (SE) | 5.64 (0.07) | 5.89 (0.07) | 5.80 (0.07) | 0.0395 |
| | 95% CI | (5.50, 5.78) | (5.75, 6.03) | (5.66, 5.94) | |
| | p-value[b] | | 0.0229 | 0.1873 | |
| 4 Hours | LS Mean (SE) | 5.44 (0.09) | 5.63 (0.09) | 5.43 (0.09) | 0.2330 |
| | 95% CI | (5.26, 5.63) | (5.44, 5.82) | (5.24, 5.61) | |
| | p-value[b] | | 0.2686 | 0.9891 | |
| 8 Hours | LS Mean (SE) | 5.61 (0.09) | 5.71 (0.09) | 5.76 (0.09) | 0.4860 |
| | 95% CI | (5.43, 5.79) | (5.54, 5.89) | (5.58, 5.94) | |
| | p-value[b] | | 0.6350 | 0.3931 | |

[a]p-value is based on an ANOVA model F-test (Ho: No differences between the 3 treatment groups)
[b]Dunnett's test p-values, for multiple comparisons, each treatment is compared to STC

single product use at Day 1. Statistical differences, however, showed a more general trend at Day 8, with STC statistically different from LCM at all timepoints, and from MTB up to four hours following product use.

The analysis of anaerobic culture outcomes indicates reductions in concentration for all three treatment groups, at each visit. However, it did not indicate statistical dominance for differences between any of the products, though the STC group appears to trend lower throughout.

This study was designed and powered to determine whether organoleptic distinctions could be made between the three regimens. The addition of the $H_2S$ and microbial count endpoints was exploratory in nature and intended to help elicit how each treatment modified oral malodor (by reducing sulfide gas in breath, or by affecting the bacterial ecology of the tongue). Also, they were included to determine whether these additional endpoints tracked with the observed changes in organoleptic score. A cursory look at trends is difficult to interpret. As such, a supplementary correlation analysis was completed in order to evaluate these surrogate endpoints, relative to the organoleptic measure. The r-squared value for OL and $H_2S$ was 0.11; for OL and aerobes, it was 0.10; and for OL and anaerobes, it was -0.09. Similarly, low r-squared values were observed at other timepoints and also between $H_2S$, aerobes, and anaerobes. In general, neither $H_2S$ nor microbial counts were compelling surrogate markers for organoleptic oral malodor detection in this trial. The use of these measures as surrogates for OL in any subsequent study should be initiated with caution, with the appropriate statistical and population eligibility requirements carefully planned.

As the focus of this study was primarily on organoleptic effects of each regimen up to eight hours following use on a given treatment day, the effect of each regimen following the home use period was not an explicit objective. That said, each of the regimens do appear to have an effect following the seven-day home use period. A cursory look at the pre-treatment value at Day 1 appears different than the pre-treatment value at Day 8. To explore this further, a *post hoc* analysis was completed to evaluate the extent to which statistical differences in OL, following repeat use, may exist. Table VII provides this analysis. In particular, daily use of the tongue-brushing regimens (STC, MTB) appear to be most effective over the seven-day home use period, with STC exhibiting an LS mean (95% CI) reduction of 15.77% (10.91%, 20.64%), and MTB exhibiting an 8.46% (3.50%, 13.32%) reduction. Whereas, reduction over time following use of LCM was 5.49% (0.65%, 10.34%). In this *post hoc* analysis, the difference between STC and LCM was statistically significant (p-value = 0.0069). These over-time effects should be taken into consideration for future study designs. From a patient's point of view, such outcomes may prove to provide a more meaningful gauge of product efficacy, where a steady decline in malodor measures following regular and repeat product use, helps limit cyclical extremes of malodor in a given day, and over a period of days.

An additional observation is with respect to changes in quantification of the microbial population. Reductions were observed for all treatment groups at both Day 1 and Day 8. There have been variable reports of success in assessing the effects of tongue cleaning on the microbial population of the tongue dorsum following introduction of a treatment.[17-19] With its reasonably large sample size and repeated-measures approach, this study does provide some fruitful general evidence to suggest that the microbial population of the tongue is, indeed, altered following intervention. In future studies, additional sensitivities may be gained by including endpoints, and a subsequent correlation exercise, where microbial speciation analysis with PCR, rather than the more general aerobic/anaerobic quantification, is utilized.

It is also noted that this study design necessitated the standardization of the tooth brushing regimen across all treatment groups in order to isolate differences in breath regimens. The selected toothbrush, in this case, was a Sonicare powered toothbrush (PTB). This PTB has previously been demonstrated to reduce gingival inflammation and supragingival plaque in as early as two weeks.[20-22] As oral status, notably plaque and gingivitis, can be potential sources of malodor, it is not possible to determine the extent to which the standardized use of the PTB may have affected the malodor outcomes, in particular, following the seven-day period of use. Going forward, a study design that includes a negative control for both oral and breath hygiene may help elicit these effects. Ideally, plaque and gingivitis endpoints would be assessed in this model, as well.

Overall, a breath hygiene regimen that includes mechanical disruption of the tongue microflora appears to be a more effective approach for patients managing oral malodor than the use of an antimicrobial rinse alone. Combining these techniques – tongue

**Table VII**
Comparison of Organoleptic Values, Day 1 Pre-treatment to Day 8 Pre-treatment

| | Statistic | STC | LCM | MTB | p-value[a] |
|---|---|---|---|---|---|
| Day 1 | | | | | |
| Pre-treatment | LS Mean (SE) | 3.09 (0.05) | 3.16 (0.05) | 3.20 (0.05) | 0.3602 |
| | 95% CI | (2.98, 3.20) | (3.05, 3.27) | (3.09, 3.31) | |
| Day 8 | | | | | |
| Pre-treatment | LS Mean (SE) | 2.65 (0.07) | 2.94 (0.07) | 2.87 (0.07) | 0.0163 |
| | 95% CI | (2.50, 2.79) | (2.80, 3.09) | (2.72, 3.01) | |
| Reduction, Day 1 to Day 8 | | | | | |
| | LS Mean (SE) | 0.50 (0.07) | 0.21 (0.07) | 0.28 (0.07) | 0.0163 |
| | 95% CI | (0.36, 0.65) | (0.06, 0.35) | (0.14, 0.43) | |
| | p-value[b] | | 0.0110 | 0.0712 | |
| Percent Reduction, Day 1 to Day 8 | | | | | |
| | LS Mean (SE) | 15.77 (2.46) | 5.49 (2.45) | 8.46 (2.46) | 0.0112 |
| | 95% CI | (10.91, 20.64) | (0.65, 10.34) | (3.60, 13.32) | |
| | p-value[b] | | 0.0069 | 0.0690 | |

[a] p-value is based on an ANOVA model F-test (Ho: No differences between the 3 treatment groups)
[b] Dunnett's test p-values, for multiple comparisons, each treatment is compared to STC

brushing with antimicrobial rinse application – does appear to have the most impactful effect on the organoleptic character of breath. Prior studies [12,23] have also reported improvements in oral malodor using a combined treatment approach, though the small sample size in both of these cited studies and the lack of a comparator in the latter, are noted. The current study, however, was not limited by these constraints, and for patients who suffer from oral malodor this two-pronged approach may provide a more pronounced immediate and, so-called "all-day" (*i.e.*, eight-hour), benefit.

**Acknowledgments:** The study was supported by a grant from Philips Oral Healthcare. The authors would like to express their appreciation to the following individuals for their assistance with the conduct and reporting of the study: Raydolfo Aprecio, Michele Arambula, Weixue Chen, Wei Jiang, Marilou Laserna, Omaima Mohamed, Min Meng, Alaina Piper, Mayra Orozco, Danying Tao, Dongjiao Zhang, Zhe Zhong, Krishnakant Nammi, and Wendy Jenkins.

**Conflict of Interest:** Yiming Li, Sean Lee, Joni Stephens, Wu Zhang, and Montry Suprono are employed by Loma Linda University. Anthony Mwatha, Marilyn Ward, and Farah Mirza are employed by Philips Oral Healthcare.

***For correspondence with the authors of this paper, contact Dr. Yiming Li -- yli@llu.edu.***

## References

1. Tonzetich J. Production and origin of oral malodor: a review of mechanisms and methods of analysis. *J Periodontol* 1977;48:13-20.
2. De Boever EH, Loesche WJ. Assessing the contribution of anaerobic microflora of the tongue to oral malodor. *J Am Dent Assoc* 1995;126:1384-93.
3. Bosy A, Kulkami GV, Rosenberg M, McCulloch CAG. Relationship of oral malodor to periodontitis: evidence of independence in discrete subpopulations. *J Periodontol* 1994;65:37-46.
4. Kazor CE, Mitchell PM, Lee AM, Stokes LN, Loesche WJ, Dewhirst FE, Paster BJ. Diversity of bacterial populations on the tongue dorsa of patients with halitosis and healthy patients. *J Clin Microbiol* 2003;41:558-63.
5. Yaegaki K, Sanada K. Volatile sulfur compounds in mouth air from clinically healthy subjects and patients with periodontal disease. *J Periodont Res* 1992;27(4 Pt 1):233-8.
6. Iatropoulos A, Panis V, Mela E, Stefaniotis T, Madianos PN, Papaioannou W. Changes of volatile sulphur compounds during therapy of a case series of patients with chronic periodontitis and halitosis. *J Clin Periodontol* 2016; 43(4): 359–365
7. Figueiredo LC, Rosetti EP, Marcantonio E, Marcantonio RAC, Savador SL. The relationship of oral malodor in patients with or without periodontal disease. *J Periodontol* 2002; 73:1338-42.
8. Borden LC, Chaves ES, Bowman JP, Fath BM, Hollar GL. The effect of four mouthrinses on oral malodor. *Compend Contin Educ Dent* 2002;23:531-36, 538.
9. Young A, Jonski G, Rölla G. Inhibition of orally produced volatile sulfur compounds by zinc, chlorhexidine or cetylpyridinium chloride – effect of concentration. *Eur J Oral Sci* 2003;111: 400-4.
10. Roldán S, Winkel EG, Herrera D, Sanz M, VanWinkelhoff AJ. 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;30:427–34.
11. Acar B, Berker E, Tan C, larslan YD, Tekçiçek M, Tezcan I. Effects of oral prophylaxis including tongue cleaning on halitosis and gingival inflammation in gingivitis patients—a randomized controlled clinical trial. *Clin Oral Investig* 2018; Sep 13. doi: 10.1007/s00784-018-2617-5. [Epub ahead of print]
12. Ademovski SE, Lingström P, Winkel E, Tangerman A, Persson GR, Renvert S. Comparison of different treatment modalities for oral halitosis. *Acta Odontol Scand* 2012;70:224-33.
13. Rosenberg M, Kulkarni GV, Bosy A, MacCullough CAG. Reproducibility and sensitivity of oral malodor measurements with a portable sulphide monitor. *J Dent Res* 1991;70:1436-40.
14. Hartley MG, El-Maaytah MA, McKenzie C, Greenman J. Assessment of an impressed toothbrush as a method of sampling tongue microbiota. In: van Steenberghe D, Rosenber M, eds. *Bad Breath: a Multidisciplinary Approach*. Belgium: Leuven University Press, 1996; pp. 123-33.
15. Hartley G, McKenzie C, Greenman J, El-Maaytah, MA, Scully C, Porter S. Tongue microbiota and malodor: effects of metronidazole mouthrinse on tongue microbiota and breath odour. *Microb Ecol Health Dis* 1999;11:226-33.
16. Saad S, Gomez-Pereira P, Hewett K, Horstman P, Patel J, Greenman J. Daily reduction of oral malodor with the use of a sonic tongue brush combined with an antibacterial tongue spray in a randomized cross-over clinical investigation. *J. Breath Res* 2016;10:016013.
17. Ademovski SE, Persson GR, Winkel E, Tangerman A, Lingström P, Renvert S. The short-term treatment effects on the microbiota at the dorsum of the tongue in intra-oral halitosis patients—a randomized clinical trial. *Clin Oral Investig* 2013;17:463–73.
18. Laleman I, Koop R, Teughels W, Dekeyser C, Quirynen M. Influence of tongue brushing and scraping on the oral microflora of periodontitis patients. *J Periodontol Res* 2018;53:73-9.
19. Kurata H, Awano S, Yoshida A, Ansai T, Takehara T. The prevalence of periodontopathogenic bacteria in saliva is linked to periodontal health status and oral malodour. *J Med Microbiol* 2008;57(Pt 5):636-42.
20. Delaurenti M, Ward M, Souza S, Jenkins W, Putt M, Milleman J, Milleman K. The effect of use of a sonic power toothbrush and a manual toothbrush control on plaque and gingivitis. *J Clin Dent* 2017;28(Spec Iss A):A1-6.
21. Jenkins W, Souza S, Ward M, Defenbaugh J, Milleman K, Milleman J. An evaluation of plaque and gingivitis reduction following home use of sonicare flexcare platinum with premium plaque control brush head and a manual toothbrush. *J Clin Dent* 2017;28(Spec Iss A):A7-12.
22. De Jager M, Rmaile A, Darch O, Bikker JW. The effectiveness of manual versus high-frequency, high-amplitude sonic powered toothbrushes for oral health: a meta-analysis. *J Clin Dent* 2017;28(Spec Iss A):A13-28.
23. Roldán S, Herrera D, O'Connor A, González I, Sanz M. A combined therapeutic approach to manage oral halitosis: a 3-month prospective case series. *J Periodontol* 2005;76:1025-33.