Prevention and therapy of periodontal diseases and oral malodour

van der Sluijs, E.

6

The effect of a tooth/tongue gel and mouthwash regimen on morning oral malodour. A three week single-blind randomized clinical trial

E. van der Sluijs
G.A. van der Weijden
N.L. Hennequin-Hoenderdos
D.E. Slot

International Journal of Dental Hygiene
2017 accepted

Abstract

Aim

To compare the effects of a regimen consisting of a tooth/tongue gel, tongue cleaner and mouthwash with the effects of using standard fluoride dentifrice on the organoleptic oral malodour score (ORG) and volatile sulphur compounds (VSCs).

Methods

A total, 66 non-dental students participated in a 3-week parallel, single-blind, randomized, controlled clinical trial. The test group used a tongue cleaner, a tooth/tongue gel and mouthwash containing amine fluoride/stannous fluoride and zinc lactate as oral malodour counteractive. The control group used a standard fluoride dentifrice. Measurements were performed in the morning at baseline, at days 1, 7, and 21. The primary outcome was the ORG score. The secondary outcome, the VSC measurement, was assessed using OralChroma™ (H_2S, CH_3SH, $(CH_3)_2S$)) and Halimeter®. Tongue coating thickness and tongue discoloration were scored. At baseline and day 21, the participants' self-perceptions were assessed.

Results

At day 1 for the ORG, H_2S, CH_3SH and Halimeter® readings, a significant decrease was observed in the test group. At day 21, the decrease in H_2S and the Halimeter® outcomes were maintained for the test group, and a significant increase in tongue surface discoloration was observed. The test group evaluated their 'morning breath upon awakening' as significantly better ($P = 0.001$) after 21 days.

Conclusion

A significant overnight effect on morning oral malodour was observed for most of the parameters in favour of the test group. At day 21, the effect of prolonged use was significant for H_2S and the Halimeter® readings, although not for the primary ORG outcome parameter.

Introduction

In general, people are becoming more aware of oral malodour, which remains one of the greatest taboos in society (1). According to the European Federation of Periodontology guidelines for management of malodour (2), the dental care professional should provide personalized advice on bad breath when appropriate. This advice includes optimizing patient oral hygiene practices, such as tooth brushing and interdental cleaning. In addition, it is suggested that patients should be instructed and motivated to engage in tongue cleaning when tongue coating thickness is present and that the use of chemical agents with proven efficacy be recommended (3). The systematic review by Slot et al. (4), which supported these recommendations, included three experiments (5-7) evaluating regimens in which the use of a tongue cleaner was considered together with a chemical agent. Earlier, Blom et al.(8) have shown that nearly all mouth rinses with active ingredients have some beneficial effect with respect to oral malodour in short- and longer-term studies. Most evidence was available for chlorhexidine mouth rinses, and combinations with cetylpyridinium chloride and zinc provided the best evidence profiles.

One of the other commercially available mouth rinse products contains amine fluoride/stannous fluoride and zinc lactate as oral malodour counteractive. Clinical trials evaluating morning oral malodour have shown that this mouth rinse was effective compared to a negative control (9,10). In addition, there is also a tooth/tongue gel available that contains the same ingredients as the commercially available mouth rinse. This gel can be used in combination with a tongue cleaner. Two recent studies (11,12) showed that the combination of a tongue cleaner with the specific tooth and tongue gel had instant effects (12) that remained measurable after 7 days of use (11).

Mechanical approaches to cleaning the dorsum of the tongue, such as tongue brushing or tongue scraping, have the potential to successfully reduce tongue coating thickness and oral malodour (13). A recent observational survey evaluated participants' preferences and perceptions of the effectiveness of several commercially available tongue cleaners. Compared to other tongue cleaner products, the participants found two scrapers to be the most comfortable and effective (14). One of these scrapers was the product used in a regimen with a tooth and tongue gel (11,12). The clinical efficacy of using this regimen with in addition to a mouthwash that contains similar active ingredients has not yet been evaluated.

Therefore, the aim of this study was to investigate the proposed regimen of a tooth/brushing gel, a tongue scraper with gel, and a mouth rinse compared to brushing with a standard fluoride dentifrice regarding its effect on organoleptic scores and the concentration of volatile sulphur compounds (VSCs).

Materials and methods

The recommendations for strengthening the reporting were followed as suggested by the guideline Consolidated Standards of Reporting Trials (CONSORT) (15) and the checklist Template for Intervention Description and Replication (TIDieR) (16).

Ethical procedures

This study followed the Good Clinical Practice (CPMP/ICH/135/95) guidelines, in agreement with the ethical principles of the Declaration of Helsinki (October 2008, Brazil) and in accordance with the Medical Research Involving Human Subjects Act (WMO) and applicable local regulations. The study was approved by the medical ethical committee at Amsterdam Medical Centre (2011_301#B201283) and was registered at the Dutch Trial Register (#3240). The study was conducted from January 2012 until February 2013 at the Department of Periodontology Academic Centre for Dentistry of Amsterdam (ACTA), the Netherlands. Before enrolment, all of the volunteers were provided with verbal and written information regarding the aim, rationale and duration of the study. The investigator explained the details of the trial and the potential risks involved. Prior to the study procedures, an informed consent form was signed by all eligible subjects who had agreed to participate, and they were informed that they were free to withdraw at any time.

Sample size

Sample size calculations were performed using the PS Power and Sample Size Program (17,18). The *a priori* sample size calculation, based on pooled data of the mean differences between interventions from a previous study (19), showed that, for a detectable difference of 0.4 in the primary outcome between day 0 and day 21 with a pooled standard deviation of 0.5 and an alpha of 0.05 in conjunction with a beta of 80%, a total of 25 participants were needed per group. To account for a dropout rate of 20%, 30 participants per group needed to be enrolled, which was in agreement with ADA guidelines (20).

Recruitment and inclusion

The participants were students from different universities and colleges in and around Amsterdam. They were notified by e-mail and flyer advertisement to participate in this study. Dental care professionals and dental students were excluded from participation. Those who responded to the invitation first completed an online standard questionnaire with questions related to eligibility. If considered potentially eligible, they were invited to be screened by a dental hygienist (EVDS). To qualify for inclusion, the subjects were required to be ≥18 years old, classified as systemically healthy as assessed by the medical questionnaire, have ≥5 teeth per quadrant, be periodontally healthy as clinically assessed by scoring the Dutch Periodontal Screening Index with inclusion criterion scores ≤3 minus (21,22) a mean organoleptic score of 2 at screening and baseline, and be willing to adhere to specific regulations with respect to food constituents and perfume products and the provisions to be undertaken prior to the breath analyses, as described in detail by Van der Sluijs et al. 2016 (19) (see online appendix S1). Excluded were those who presented with any of the following: orthodontic appliances, removable (partial) dentures, smoking, overt caries, any pathological alterations of the oral mucosa, pregnancy, acute sinusitis or severe oral-pharyngeal infections, any medications that can cause malodour, and reduced salivary flow. Additionally, participation in a clinical trial within the previous 30 days was not allowed. At baseline details of, the participants' regular daily oral hygiene routine were recorded.

Design

This was a 3-week single (examiner) blind, parallel, randomized controlled clinical trial (RCT). Randomization was performed using true random numbers, which were generated by sampling, and by processing a source of atmospheric noise (23). Every participant received a unique subject identification number. The study supervisor used the obtained allocation randomization sequence to assign the participants to the test or control group. For details of the regimens, see table 1. No stratification was applied. To conceal the intervention regimen to the examiners, the participants were instructed not to reveal their intervention regimen in any manner. Records of earlier examinations were not available to the examiners at any time point.

Table 1 The following regimens were assigned by groups.

	Test regimen	**Control regimen**
Toothbrush	GABA International Meridol® toothbrush	
Brushing instructions	Brush twice per day for 2 minutes with pea-sized amount of dentifrice (29)	
Dentifrice	Meridol® Halitosis tooth and tongue gel GABA International Therwil, Switzerland	Everclean, HEMA Amsterdam, the Netherlands
Ingredients of dentifrice	Amine fluoride, stannous fluoride, zinc lactate, oral malodour counteractive	Sodium fluoride
Tongue cleaner	Meridol® Halitosis, Therwil, Switzerland	NA
Instructions for tongue cleaner	Once per day in the evening after brushing. Distribute gel on the tongue surface. The tongue is divided with two imaginary vertical lines, resulting in three parts. Each part is cleaned with the tongue cleaner with tooth and tongue gel. Described in detail by Wilhelm et al. (11)	
Mouthwash	Meridol® Halitosis mouthwash containing GABA International Therwil, Switzerland	NA
Ingredients in mouthwash	Amine fluoride, stannous fluoride, zinc lactate, oral malodour counteractive	
Instructions for mouthwash	Once per day in the evening after tongue cleaning Rinse with 15 ml for one minute. For the last 10 s, gargle while sticking out the tongue, and bend the head slightly backwards	
After use instructions	Do not eat or drink within 30 minutes after the oral hygiene intervention	

* NA = not applicable

Outcomes variables

The primary outcome variable was the organoleptic score, assessed by a standardized procedure on a scale of 0-5, and was defined as follows: score 0= absence of odour, 1= barely noticeable odour, 2= slight odour, 3= moderate odour, 4= strong odour, and 5= extremely strong odour. The trained organoleptic judge (EVDS) used as a memory refresher the Smell Identification Test® to optimize the ability to distinguish odours (Sensonics Inc., Haddon Heights, NJ, USA) and especially to be able to detect odours at low concentrations (24-27).
The secondary outcomes were the volatile sulphur compounds (VSCs) assessed using two specific measuring devices: a Halimeter® RH-17 (Interscan Corporation, Simi Valley, CA, USA) and the OralChroma™ with data management software: ABIMEDICAL for Windows, version 3.5.0 (FIS Inc. Itami, Hyogo, Japan). The devices differ by the sensor principle of operation, the OralChroma™ uses a gas chromatography technology which separates the concentration values of hydrogen sulphide (H_2S), methyl mercaptan (CH_3SH), and dimethyl sulphide ($(CH_3)_2S$). The Halimeter® uses a electrochemical voltammetric technique which displays a peak concentration of VSCs. The second examiner (SCS) was responsible for operating this equipment in the absence of the organoleptic judge to avoid bias to the organoleptic assessment based on the VCS outcomes. At 24 hours before each visit, both apparatuses were switched on to enable them to acclimatize and self-calibrate to the ambient air. The Halimeter® and OralChroma™ were checked and professionally calibrated by the manufacturers prior to the study. Both procedures are further described in detail by Van der Sluijs et al. (19).
Additionally, the tongue coating thickness and tongue surface discoloration were scored (EVDS) according to the methods described in detail by Mantilla Gómez et al. (21). The tongue surface was assessed from the circumvallate papilla to the tip, i.e., back third, middle third, and front third.
For patient reported outcomes participants were questioned about their self-perceptions of morning oral malodour before and at the end of the clinical trial, with the aid of a visual analogue scale (VAS) (28). They were requested to place a vertical mark on a 10-cm-long uncalibrated line (0-10). The left extreme represented the negative (stale, score 0), whereas the right extreme represented the positive (very fresh, score 10). In addition, feedback was requested regarding their perceptions of the assigned regimen.

Procedure

Prior to the assessment, the participants were requested to adhere to specific regulations with respect to food constituents and perfume products to avoid factors that could influence the breath analyses, as suggested in the ADA guidelines (20) and further described in Van der Sluijs et al. (19) (see appendix S1). Instructions about these regulations were provided. Text messages (SMS-Short Message Service) were sent to remind each participant within 48 hours before each visit. Clinical assessments were performed between 7:00 and 12:00 am with a follow-up of 21 days and took place in the same room under the same conditions. At visit 1 (baseline), the participant used the randomly assigned products under supervision after verbal instructions from the coordinator (NLHH). As a reminder, short written instructions were provided, and the participants were asked to mark their use of

each regimen on a calendar. The participants received further specific instructions according to their assigned test and control regimen products.

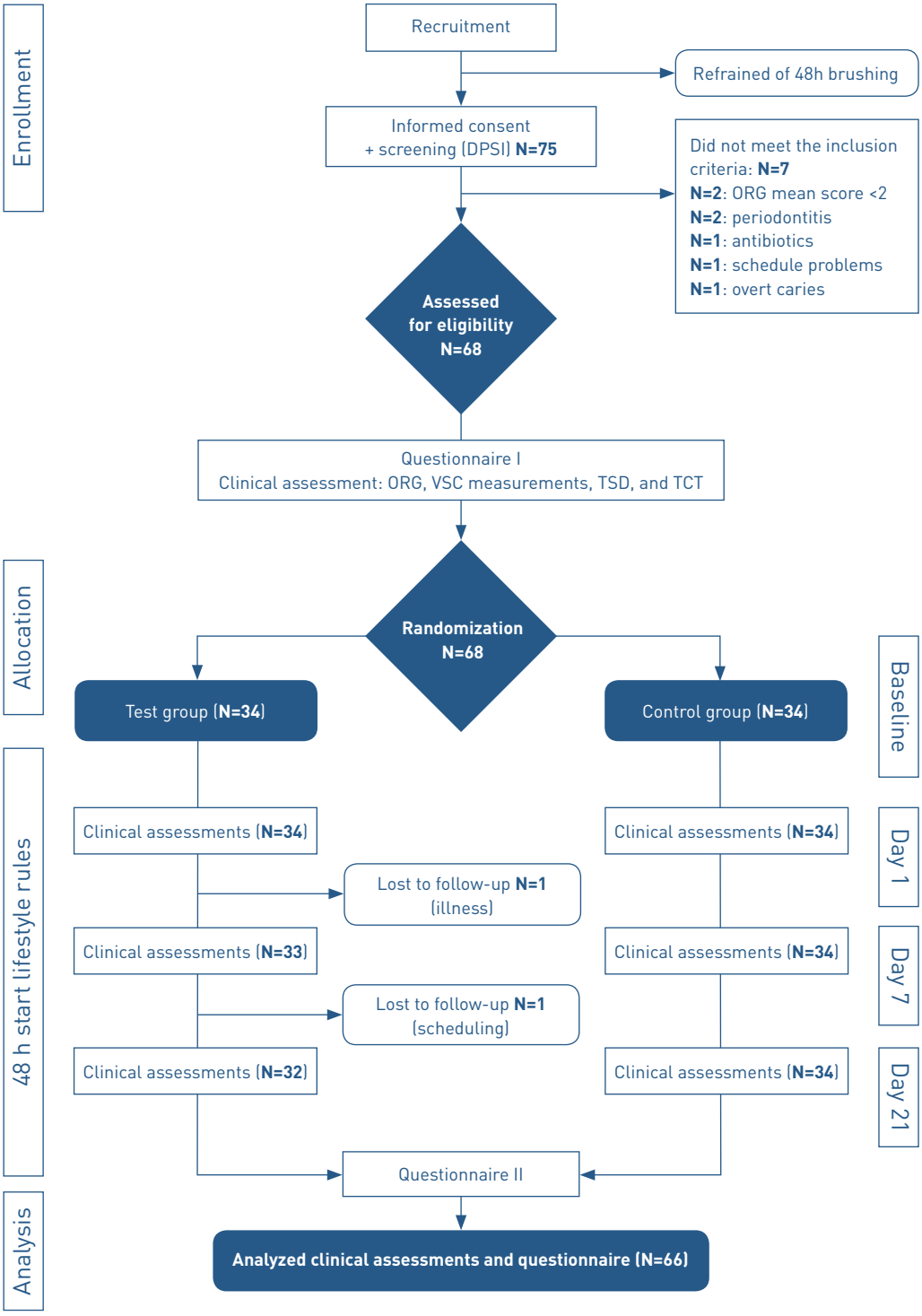
Monitoring of compliance and adverse events

At the end of the study, the participants were asked to return their (empty or not) tubes of dentifrice and bottles of mouth rinse to measure the remaining paste/fluid. The calendars were checked for compliance. Adverse events were reported to the study coordinator.

Statistical analysis

The statistical analysis, using SPSS software (Statistical Package for Social Sciences, IBM Corp. Released 2013. IBM SPSS Statistics for Windows, Version 22.0. Armonk, NY, USA: IBM Corp.), was performed before the randomization code was broken. The population demographics are presented by test and control regimen. Oral hygiene practices and the presence of tongue fissures are reported as absolute numbers and as frequencies. Furthermore, both organoleptic scores and VSC readings at any time point were summarized with their means and standard deviation (SD). Additionally, means and SDs were calculated for the VAS scores of the questionnaires. Data were tested for normal distribution by the Kolmogorov-Smirnov test and were found not to correspond with the normal distribution assumption ($p>0.05$). Based on the non-normal distribution, the Wilcoxon's test, a non-parametric statistical approach, was used for the within group changes with Wilcoxon's test and between group differences by the Mann-Whitney U test. Changes within groups and differences between groups were calculated. For the questionnaire, a parametric statistical technique was applied with the independent T-test. Values $p\leq0.05$ were accepted as statistically significant. Correlations between tongue surface discoloration and the used volume of mouth rinse were analysed using Pearson's correlation. The relationship between tongue surface characteristics and organoleptic scores were analysed using Spearman's rho.

Figure 1 Study flow of the study design and population.



ORG = Organoleptic score; TCD = Tongue surface discoloration; TCT = Tongue coating thickness;
VSC measurements = using OralChroma™ and Halimeter® apparatuses;
Test group = used toothbrush, tongue cleaner, tooth gel and tongue gel, mouthwash;
Control group = used toothbrush, standard fluoride dentifrice

Results

Participants

As proposed by CONSORT (15), the study flow is shown in figure 1. Throughout the study, one single adverse event was detected by the examiner on the palate of the oral cavity that was not related to use of the study products. The details were presented separately in a case report (30).

Table 2 presents the demographics of the participants and their daily oral hygiene habits. In terms of baseline characteristics, the two groups were well balanced (P>0.05); however, more female subjects participated. The majority of participants used a manual toothbrush. Almost 25% of the participants used interdental devices, e.g., floss and woodsticks. Approximately 40% of the participants reported performing some form of tongue cleaning.

Table 2 Subjects' demographic and oral hygiene practices according to the assigned group regimen.

N=66			Test regimen N=32	Control regimen N=34
Population		Mean age (SD) Range	22.84 (3.22) 19-32	22.21 (3.44) 19-39
		♀ Female	25 (78.1%)	24 (70.6%)
Tongue fissure		Present	1 (3.1%)	2 (5.9%)
Hygiene devices	Tooth Brush	Manual tooth brush	20 (62.5%)	23 (67.6%)
		Electric tooth brush	4 (12.5%)	4 (11.8%)
		Both: manual toothbrush and electric toothbrush	8 (25%)	7 (20.6%)
	Interdental devices	None	12 (37.5%)	13 (38.2%)
		Floss	6 (18.8%)	10 (29.4%)
		Woodsticks	9 (28.1%)	10 (20.6%)
		Interdental brushes	0	0
		Mouthwash	3 (9.4%)	7 (20.25%)
		Other	5 (15.6%)	4 (11.8%)
	Tongue cleaning	None	20 (62.5%)	18 (52.9%)
		Toothbrush	10 (32.3%)	5 (44.1%)
		Scraper	2 (6.2%)	2 (5.8%)
		Spoon	0	0
		Other	0	0

Primary outcome

Table 3 provides a summary of the outcomes of the organoleptic scores for the test and control regimen. At baseline, no significant differences were observed. The overnight effect within the test group showed the largest decrease of 0.47 in the ORG score (P = 0.015). This effect persisted at day 7 and day 21 (P<0.05). No significant differences could be found between the time points (day 1-7, day 7-21) (see for details appendix S2). For the control groups, there was no significant difference at any of the time points. Between the test and control groups, there was a significant difference at day 1 and day 7 (P<0.05).

Table 3 Mean (standard deviation) of organoleptic score during the study for both treatment modalities (test regimen N=32 and the control regimen N=34). In addition, the outcome of the statistical analysis is presented.

N=66		Organoleptic score on a 6-point scale		Statistical analysis between groups
		Test regimen	Control regimen	
Time point	Baseline	2.41 (0.56)	2.32 (0.48)	0.604
	Day 1	1.94 (0.67) ■	2.32 (0.68)	0.015*
	Day 7	2.03 (0.74) ■	2.32 (0.68)	0.050*
	Day 21	2.13 (0.66) ■	2.15 (0.78)	0.847

* Significant difference (between groups); Mann-Whitney U test (P<0.05)
■ Significant as compared to baseline (within groups) Wilcoxon test (P<0.05)

Secondary outcomes

Table 4 shows the results from the breath analysis instruments for both regimens. The VSC data showed a non-normal distribution of the data at several time points. At baseline, both apparatuses showed no significant differences between the test and control regimens. The test group showed a significant decrease in hydrogen sulphide and methyl mercaptan at day 1. Additionally, the control group showed numerical decreases in hydrogen sulphide and methyl mercaptan, but these were not statistically significant. Dimethyl sulphide measurements during the experimental period decreased but did not reach significance in the test group. The control group observed for dimethyl sulphide a numerical increase over time with a mean of 25.50 (145.50) which was not significant. The Halimeter® readings showed a significant decrease (P<0.001) in the test group at each time point as compared to baseline.

Between the test and control regimens, a significant effect on hydrogen sulphide outcomes occurred in favour of the test regimen at all visits. Regarding dimethyl sulphide, the measurements overnight and at day 7 showed significant differences (P<0.05) in favour of the test group, while methyl mercaptan presented a significant difference only at day 21 (P = 0.012). The Halimeter® readings showed significant differences between groups at all of the time points (for details see appendix S3-S6).

Table 4 Mean (standard deviation) of oral malodour apparatus assessments (OralChroma™ and Halimeter®) during the study for both modalities (test regimen N=32 and control regimen N=34). In addition, the outcome of the statistical analysis is presented.

N=66		OralChroma™						Halimeter®	
		Test regimen			Control regimen			Test regimen	Control regimen
		H_2S	CH_3SH	$(CH_3)_2S$	H_2S	CH_3SH	$(CH_3)_2S$		
Time point	Baseline	406.75 (443.56)	205.81 (293.01)	102.06 (157.91)	299.88 (394.62)	93.15 (120.53)	38.21 (57.51)	185.69 (110.74)	140.62 (92.67)
	Day 1	61 (116.11)*■	41.97 (143.39) ■	80.34 (129.31)*	199.41 (262.60)*	64.79 (86.82)	53.47 (75.24)*	77.69 (27.67)*■	157.88 (131.57)*
	Day 7	46.94 (70.99)*■	47.62 (94.89) ■	80.59 (108.40)*	217.76 (235.07)*	62.29 (68.63)	73.32 (106.18)*■	75.69 (21.59)*■	134.59 (69.89)*
	Day 21	85.41 (255.37)*■	63.72 (138.72)*■	64.69 (111.54)	157.44 (170.54)*	84.56 (176.79)*	63.71 (126.26)	73.06 (31.67)*■	122.26 (64.62)*

H_2S= hydrogen sulphide CH_3SH= methyl mercaptan $(CH_3)_2S$= dimethyl sulphide

* Significant difference (between groups); Mann-Whitney U test (P<0.05)

■ Significant as compared to baseline (within groups) Wilcoxon test (P<0.05)

Table 5 Mean (standard deviation) of the appearance of the dorsum of the tongue; surface discoloration score and tongue coating thickness score during the study for both modalities (test regimen N=32 and control regimen N=34). In addition, the outcome of the statistical analysis is presented.

N=66		Tongue surface discoloration score		Statistical analysis between groups	Tongue coating thickness score		Statistical analysis between groups
		Test regimen	Control regimen		Test regimen	Control regimen	
Time point	Baseline	8.38 (2.74)	7.53 (3.53)	0.222	5.53 (3.21)	5.38 (3.82)	0.712
	Day 1	6.84 (3.57) ■	6.41 (3.43)	0.619	4.25 (3.04) ■	4.71 (3.94)	0.865
	Day 7	11.44 (4.05) ■	7.26 (3.86)	0.000*	6.50 (3.85)	5.03 (3.58)	0.139
	Day 21	10.66 (0.94) ■	7.59 (3.26)	0.008*	6.19 (2.30)	4.61 (3.32)	0.146

* Significant difference (between groups); Mann-Whitney U test (P<0.05)

■ Significant as compared to baseline (within groups) Wilcoxon test (P<0.05)

Tongue surfaces

Table 5 shows the mean scores for tongue surface discoloration and tongue coating thickness separated by group. The mean score for each of the 9 sections of the tongue surface was calculated. For the test regimen, the tongue surface discoloration showed a significant increase at day 1 and a decrease compared to the baseline values during the other assessments (P<0.05). At baseline, for tongue surface discoloration, there was no difference between the two groups (P = 0.222). At days 7 and 21, there was a significant difference between the two groups (P<0.05). Within group analysis showed, for the test group after a single

use, a significant decrease (P<0.05) (see for details appendix S7-S8). Between group analysis for tongue coating thickness scores showed no difference between the test and control groups at any time point.

Tongue surface discoloration and tongue coating thickness scores are also presented in percentages of the participants positive per section of the tongue surface (see appendix S9-S16) separated by test or control group. In both groups, tongue discoloration had the highest prevalence of score 2 (yellow/light brown) in the posterior and mid-dorsal regions of the tongue. Tongue coating thickness scores had the highest prevalence of score 2 (heavy thick) in the posterior dorsal sections of the tongue. Both the tongue surface discoloration score and the tongue coating thickness score showed, mostly for the anterior sections, a score of 0 (respectively pink and no coating). A tongue surface discoloration score of 3 was mostly seen in the test group during the study.

Correlation tongue surface characteristics and organoleptic scores

Table 6 presents the correlation between tongue surface discoloration/tongue coating thickness and organoleptic scores of baseline and day 21. Correlation coefficients were weak and showed no significantly relation between the tongue surface characteristics and organoleptic scores at baseline and day 21.

Table 6 Correlation between tongue surface discoloration/ tongue coating thickness and organoleptic scores for both oral hygiene regimens at baseline and day 21. Also the level of significance is indicated.

		Correlation coefficient *	P-value
Tongue surface discoloration	**Test regimen**		
	Baseline	0.216	0.235
	Day 21	0.186	0.308
	Control regimen		
	Baseline	0.324	0.062
	Day 21	0.089	0.618
Tongue coating thickness	**Test regimen**		
	Baseline	0.194	0.288
	Day 21	0.276	0.126
	Control regimen		
	Baseline	0.276	0.192
	Day 21	0.001	0.994

* Spearman's rho

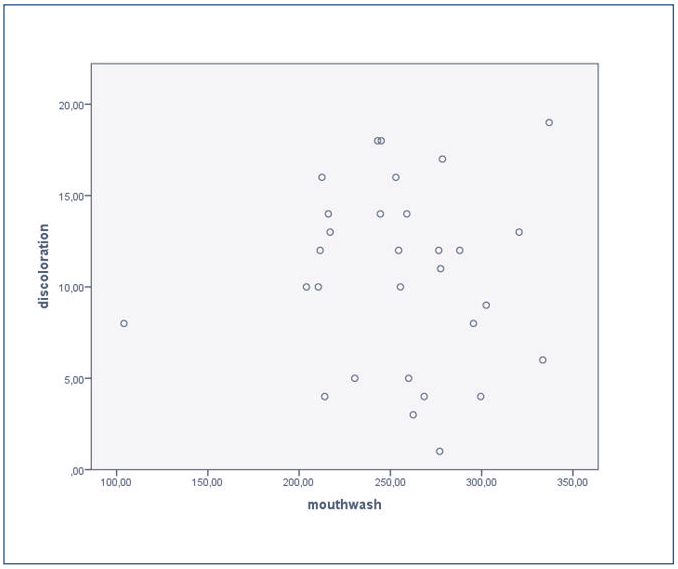


Figure 2 Used volume of mouthwash (in ml) correlated with tongue coating discoloration for the test group regimen (N=3) at day 21.

Used products

The returned bottles of mouth rinse (in millilitres) and tubes of dentifrice (in milligrams) of the study participants were evaluated. The test group used on average 255 ml volume of mouth rinse and 29 mg of tongue gel and tooth gel. The control group used 19 mg of dentifrice (for details see appendix S17).

Following the results of tongue discoloration (table 5), figure 2 provides a scatterplot of the tongue surface discoloration and mouth rinse volume (in millilitres) used in the test regimen per individual after a period of 21 days. The correlation between tongue surface discoloration and the used quantity of mouth rinse in the test group was 0.001 (P = 0.996), showing that there was no relationship between the two variables.

Patient reported outcomes

A questionnaire was administered, and the self-report findings are shown in table 7. At day 21 the test regimen participants evaluated their breath at awakening as being significantly better (P = 0.001) than the control group. In addition, the incremental difference between baseline and day 21 was in favour of the test regimen (P = 0.001). Additionally, the participants evaluated their mouth and tongue feeling at their appointments. Both these scores significantly more positive for the test group than the control group at day 21 as compared to the baseline assessments P<0.05). At day 21, the participants were able to provide feedback in the form of free comments regarding the study product used. In total, 25% (N=8) of the test regimen group mentioned that gargling with the rinse was uncomfortable. The volume of mouth rinse and the foaming effect was considered to be a negative aspect by 16% of the participants. Three participants in the control group expressed that 2 minutes of brushing is too short (for details see appendix S18-S19).

Table 7 Mean (standard deviation) of participants' self-perception regarding the morning oral malodour questionnaire as assessed at baseline and end of the study for both modalities (test regimen N=32 and control regimen N=34). In addition, the outcome of the statistical analysis is presented.

N=66	Extreme		Time point	Test regimen	Control regimen	Statistical analysis
	From	To				
1. How do you rate your breath today when you woke up?	Stale	Fresh	Baseline	2.26 (1.26)	2.46 (1.60)	0.583
			End	4.83 (2.04)	3.28 (1.69)	0.001*
			Difference	2.57 (2.36)	0.82 (1.54)	0.001*
2. At this moment, my mouth feels...	Not clean	Very clean	Baseline	2.24 (1.77)	1.91 (1.34)	0.402
			End	5.40 (2.32)	3.33 (1.73)	0.000*
			Difference	3.16 (2.31)	1.42 (1.91)	0.001*
3. At this moment, my tongue feels...	Not clean	Very clean	Baseline	2.89 (1.80)	2.90 (1.79)	0.983
			End	5.32 (1.99)	3.64 (1.65)	0.000*
			Difference	2.43 (2.02)	0.74 (2.44)	0.003*

* Significant difference (between groups); independent T-test (P<0.05)

Discussion

The current study evaluated a regimen of a mouth wash with active ingredients (9,10), and a tooth/tongue gel (11,12) combined with a tongue cleaner (14). Overall, the test regimen had a significant overnight effect but lacked a prolonged effect with respect to the primary outcome parameter of organoleptic scores. In the present study, on day 21, this effect of the product was maintained for Halimeter® readings and hydrogen sulphide concentrations.

Subjective

For oral malodour measurements, there are objective and subjective procedures. The organoleptic score is considered the gold standard, though instrumental methods have been introduced. While the apparatuses can only detect VSC gasses, the organoleptic judge can recognize and name the specific mixtures of substances (31). The organoleptic technique is inexpensive and easy to use in daily practice. The scores reflect human perception, and only a human can judge the acceptability of the mixed odours emanating from the mouth. There is no single analytic machine that can detect all volatiles in the breath that cause smells. However, humans are limited by the exponential nature of smell reception, whilst an instrument can provide concentrations of measurements on a linear scale and can delineate a halving or doubling of gas concentrations fairly precisely. To move from an organoleptic score '1' to a '2' or '2' to a '3' probably requires fivefold or more increases in gas concentrations. A mere doubling can hardly be detected by the human nose (32). Despite this, organoleptic measurements are the most popular diagnostic procedure, although it is a subjective method that renders the quantification of odour intensity notoriously difficult (33). A challenge of the organoleptic method is to ensure consistency in large numbers of

judgements (34), particularly over time. For the calibration and training of judges, standard compounds (such as n-butanol) are utilized, but these compounds cannot necessarily predict odour judges' responses to VSCs (35). Therefore, in the ideal organoleptic method, no opinion about the quality of odour is provided, only the intensity or strength of the target 'smell' is given (30).

The present study showed a numerical decrease in organoleptic scores for both groups compared to the baseline values. Two earlier studies (11,12), using organoleptic outcomes have evaluated the same tongue and tooth gel compared to no active tongue cleaning but tooth brushing alone (12). These studies found an effect after 5 minutes of use (11,12) and also after 7 days (11). In addition, the tested mouth rinse showed a positive effect compared to a negative control rinse overnight and at 21 days (10). A possible explanation for the difference with the present finding may be in the method used for organoleptic scoring. The research group conducting these studies (9,11,12) used the same 6-point scale but positioned the study participants in a room adjacent to the room with the organoleptic judges. The participants exhaled through a hole, and on the opposite side of the dividing wall, the judges smelled and rated their breath. This procedure differed fundamentally from the method used in the present study. In fact, the other studies measured exhaled breath, while we assessed the malodour of the oral cavity. Furthermore, the previously discussed studies (10-12) used the same inclusion criterion for the organoleptic score (≥2) as the current study. However, the organoleptic baseline value of the other studies was >3, while on average the participants in the current study had baseline values <2.5. This difference therefore might have influenced the potential of an effect being found in the present study. Dadamio et al. (10) also evaluated the same rinse assessed by the same 6-point scale, but the breath was scored at rest, as described by Rosenberg (35) (open mouth without breathing), while the patients counted from 1 to 11. Only a mean score of organoleptic breath was presented; it is unclear how the data were calculated or which data are presented. Nevertheless, these two methods again differed fundamentally from the method presented in the current study and from the measurements from the studies executed by the proDERM Institute (11, 12). Consequently, it is complex to devise a clear definition of success when the organoleptic score is assessed using a similar scale but is obtained by different methods.

Objective

It was suggested that objectively measuring VSCs with an apparatus is a much more reliable method of judging their strength than simple organoleptic measurement (32). It is clinically recommended to use a device to measure VSC and this is considered simple and fast (36). Most breath clinics use portable sulphide monitors, such as the Halimeter® (32). However, the Halimeter® apparatus cannot distinguish between the different sulphur gases and shows different sensitivity for each of these gases (10). The Halimeter® is more sensitive to hydrogen sulphide then to methyl mercaptan. It is least sensitive to dimethyl sulphide (37). It underestimates methyl mercaptan by approximately 31%, and markedly underestimates dimethyl sulphide concentrations by approximately 70% (37). Hydrogen sulphide is a major component of physiological oral malodour, whereas methyl mercaptan

is a cause of pathological oral halitosis (32). A recent clinical study (38) defined the following threshold values for successfully treated oral-malodour: a hydrogen sulphide value <112 ppb and a methyl mercaptan value <26 ppb. The Halimeter® manufacturer proposed that values between 90 ppb and 140 ppb should be considered normal. In the present study, the mean Halimeter® scores at baseline were 186 and 141 in the test group and control group respectively, which confirmed that the participants had oral malodour at baseline. In the course of the study in the test group the mean score dropped well below '90'.

Patient perception

According to the questionnaire (table 7), the participants' perception showed with respect to each question a significant difference. This implies that the results indicate that at day 21 there was a subjective perception of a difference between the groups while based on the organoleptic score this was not significant. An explanation of this discrepancy might be that by removing food debris from the tongue surface, taste buds are able to function more effectively and detect the subtle flavors in food (39). Result suggests that mild tongue brushing may enhance taste sensitivity of saltiness and sourness which may give a sense of freshness (40). Also self-estimation of oral malodour has been reported to be reliable and to correlate with the objective assessment. These were indeed also significantly different in favour of the test group in the present study of the compounds H_2S and CH_3SH (41).

Working mechanisms of test group regimen

With respect to the ingredients available in the tooth and tongue gel and mouthwash several possible mechanisms can be considered. First of all a reduction in the number of bacteria able to produce the malodour compounds. Secondly the neutralization of the sulphur compounds. Regarding anti-microbial products, several studies have reported the effectiveness of AmF/ SnF_2 on plaque control (10,42-44). Zinc lactate is effective by converting H_2S and CH_3SH to nonvolatile sulphides (11,45-47). Another approach to control oral malodour is remove tongue coatings and reduce the bacterial load (42-44). A previous study in 2002 has indicated that several VSCs producing bacteria have the ability to colonize on the tongue surface in periodontally healthy subjects (48,49). It found that whenever there is large amount of coating on the dorsal surface of the tongue, participants' most likely have halitosis. The added effect of the tongue surface cleaning has been shown to contribute marginally on a reduction indices of oral malodour (11).

The mechanical action of rinsing in the test group may have contributed to the reduction of oral malodour. Therefore it is not possible to definitively attribute the reduction in malodour specifically to the mouthwash ingredients. Previous results have shown that there is a significant decrease of morning bad breath after rinsing with water (19). Whether rinsing has the same impact in combination with a brushing regimen remains unknown.

Correlation tongue surfaces and organoleptic scores.

It has been suggested that a possible side effect of some of the active ingredients in the test regimen group (i.e., amine fluoride/stannous fluoride) could be stain development (50-54). Paraskevas et al. (55) showed that an amine fluoride/stannous fluoride dentifrice and

mouth rinse combination resulted in an increase in staining of tooth surfaces at 6 months compared with baseline. The present study evaluated tongue surface discoloration and observed a significant increase within the test group. At day 7 and day 21 the amount of tongue surface discoloration was significantly higher than in the control group. Consequently it was of interest whether the volume of mouthwash used had an impact on the discoloration scores. The statistical analysis however showed no correlation between tongue surface discoloration and the volume of mouthwash used (figure 2). Similarly a previous study showed a (very) weak correlation between tongue surface coating characteristics and organoleptic scores (19). Similar results were observed in the present study. Where no correlation was found between oral malodour and tongue surface characteristics. However the dental care professional should be aware that discoloration can occur following the use of a product containing amine fluoride/stannous fluoride.

The observation that the self-perception of the participants showed a significant positive effect compared to the control group supported the Halimeter® and H_2S and CH_3SH outcomes. That changes in malodour are not detected organoleptically but were apparent with the Halimeter® and OralChroma™ was also observed in a recent study in which the effect of rinsing with water or drinking water on morning oral malodour was evaluated. The same trend was observed, namely, organoleptic scores and VSC readings showed no benefit for the rinse group, while the participants' self-perception was that morning oral malodour significantly improved (19). Because patient-reported outcomes directly reflect how patients feel or function in relation to a health condition and its therapy without interpretation by healthcare professionals or anyone else, these outcomes are also meaningful and important (56).

Limitations and directions for further research

The questionnaires provided participants an opportunity to comment on the instruction of using the instructed use of mouthwash. The participants experienced gargling while sticking out the tongue as extremely unpleasant. Also due to foam formation, the volume of the mouthwash was reported to increase and become difficult to manage for some of the participants. This may have impacted on compliance.
For further research a clear definition of successful treatment of patients with morning oral malodour and the definition of corresponding parameters and methods to assess morning oral malodour appears to be essential.

Conclusion

Morning oral malodour was significantly reduced overnight for most breath parameters with the test regimen. At day 21, the prolonged effect of the test regimen was detectible with the Halimeter® readings and H_2S and CH_3SH concentrations. Participants' self-perceptions supported these findings.

Acknowledgements

The Department of Periodontology received products from GABA International. The test group used Meridol® Halitosis tooth and tongue gel, a Meridol® Halitosis tongue cleaner, and Meridol® Halitosis mouth rinse, and a Meridol® manual toothbrush. The control group used a Meridol® manual toothbrush. Currently GABA International is part of the Colgate-Palmolive Company. The authors gratefully acknowledge the second judge Sam Supranoto, who helped to conduct the study and who was responsible for operating the Halimeter® and OralChroma™ apparatuses. In addition, Esther Martin and Danielle Ekkelboom are acknowledged for their help during the clinical study.

Conflict of interest and source of funding statement

The authors have stated explicitly that there are no conflicts of interest. Van der Weijden, Slot and their research team at ACTA have previously received external advisor fees, lecturer fees or research grants from toothbrush and dentifrice manufacturers. These manufacturers have included Colgate, Dentaid, GABA, Lactona, Meda Pharma, Oral-B, Procter & Gamble, Sara Lee, Sunstar and Unilever. The study was performed with a commission from ACTA Dental Research BV. ACTA Research BV received financial support from GABA for its commitment to award this project to the Department of Periodontology of ACTA.

Authors contributions

Conception or design of the study: EVDS, GAW, DES

Analysis and/or interpretation of the data: EVDS, GAW, NLHH, DES

Drafted the manuscript: EVDS, DES

Critically revised the manuscript: GAW, NLHH, DES

All authors gave their final approval and agreed to be accountable for all aspects of the work

Supporting information

Additional supporting information may be found after the references of this article on page 108.

References

1. Dadamio J, Laleman I, Quirynen M. The role of toothpastes in oral malodour management. Monogr Oral Sci 2013; 23: 45-60.
2. European Federation of Periodontology. Guidelines for dentists and dental hygienist on efficacy of mechanical an/or chemical agents for treating halitosis. Available at: http://prevention.efp.org/wp-content/uploads/2015/12/Management-of-malodour.pdf [assessed 12 December 2016].
3. Sanz M, Bäumer A, Buduneli N, Dommisch H, Farina R, Kononen E, Linden G, Meyle J, Preshaw PM, Quirynen M, Roldan S, Sanchez N, Sculean A, Slot DE, Trombelli L, West N, Winkel E. Effect of professional mechanical plaque removal on secondary prevention of periodontitis and the complications of gingival and periodontal preventive measures: consensus report of group 4 of the 11th European Workshop on Periodontology on effective prevention of periodontal and peri-implant diseases. J Clin Periodontol 2015; 42: 214-220.
4. Slot DE, De Geest S, van der Weijden FA, Quirynen M. Treatment of oral malodour. Medium-term efficacy of mechanical and/or chemical agents: a systematic review. J Clin Periodontol 2015; 42: 303-316.
5. 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: 224-233.
6. 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: 463-473.
7. Quirynen M, Zhao H, Soers C, Dekeyser C, Pauwels M, Coucke W, Steenberghe Dv. 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.
8. Blom T, Slot DE, Quirynen M, Van der Weijden GA. The effect of mouthrinses on oral malodour: a systematic review. Int J Dent Hyg 2012; 10: 209-222.
9. Wigger-Alberti W, Gysen K, Axmann EM, Wilhelm KP. Efficacy of a new mouthrinse formulation on the reduction of oral malodour in vivo. A randomized, double-blind, placebo-controlled, 3 week clinical study. Breath Res 2010; 4: 017102.
10. Dadamio J, Van Tournout M, Teughels W, Dekeyser C, Coucke W, Quirynen M. Efficacy of different mouthrinse formulations in reducing oral malodour: a randomized clinical trial. J Clin Periodontol 2013; 40: 505-513.
11. Wilhelm D, Himmelmann A, Axmann EM, Wilhelm KP. Clinical efficacy of a new tooth and tongue gel applied with a tongue cleaner in reducing oral halitosis. Quintessence Int 2012; 43: 709-718.
12. Wilhelm D, Himmelmann A, Krause C, Wilhelm KP. Short term clinical efficacy of new meridol HALITOSIS tooth & tongue gel in combination with a tongue cleaner to reduce oral malodour. J Clin Dent 2013; 24:12-19.
13. Van der Sleen MI, Slot DE, Van Trijffel E, Winkel EG, Van der Weijden GA. Effectiveness of mechanical tongue cleaning on breath odour and tongue coating: a systematic review. Int J Dent Hyg 2010; 8: 258-268.
14. Beekmans DG, Slot DE, Van der Weijden GA. User perception on various designs of tongue scrapers: an observational survey. Int J Dent Hyg 2016 Feb 10. doi: 10.1111/idh.12204. [Epub ahead of print].
15. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. Ann Intern Med 2010; 152: 726–732.
16. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, Altman DG, Barbour V, Macdonald H, Johnston M, Lamb SE, Dixon-Woods M, McCulloch P, Wyatt JC, Chan AW, Michie S. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ 2014; 348: g1687.
17. Power and Sample Size Calculation. Available at: http://biostat.mc.vanderbilt.edu/wiki/Main/

PowerSampleSize. [accessed on 23 March 2016].
18. Dupont WD, Plummer WD: "Power and Sample Size Calculations: A Review and Computer Program". Controlled Clinical Trials 1990; 11: 116-128.
19. Van der Sluijs E, Slot DE, Bakker E, Van der Weijden GA. The effect of water on morning bad breath: a randomized clinical trial. Int J Dent Hyg 2016; 14: 124-134.
20. American Dental Association. Acceptance Program Guidelines. Products Used in the Management of Oral Malodour. American Dental Association, November 2003. Available at: http://www.ada.org/~/media/ADA/Science%20and%20Research/Files/guide_oral_malodar.pdf?la=en Assessed at [13 July 2016].
21. Gómez SM, Danser MM, Sipos PM, Rowshani B, van der Weijden GA. Tongue coating and salivary bacterial counts in healthy/gingivitis subjects and periodontitis patients. J Clin Periodontol 2001; 28: 970–978.
22. Van der Velden U. The Dutch periodontal screening index validation and it's application in The Netherlands. J Clin Periodontol 2009; 36: 1018-1024.
23. Randomness and Integrity Services Ltd. Dublin, Ireland. Available at: www.random.org. [assessed at: 13 July 2016].
24. Doty RL, Shaman P, Applebaum SL, Giberson R, Siksorski L, Rosenberg L. Smell identification ability: changes with age. Science 1984; 226: 1441–1443.
25. Rosenberg M, Septon I, Eli I et al. Halitosis measurement by an industrial sulphide monitor. J Periodontol 1991; 62: 487–489.
26. Rosenberg M, McCulloch CA. Measurement of oral malodour: current methods and future prospects. J Periodontol 1992; 63: 776–782.
27. Greenman J, Duffield J, Spencer P, Rosenberg M, Corry D, Saad S, Lenton P, Majerus G, Nachnani S, El-Maaytah M. Study on the organoleptic intensity scale for measuring oral malodour. J Dent Res 2004; 83: 81-85.
28. Mottola CA. Measurement strategies: the visual analogue scale. Decubitus 1993; 6: 56-58.
29. Creeth J, Bosma ML, Govier K. How much is a 'pea-sized amount'? A study of dentifrice dosing by parents in three countries. Int Dent J 2013; 63: 25-30.
30. Oliveira SC, Slot DE, Van der Weijden GA. What is the cause of palate lesions? A case report. Int J Dent Hyg 2013; 11: 306-309.
31. McKeown L. Social relations and breath odour. Int J Dent Hyg 2003; 1: 213-217.
32. Greenman J, Lenton P, Seemann R, Nachnani S. Organoleptic assessment of halitosis for dental professionals–general recommendations. J Breath Res 2014; 8: 017102.
33. Yaegaki K, Brunette DM, Tangerman A, Choe YS, Winkel EG, Ito S, Kitano T, Ii H, Calenic B, Ishkitiev N, Imai T. Standardization of clinical protocols in oral malodour research. J Breath Res 2012; 6: 017101.
34. Saad S, Greenman J, Shaw H. Comparative effects of various commercially available mouthrinse formulations on oral malodour. Oral Dis: 2011; 17: 180–186
35. Rosenberg M. Clinical assessment of bad breath: current concepts. J Am Dent Assoc 1996; 127: 475–482.
36. Seemann R, Duarte da Conceicao M, Filippi A, Greenman J, Lenton P, Nachnani S, Quirynen M, Roldán S, Schulze H, Sterer N, Tangerman A, Winkel EG, Yaegaki K, Rosenberg M. [Halitosis management by the general dental practitioner-results of an International Consensus Workshop*]. Swiss Dent J 2014; 124: 1205-1211.
37. Furne J, Majerus G, Lenton P, Springfield J, Levitt DG, Levitt MD. Comparison of volatile sulfur compound concentrations measured with a sulfide detector vs. gas chromatography. J Dent Res 2002; 81: 140-143.
38. Erovic Ademovski S, Mårtensson C, Persson GR, Renvert S. The effect of periodontal therapy on intra-oral halitosis: a case series. J Clin Periodontol 2016; 43: 445-452.
39. Quirynen M, Avontroodt P, Soers C, Zhao H, Pauwels M, van Steenberghe D. Impact of tongue cleansers on microbial load and taste. J Clin

Periodontol 2004; 31(7): 506-510.
40. Ohno T, Uematsu H, Nozaki S, Sugimoto K. Improvement of taste sensitivity of the nursed elderly by oral care. J Med Dent Sci 2003; 50: 101-117.
41. Pham TA. Comparison between self-estimated and clinical oral malodour. Acta Odontol Scand 2013; 71: 263-270.
42. Auschill, T. M., Hein, N., Hellwig, E., Follo, M., culean, A. & Arweiler, N. B. Effect of two antimicrobial agents on early in situ biofilm formation. J C Periodontology 2005; 32: 147–152.
43. Brecx, M., Macdonald, L. L., Legary, K., Cheang, M. & Forgay, M. G. Long-term effects of Meridol and chlorhexidine mouthrinses on plaque, gingivitis, staining, and bacterial vitality. J Dent Res 1993; 72: 1194– 1197.
44. Øgaard, B., Alm, A. A., Larsson, E. & Adolfsson, U. A prospective, randomized clinical study on the effects of an amine fluoride/stannous fluoride toothpaste/mouthrinse on plaque, gingivitis and initial caries lesion development in orthodontic patients. Eur J Orthod 2006; 28: 8–12.
45. Roldán S, Winkel EG, Herrera D, Sanz M, Van Winkelhoff 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 lacebo-controlled study. J Clin Periodontol 2003; 30: 427-434.
46. Van Steenberghe D, Avontroodt P, Peeters W, Pauwels M, Coucke W, Lijnen A, Quirynen M. Effect of different mouthrinses on morning breath. J Periodontol 2001; 72: 1183-1191.
47. Schmidt NF, Missan SR, Tarbet WJ. The correlation between organoleptic mouth-odor ratings and levels of volatile sulfur compounds. Oral Surg Oral Med Oral Pathol 1978; 45: 560-567.
48. Barzan A, Abeer S. Oral halitosis and oral hygiene practices among dental students J Bagh Coll Dentistry 2007; 91: 72-79.
49. Kishi M, Kimura S, Dhare-Nemoto Y. Oral malodour and periodonto-pathic microorganisms in the tongue coating of periodontally healthy subjects. Dentistry in Japan 2002; 38: 24-28.
50. Leverett DH, McHugh WD, Jensen EO. Effect of daily rinsing with stannous fluoride on plaque and gingivitis: final report. J Dent Res 1984; 63: 1083–1086.
51. Wolff LF, Pihlstrom BL, Bakdash MB, Aeppli D M, Bandt CL. Effect of toothbrushing with 0.4% stannous fluoride and 0.22% sodium fluoride gel on gingivitis for 18 months. J Am Dent Assoc 1989; 119: 283–289.
52. Boyd RL, Chun YS. Eighteenmonth evaluation of the effects of a 0.4% stannous fluoride gel on gingivitis in orthodontic patients. Am J Orthod Dentofacial Orthop 1994; 105: 35–41.
53. Mankodi S, Lopez M, Smith I, Petrone DM, Petrone ME, Chaknis P, Proskin HM. Comparison of two dentifrices with respect to efficacy for the control of plaque and gingivitis, and with respect to extrinsic tooth staining: a six-month clinical study on adults. J Clin Dent 2002; 13: 228–233.
54. Paraskevas S, Danser MM, Timmerman MF, Van der Velden U, Van der Weijden, GA. Effect of a combination of amine/stannous fluoride dentifrice and mouthrinse in periodontal maintenance patients. J Clin Periodontol 2004; 31: 177–183.
55. Paraskevas S, Versteeg PA, Timmerman MF, Van der Velden U, Van der Weijden GA. The effect of a dentifrice and mouth rinse combination containing amine fluoride/stannous fluoride on plaque and gingivitis: a 6-month field study. J Clin Periodontol 2005; 32: 757-764.
56. Higgins JPT, Green S. CCHB Cochrane Handbook for Systematic Reviews of Interventions 2012. Chapter 3 Special topics. Available at: http://handbook.cochrane.org/chapter_17/17_patient_reported_outcomes.htm [assessed at: 13 July 2016].

Appendix S1 Details of the lifestyle rules to which the participants were requested to adhere to avoid interference of food constituents or other scented products in morning oral malodour assessments.

Time before morning oral malodour assessment	Participants were requested...
>48 h	not to eat spicy food or drink alcohol
>8 h	to abstain from eating and drinking prior to the assessment
>8 h	not to use scented products (shampoo, body lotion, perfume, aftershave, lipstick, make-up)
>8 h	not to use, e.g., chewing gum and peppermint drops
>3 h	not to drink water
Morning of the assessment	to refrain from all oral hygiene practices and to come to the clinic while fasting

Appendix S2 Mean (standard deviation) and within group analysis for the primary outcome of the test regimen N=32 and the control regimen N=34.

N=66		Organoleptic score on a 6-point scale			
		Test regimen	**P-value Analysis between appointment**	**Control regimen**	**P-value Analysis between appointment**
Clinical assessment	**Baseline**	2.41 (0.56)	0.001 ◆	2.32 (0.48)	0.317
	Day 1 (overnight)	1.94 (0.67)	0.180	2.32 (0.68)	0.922
	Day 7	2.03 (0.74)	0.675	2.32 (0.68)	0.560
	Day 21	2.13 (0.66)		2.15 (0.78)	
		Test regimen	**P-value Analysis compared to baseline**	**Control regimen**	**P-value Analysis compared to baseline**
Change from baseline	**Day 1 (overnight)**	0.47 (0.67)	0.001 ■	0.00 (0.60)	0.317
	Day 7	0.38 (0.94)	0.040 ■	0.00 (0.65)	0.400
	Day 21	0.28 (0.68)	0.048 ■	0.18 (0.63)	1.000

◆ Significant between appointments (within groups) Wilcoxon test (P<0.05)
■ Significant as compared to baseline (within groups) Wilcoxon test (P<0.05)

Appendix S3 Mean (standard deviation) of OralChroma™ apparatus assessments (hydrogen sulphide) during the study. Test regimen N=32 and control regimen N=34, according to the assigned regimen, are presented.

N=66		Hydrogen sulphide			
		Test regimen	**P-value Analysis between appointment**	**Control regimen**	**P-value Analysis between appointment**
Clinical assessment	**Baseline**	406.75 (443.56)	0.000 ◆	299.88 (394.62)	0.126
	Day 1 (overnight)	61 (116.11)	0.768	199.41 (262.60)	0.925
	Day 7	46.94 (70.99)	0.864	217.76 (235.07)	0.221
	Day 21	85.41 (255.37)		157.44 (170.54)	
		Test regimen	**P-value Analysis compared to baseline**	**Control regimen**	**P-value Analysis compared to baseline**
Change from baseline	**Day 1 (overnight)**	345.75 (480.63)	0.001 ■	100.47 (377.62)	0.126
	Day 7	359.81 (407.41)	0.000 ■	82.12 (438.86)	0.726
	Day 21	321.34 (517.41)	0.000 ■	142.44 (393.37)	0.129

◆ Significant between appointments (within groups) Wilcoxon test (P<0.05)
■ Significant as compared to baseline (within groups) Wilcoxon test (P<0.05)

Appendix S4 Mean (standard deviation) and within group analysis for the OralChroma™ apparatus assessments (methyl mercaptan) during the study. Test regimen N=32 and control regimen N=34, according to the assigned regimen, are presented.

N=66		Methyl mercaptan			
		Test regimen	**P-value Analysis between appointment**	**Control regimen**	**P-value Analysis between appointment**
Clinical assessment	**Baseline**	205.81 (2.93)	0.002 ◆	93.15 (120.53)	0.289
	Day 1 (overnight)	41.97 (143.39)	0.139	64.79 (86.82)	0.878
	Day 7	47.62 (94.89)	0.567	62.29 (68.63)	0.477
	Day 21	63.72 (138.72)		84.56 (176.79)	
		Test regimen	**P-value Analysis compared to baseline**	**Control regimen**	**P-value Analysis compared to baseline**
Change from baseline	**Day 1 (overnight)**	163.84 (338.66)	0.002 ■	28.35 (183.04)	0.289
	Day 7	158.19 (287.29)	0.001 ■	30.85 (128.31)	0.405
	Day 21	142.09 (328.33)	0.012 ■	8.59 (220.91	0.102

◆ Significant between appointments (within groups) Wilcoxon test (P<0.05)
■ Significant as compared to baseline (within groups) Wilcoxon test (P<0.05)

Appendix S5 Mean (standard deviation) and within group analysis for the OralChroma™ apparatus assessments (dimethyl sulphide) during the study. Test regimen N=32 and control regimen N=34, according to the assigned regimen, are presented.

N=66		Dimethyl sulphide			
		Test regimen	P-value Analysis between appointment	Control regimen	P-value Analysis between appointment
Clinical assessment	Baseline	102.06 (157.91)	0.399	38.21 (57.51)	0.382
	Day 1 (overnight)	80.34 (129.31)	0.378	53.47 (75.24)	0.537
	Day 7	80.59 (108.40)	0.236	73.32 (106.18)	0.204
	Day 21	64.69 (111.54)		63.71 (126.26)	
		Test regimen	P-value Analysis compared to baseline	Control regimen	P-value Analysis compared to baseline
Change from baseline	Day 1 (overnight)	21.72 (210.33)	0.399	-15.27 (100.75)	0.382
	Day 7	21.47 (177.28)	0.674	-35.12 (122.42)*	0.043 ■
	Day 21	37.38 (202.17)	0.264	-25.50 (145.50)	0.649

◆ Significant between appointments (within groups) Wilcoxon test (P<0.05)
■ Significant as compared to baseline (within groups) Wilcoxon test (P<0.05)

Appendix S6 Mean (standard deviation) and within group analysis for the Halimeter® apparatus assessments during the study. Test regimen N=32 and control regimen N=34, according to the assigned regimen, are presented.

N=66		Halimeter®			
		Test regimen	P-value Analysis between appointment	Control regimen	P-value Analysis between appointment
Clinical assessment	Baseline	185.69 (110.74)	0.000 ◆	140.62 (92.67)	0.584
	Day 1 (overnight)	77.69 (27.67)	0.910	157.88 (131.57)	0.594
	Day 7	75.69 (21.59)	0.270	134.59 (69.89)	0.161
	Day 21	73.06 (31.67)		122.26 (64.62)	
		Test regimen	P-value Analysis compared to baseline	Control regimen	P-value Analysis compared to baseline
Change from baseline	Day 1 (overnight)	108.00 (103.74)	0.000 ■	-17.26 (131.70)	0.584
	Day 7	110.00 (101.56)	0.000 ■	6.03 (94.33)	0.713
	Day 21	112.63 (113.01)	0.000 ■	18.35 (103.58)	0.317

◆ Significant between appointments (within groups) Wilcoxon test (P<0.05)
■ Significant as compared to baseline (within groups) Wilcoxon test (P<0.05)

Appendix S7 Mean (standard deviation) of the appearance of the dorsum of the tongue surface discoloration score during the study. Test regimen N=32 and control regimen N=34, according to the assigned regimen, are presented.

N=66		Tongue surface discoloration score			
		Test regimen	P-value Analysis between appointment	Control regimen	P-value Analysis between appointment
Clinical assessment	Baseline	8.38 (2.74)	0.013 ◆	7.53 (3.53)	0.021 ◆
	Day 1 (overnight)	6.84 (3.57)	0.000 ◆	6.41 (3.43)	0.116
	Day 7	11.44 (4.05)		7.26 (3.86)	
	Day 21	10.66 (0.94)	0.125	7.59 (3.26)	0.612
		Test regimen	P-value Analysis compared to baseline	Control regimen	P-value Analysis compared to baseline
Change from baseline	Day 1 (overnight)	1.53 (2.99)	0.000 ■	1.12 (2.77)	0.021 ■
	Day 7	3.06 (4.21)	0.000 ■	0.26 (3.52)	0.403
	Day 21	2.28 (4.89)	0.000 ■	0.06 (3.75)	0.763

◆ Significant between appointments (within groups) Wilcoxon test (P<0.05)
■ Significant as compared to baseline (within groups) Wilcoxon test (P<0.05)

Appendix S8 Mean (standard deviation) and within group analysis for the appearance of the dorsum of the tongue coating thickness score during the study. Test regimen N=32 and control regimen N=34, according to the assigned regimen, are presented.

N=66		Tongue coating thickness score			
		Test regimen	P-value Analysis between appointment	Control regimen	P-value Analysis between appointment
Clinical assessment	Baseline	5.53 (3.21)	0.012 ◆	5.38 (3.82)	0.206
	Day 1 (overnight)	4.25 (3.04)	0.001 ◆	4.71 (3.94)	0.435
	Day 7	6.50 (3.85)		5.03 (3.58)	
	Day 21	6.19 (2.30)	0.388	4.61 (3.32)	0.240
		Test regimen	P-value Analysis compared to baseline	Control regimen	P-value Analysis compared to baseline
Change from baseline	Day 1 (overnight)	1.28 (2.73)	0.012 ■	0.68 (2.88)	0.206
	Day 7	0.97 (3.40)	0.139	0.35 (2.16)	0.333
	Day 21	0.66 (0.79)	0.347	0.82 (3.99)	0.314

◆ Significant between appointments (within groups) Wilcoxon test (P<0.05)
■ Significant as compared to baseline (within groups) Wilcoxon test (P<0.05)

Appendix S9 Percentages of subjects with the percentages for tongue surface discoloration score according to section of the tongue at visit 1. Test regimen (T) (N=32) control regimen (C) (N=34).

T 0=3% 1=44% 2=53% 3=0%	**C** 0=3% 1=50% 2=47% 3=0%	**T** 0=0% 1=22% 2=75% 3=3%	**C** 0=6% 1=9% 2=55% 3=0%	**T** 0=3% 1=44% 2=53% 3=0%	**C** 0=6% 1=47% 2=47% 3=0%
T 0=13% 1=75% 2=13% 3=0%	**C** 0=32% 1=50% 2=18% 3=0%	**T** 0=6% 1=47% 2=47% 3=0%	**C** 0=12% 1=53% 2=35% 3=0 %	**T** 0=13% 1=75% 2=13% 3=0%	**C** 0=35% 1=47% 2=18% 3=0%
T 0=97% 1=3% 2=0% 3=0%	**C** 0=94% 1=6% 2=0% 3=0%	**T** 0=94% 1=6.3% 2=0% 3=0%	**C** 0=94% 1=6% 2=0% 3=0%	**T** 0=94% 1=6% 2=0% 3=0%	**C** 0=94% 1=6% 2=0% 3=0%

Appendix S10 Percentages of subjects with the percentages for tongue coating thickness score according to section of the tongue at visit 1. Test regimen (T) (N=32) control regimen (C) (N=34).

T 0=22% 1=56% 2=22%	**C** 0=32% 1=38% 2=29%	**T** 0=0% 1=47% 2=53%	**C** 0=6% 1=41% 2=53%	**T** 0=22% 1=59% 2=18%	**C** 0=35% 1=5% 2=29%
T 0=53% 1=43.8% 2=3%	**C** 0=56% 1=38% 2=6%	**T** 0=25% 1=56% 2=19%	**C** 0=32% 1=35% 2=32%	**T** 0=56% 1=41% 2=3%	**C** 0=56% 1=38% 2=6%
T 0=97% 1=3 % 2=0%	**C** 0=100% 1=0% 2=0%	**T** 0=100% 1=0% 2=0%	**C** 0=100% 1=0% 2=0%	**T** 0=100% 1=0% 2=0%	**C** 0=100% 1=0% 2=0%

Appendix S11 Percentages of subjects with the percentages for tongue surface discoloration score according to section of the tongue at visit 2. Test regimen (T) (N=32) control regimen (C) (N=34).

T 0=13% 1=56% 2=31% 3=0%	**C** 0=15% 1=50% 2=35% 3=0%	**T** 0=0% 1=47% 2=41% 3=13%	**C** 0=0% 1=44% 2=56% 3=0%	**T** 0=13% 1=56.3% 2=31% 3=0%	**C** 0=15% 1=50% 2=35% 3=0%
T 0=28% 1=56% 2=15% 3=0%	**C** 0=41% 1=53% 2=6% 3=0%	**T** 0=19% 1=56% 2=25% 3=0%	**C** 0=24% 1=47% 2=29% 3=0%	**T** 0=28% 1=56% 2=16% 3=0%	**C** 0=41% 1=53% 2=6% 3=0%
T 0=100% 1=0% 2=0% 3=0%	**C** 0=97% 1=3% 2=0% 3=0%	**T** 0=100% 1=0% 2=0% 3=0%	**C** 0=97% 1=3% 2=0% 3=0%	**T** 0=100% 1=0% 2=0% 3=0%	**C** 0=97% 1=3% 2=0% 3=0%

Appendix S12 Percentages of subjects with the percentages for tongue coating thickness score according to section of the tongue at visit 2. Test regimen (T) (N=32) control regimen (C) (N=34).

T 0=44% 1=44% 2=13%	**C** 0=41% 1=32% 2=27%	**T** 0=6% 1=59% 2=34%	**C** 0=0% 1=56% 2=44%	**T** 0=41% 1=47% 2=13%	**C** 0=41% 1=47% 2=13%
T 0=63% 1=38% 2=0%	**C** 0=74% 1=18% 2=9%	**T** 0=41% 1=41% 2=19%	**C** 0=41% 1=44% 2=15%	**T** 0=63% 1=38% 2=0%	**C** 0=74% 1=18% 2=9%
T 0=100% 1=0% 2=0%	**C** 0=97% 1=3% 2=0%	**T** 0=97% 1=3% 2=0%	**C** 0=97% 1=3% 2=0%	**T** 0=100% 1=0% 2=0%	**C** 0=97% 1=3% 2=0%

Appendix S13 Percentages of subjects with the percentages for tongue surface discoloration score according to section of the tongue at visit 3. Test regimen (T) (N=32) control regimen (C) (N=34).

T 0=0% 1=19% 2=69% 3=13%	**C** 0=12% 1=44% 2=44% 3=0%	**T** 0=0% 1=13% 2=38% 3=50%	**C** 0=0% 1=47% 2=47% 3=6%	**T** 0=19% 1=69% 2=13% 3=0%	**C** 0=12% 1=44% 2=44% 3=0%
T 0=13% 1=34% 2=44% 3=9%	**C** 0=29% 1=53% 2=18% 3=0%	**T** 0=3% 1=41% 2=34% 3=22%	**C** 0=24% 1=47% 2=29% 3=22%	**T** 0=13% 1=34% 2=47% 3=6%	**C** 0=29% 1=53% 2=17% 3=0%
T 0=84% 1=16% 2=0% 3=0%	**C** 0=97% 1=3% 2=0% 3=0%	**T** 0=84% 1=16% 2=0% 3=0%	**C** 0=94% 1=3% 2=3% 3=0%	**T** 0=84% 1=16% 2=0% 3=0%	**C** 0=97% 1=3% 2=0% 3=0%

Appendix S14 Percentages of subjects with the percentages for tongue coating thickness score according to section of the tongue at visit 3. Test regimen (T) (N=32) control regimen (C) (N=34).

T 0=19% 1=53% 2=28%	**C** 0=32% 1=41% 2=27%	**T** 0=3% 1=41% 2=56%	**C** 0=9% 1=47% 2=44%	**T** 0=19% 1=53% 2=28%	**C** 0=32% 1=41% 2=27%
T 0=38% 1=50% 2=13%	**C** 0=56% 1= 44% 2=0%	**T** 0=25% 1=50 % 2=25%	**C** 0=38% 1=32% 2=29%	**T** 0=38% 1=47% 2=16%	**C** 0=56% 1=44% 2=0%
T 0=91% 1=9% 2=0%	**C** 0=100% 1=0% 2=0	**T** 0=91% 1=9% 2=0%	**C** 0=100% 1=0% 2=0%	**T** 0=94% 1=6% 2=0%	**C** 0=100% 1=0% 2=0%

Appendix S15 Percentages of subjects with the percentages for tongue surface discoloration score according to section of the tongue at visit 4. Test regimen (T) (N=32) control regimen (C) (N=34).

T 0=3% 1=31% 2=53% 3=13%	**C** 0=6% 1=47% 2=47% 3=0%	**T** 0=0% 1=28% 2=31% 3=41%	**C** 0=0% 1=38% 2=53% 3=9%	**T** 0=3% 1=31% 2=53% 3=13%	**C** 0=6% 1=47% 2=47% 3=0%
T 0=16% 1=28% 2=41% 3=16%	**C** 0=29% 1=56% 2=15% 3=0%	**T** 0=19% 1=25% 2=28% 3=28%	**C** 0=9% 1=59% 2=32% 3=0%	**T** 0=19% 1=25% 2=41% 3=16%	**C** 0=27% 1=59% 2=15% 3=0%
T 0=91% 1=9% 2=0% 3=0%	**C** 0=97% 1=3% 2=0% 3=0%	**T** 0=91% 1=9% 2=0% 3=0%	**C** 0=97% 1=3% 2=0% 3=0%	**T** 0=91% 1=9% 2=0 3=0%	**C** 0=97% 1=3% 2=0% 3=0%

Appendix S16 Percentages of subjects with the percentages for tongue coating thickness score according to section of the tongue at visit 4. Test regimen (T) (N=32) control regimen (C) (N=34).

T 0=34% 1=31% 2=34%	**C** 0=32% 1=44% 2=21%	**T** 0=6% 1=47% 2=47%	**C** 0=5.9% 1=60% 2=35%	**T** 0=31% 1=34% 2=34%	**C** 0=32% 1=47% 2=1%
T 0=38% 1=47% 2=16%	**C** 0=32% 1=56% 2=12%	**T** 0=34% 1=28% 2=38%	**C** 0=65% 1=29% 2=6%	**T** 0=41% 1= 42% 2=16%	**C** 0=65% 1=29% 2=6%
T 0=94% 1=6% 2=0%	**C** 0=100% 1=0% 2=0%	**T** 0=94% 1=6% 2=0%	**C** 0=100% 1=0% 2=0%	**T** 0=94% 1=6% 2=0%	**C** 0=100% 1=0% 2=0%

Appendix S17 Mean (SD) of used products according to the assigned group regimen, test regimen N=32 and control regimen N=34. Amount of mouthwash tongue gel, tooth gel and toothpaste in ml.

	Test regimen		**Control regimen**
	Mouthwash N=30 ○	Tongue and tooth gel N=30 ○○	Toothpaste N=30 □
Baseline	453.90 (0.70)	107.98 (0.62)	104.90 (0.24)
End	198.85 (46.47)	78.79 (11.65)	85.6 (7.54)
Used products	255.05 (46.51)	29.19 (11.75)	19.30 (7.57)

○ = 3 missing values
○○ = 2 missing values
□ = 4 missing values

Appendix S18 Participants' feedback regarding the test regimen products and instructions at day 21 (N=13).

Feedback		
Mouth rinse	1	Too much volume of mouthwash, and it foams too much during gargling
	2	Gargling was unpleasant. The experience was really unpleasant when gargling. However, the rinsing was not a problem. There was nausea after one week of using the products, usually in the morning but also spontaneously in the afternoon. This felt very uncomfortable
	3	Too much volume of mouthwash and was not able to gargle. 10 s of gargling with tongue out is very difficult, and I did not achieve this for 10 s or more. The taste of the mouth rinse became really bad. The tongue gel did not spread easily on my tongue surface
	4	The gargling of the volume of mouth rinse was uncomfortable. I would never do this in general
		Gargling was very uncomfortable. Suggestion: you could reduce the amount of mouthwash to obtain more compliance
	5	Mouthwash was quite tasty
	6	The mouthwash was very enjoyable. Gargling was very difficult because the solution foamed too much
	7	After using the products, it left residue in my mouth, which was unpleasant. Spreading the tongue gel on the tongue was tricky
	8	Gargling was not pleasant
	9	The mouthwash foamed, and my clothes got dirty easily. During the research, I even had the impression that the effect (of fresh breath) disappeared
Tongue cleaner and tongue gel	10	Tongue cleaning was too short with too much volume of mouth rinse, and it became a mess
	11	The tongue cleaning with tongue gel was difficult. The gel did not spread well, and I felt that this was not used at all during brushing
	12	Tongue cleaner with tongue gel was difficult to use. On paper, the instructions appeared to be simple, but in practice, it was difficult
	13	Cleaning the tongue I found very to be difficult, and in general, I would never do this

Appendix S19 Participants' feedback regarding the control regimen products and instructions at day 21 (N=4).

Feedback		
Toothpaste	**1**	The toothpaste was tasteless and watery
	2	The toothpaste disappeared quickly during brushing. Two minutes of brushing was short; Normally, I brush for at least three minutes
	3	The toothpaste foamed, but it has no intense mint flavour, which I prefer more. The fresh feeling disappeared quickly after brushing
Toothbrush	**4**	The toothbrush was too big
	5	Toothbrush was very nice. Two minutes of brushing was too short
	6	Two minutes of brushing was too short