STATISTICAL ANALYSIS PLAN

This document contains confidentiality statements that are not relevant for this publicly available version

A Randomized, Controlled, Examiner-Blind Clinical Study Investigating the Effects of a Dentifrice Containing 67% Sodium Bicarbonate When Used Twice Daily for 12 Weeks on Gingivitis Treatment and Plaque Removal

Protocol Number: 300029

Phase: N/A

Document History

Document	Version Date	Summary of Changes (New analysis or Change in planned analysis)
Original Analysis Plan	08-Jun-2023	Not applicable

Table of contents

Document History 3
Table of contents 4
List of Tables 5
Abbreviations 6
1 Summary of Key Protocol Information 7
1.1 Study Design 8
1.2 Study Objectives 11
1.3 Treatments 13
1.4 Sample Size Calculation 13
2 Planned Analyses 14
2.1 Interim Analysis 14
2.2 Final Analyses 14
3 Considerations for data analyses and Data Handling Conventions 14
3.1 Baseline Definition 14
3.2 Subgroups/Stratifications 14
3.3 Centers Pools 14
3.4 Timepoints and Visit Windows 14
4 Data Analysis 15
4.1 Populations for Analysis 15
4.1.1 Subject Disposition 15
4.1.2 Protocol Deviations 16
4.1.3 Analysis Populations 16
4.2 Subject Demographics and Other Baseline Characteristics 17
4.2.1 Demographic Characteristics 17
4.2.2 General Medical History 18
4.3 Treatments (Study Product, Rescue Medication, other Concomitant Therapies, Compliance) 18
4.3.1 Study Product Compliance and Exposure 18
4.3.2 Prior and Concomitant Medication 19
4.4 Analysis of Efficacy 19
4.4.1 Primary Efficacy Endpoint 19
4.4.2 Secondary Efficacy Endpoints 21

4.4.3 Handling of Missing Values/Censoring/Discontinuations 27
4.5 Analysis of Safety 27
4.5.1 Adverse Events and Serious Adverse Events 28
4.5.2 Other Safety Variables 28
4.6 Analysis of Other Variables 29
4.6.1 Repeatability of the Examiner 29
5 Changes to the Protocol Defined Statistical Analysis Plan 30
Appendix 1: List of Data Displays 32

List of Tables

Table 1-1 Schedule of Activities 10
Table 1-2 Study Objectives and Endpoints 12
Table 1-3 Investigational/Study Product Supplies 13
Table 4-1 Bleeding Index Scoring System 20
Table 4-2 Modified Gingival Index Scoring System 24
Table 4-3 Turesky Plaque Index Scoring System 25
Table 5-1 Changes to Protocol Defined Statistical Analysis Plan 31

Abbreviations

Abbreviation	Term
AE	Adverse Event
ANCOVA	Analysis of Covariance
BDRM	Blind Data Review Meeting
BI	Bleeding Index
BOP	Bleeding on Probing
CI	Confidence Interval
COVID-19	Coronavirus Disease of 2019
CRF	Case Report Form
CS	Compound Symmetry
GSK CH	GlaxoSmithKline Consumer Healthcare
ICF	Informed Consent Form
LOCF	Last Observation Carried Forward
LS	Least Square
MedDRA	Medical Dictionary for Regulatory Activities
MFC	Master Formulation Code
MGI	Modified Gingival Index
mITT	Modified Intent-to-Treat
MMRM	Mixed Model for Repeated Measures
OHT	Oral Hard Tissue
OST	Oral Soft Tissue
PP	Per-Protocol
PT	Preferred Term
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SD	Standard Deviation
SE	Standard Error
SOC	System Organ Class
TEAE	Treatment Emergent Adverse Event
TPI	Turesky Plaque Index
w/w	Weight for Weight
WHODD	World Health Organization Drug Dictionary

The purpose of this Statistical Analysis Plan (SAP) is to describe the planned analyses and outputs to be included in the Clinical Study Report for Protocol 300029 (Version 3.0, dated 10-May-2023).

1 Summary of Key Protocol Information

This will be a single center, controlled, single blind (examiner blind), randomized, two-treatment arm, parallel study. Study subjects will be over 18 years old, non-smokers, in good general health with clinically measurable levels of gingivitis (defined as those with 10-30% bleeding sites on probing) that meet all study criteria at the Screening and Baseline visits.

Approximately 200 subjects (n=100 per group) will be randomized to ensure 188 evaluable subjects (94 per group) complete the study.

This study will consist of 5 study visits: Screening, Baseline, Week 3, 6 and 12. Gingivitis will be assessed using the Modified Gingival Index (MGI) and the Bleeding Index (BI). Plaque will be assessed by the Turesky modification of the Quigley Hein Index (Turesky Plaque Index [TPI]). All evaluable teeth (in relation to the inclusion/ exclusion general dentition criteria) will be assessed.

At the Screening visit (Visit 1; in relation to subject screening), subjects will provide their written informed consent to participate in the study. Demographics, medical history, and current medications will be recorded, followed by an oral examination (oral soft tissue (OST), oral hard tissue (OHT) examination and dentition exclusions) and a gross gingival assessment. Eligible subjects will be given washout toothpaste and toothbrush to use in between the Screening and Baseline Visits. They will also be issued with a study diary on product usage. For the subjects who consent to use the electronic study diary application during the study period, the e-diary app will be installed on the subject's own mobile device and the diary data will be collected remotely through the study app.

Within 14±3 days of Screening (Visit 1), eligible subjects will return to the site for the Baseline visit (Visit 2) with overnight plaque (subjects will be instructed to abstain from oral hygiene for 12 hours [+6 hours; -2 hours] i.e., overnight immediately before the visit). At the Baseline visit, subjects will undergo, in the following order, a full OST and OHT examination, assessments of gingival inflammation (MGI), gingival bleeding (BI), pocket depth and supra-gingival plaque (TPI). Subjects with bleeding on probing (BOP; derived from BI assessment), pocket depth or TPI score outside the study range will be discontinued from the study at this visit. Eligible subjects will then be randomized to study product and undergo a supervised brushing, where they will be instructed to brush for a least a minute at site with their assigned study product, after which they will be instructed to continue using their product twice daily (morning and evening until their next visit), at least a minute/each time.

After using the study dentifrice for 3, 6, and 12 weeks, subjects will return to the study site (Visits 3, 4, and 5, respectively) with overnight plaque (subjects will be instructed to abstain

from overnight toothbrushing for 12 hours [+6 hours; -2 hours] immediately before each assessment visit), at approximately the same time of day as the Baseline visit. The study dentifrice and the diary will be reviewed to determine treatment compliance. Subjects will have a full OST examination and then undergo, in the following order, MGI, BI, and TPI assessments. At Visit 5, subjects will also have a full OHT examination, return all study supplies and have a dental prophylaxis if deemed appropriate by the investigator or examiner. The study diary application will be deactivated at Visit 5 for those subjects who had the application installed on their phones.

At Visits 2, 3, 4, and 5, repeatability data will be generated for MGI and TPI assessments from replicate examinations on the same subject (the subject for the MGI repeat assessment and the subject for the TPI assessment can be different). Depending on subject visit scheduling, every effort will be made to complete one repeatability examination for each clinical measure during each clinical assessment day, that is, at least one MGI and one TPI repeat assessment on each assessment day. Repeatability examinations will be separated by a minimum of 10 minutes and, where possible, separated by another subject. Due to the invasive nature of the BI assessment, it is not feasible to conduct an accurate repeatability assessment for this index.

Adverse events (AEs) and incidents will be recorded from informed consent and at the end of each study visit.

1.1 Study Design

This will be a single center, controlled, single blind (examiner blind), randomized, two treatment arm, parallel study in volunteers with clinically measurable levels of gingivitis, defined as those with 10-30% bleeding sites on probing. Study subjects will be over 18 years old, non-smokers, in good general health that meet all study criteria at the Screening and Baseline visits.

Approximately 200 subjects (n=100 per group) will be randomized to ensure 188 evaluable subjects (94 per group) complete the study.

This study will consist of 5 study visits (Screening, Baseline, Week 3, 6, and 12). At Visit 1, Screening, after signing informed consent, subjects will undergo an OST examination, an OHT examination and a gross assessment of gingival health, in addition to the standard procedures (inclusion, exclusion, medical history, demographics, prior/current medications) to assess eligibility for the study. Subjects will return to site 14±3 days after the Screening Visit, for Visit 2, Baseline.

At the Baseline Visit, subjects will undergo, in the following order, a full OST/ OHT examination followed by assessments of gingival inflammation (MGI), gingival bleeding (BI) including BOP which is derived from BI assessment, periodontitis status (pocket depth), and supra-gingival plaque using the TPI. Subjects with BOP, pocket depth or TPI score outside the study range will be discontinued from the study at this visit. To control inter-examiner

variability, the same examiner will be used throughout the study for all clinical assessments (MGI, TPI, BI).

After all clinical assessments, subjects will be instructed to brush for at least a minute, in their usual manner, with their assigned study product and then instructed to continue using their assigned product twice daily (morning and evening), for 3, 6 and 12 weeks after which they will return for their Visits 3, 4 and 5 assessments respectively. Subjects will record all brushing events in the diary provided (paper or electronic), and this will be reviewed by the site at each study visit.

After the Week 12 visit, study closeout procedures (return of study product etc.) will take place and subjects may undergo a prophylaxis if it is deemed necessary by the examiner. Safety and oral tolerability of the study products will be monitored over the 12-week treatment period by review of reported AEs and Incidents.

Table 1-1 presents the schedule of activities.

Table 1-1 Schedule of Activities

Procedure/Assessment	Screening	Study Period			
	Visit 1	Visit 2 Day 0 Baseline[1]	Visit 3 Day 21±3 (Week 3)[1]	Visit 4 Day 42±4 (Week 6)[1]	Visit 5 Day 84±4 (Week 12)[1]
Informed consent	X				
Medical history	X				
Demographics	X				
Current/prior/concomitant medication review	X	X	X	X	X
OST examination	X	X	X	X	X
OHT examination	X	X			X
Gross gingival assessment[2]	X				
Review of Inclusion/exclusion criteria	X	X			
Subject eligibility	X	X			
Subject continuance			X	X	X
Dispense washout toothpaste, toothbrush, and washout diary[3]	X				
Return washout toothpaste/toothbrush		X			
Modified Gingival Index (MGI) assessment		X	X	X	X
MGI repeat assessment[4]		X	X	X	X
Bleeding Index (BI) assessment[5]		X	X	X	X
Pocket Depth assessment		X			
Plaque disclosure and plaque index (TPI) assessments		X	X	X	X
TPI repeat assessment[6]		X	X	X	X
Randomization		X			
Dispense study products, new toothbrush, and timer		X			
Dispense study diary & instructions on how to complete		X			
Oral hygiene instruction[7] & supervised brushing with study dentifrice		X	X	X	
Review diary completion[8]		X	X	X	X

Washout period 14±3 days

Procedure/Assessment	Screening	Study Period			
	Visit 1	Visit 2 Day 0 Baseline[1]	Visit 3 Day 21±3 (Week 3)[1]	Visit 4 Day 42±4 (Week 6)[1]	Visit 5 Day 84±4 (Week 12)[1]
Collect study products, assess compliance, and return to subject			X	X	
Return study products (end of study)[9]					X
End of study dental prophylaxis (optional)					X
Adverse events review[10]	X	X	X	X	X
Medical device incidents review[8]		X	X	X	X
Study conclusion					X

Abbreviations: OST: Oral Soft Tissue, OHT: Oral Hard Tissue, MGI: Modified Gingivitis Index, BI: Bleeding Index, TPI: Turesky Plaque Index, BOP: bleeding on probing

Footnotes:

1. Subjects will abstain from overnight toothbrushing for a minimum of 12hrs (+6hr, -2hr) immediately prior to the assessment visits (Visits 2, 3, 4 and 5).
2. In relation to the general dentition inclusion/exclusion criteria.
3. An electronic study diary will be installed on the personal mobile phones of subjects who consent to use it during the study period. A paper diary will be dispensed to subjects who are unable to adhere to completing their electronic diary.
4. One MGI repeat assessment during each clinical assessment day.
5. BOP and number of bleeding sites are derived from BI assessment.
6. One TPI repeat assessment during each clinical assessment day.
7. Instruction.
8. Electronic or paper diary, dependent on the type of diary that subject receives.
9. If installed at screening the electronic study diary will be deactivated at the last visit. Diaries (paper or electronic diary) will be reviewed for completion and then collected.
10. Adverse Events (AEs), Serious Adverse Events (SAEs) and medical device incidents collected immediately after subject provides consent to participate in the study by the completion of the Informed Consent Form (ICF). Medical device in this study is the test product.

1.2 Study Objectives

Study objectives and endpoints are defined in Table 1-2.

Table 1-2 Study Objectives and Endpoints

Objectives	Endpoints
Primary Objective	**Primary Endpoint**
To evaluate the efficacy of a dentifrice containing 67% w/w sodium bicarbonate and 0.310% w/w sodium fluoride, compared to its baseline, for the assessment of gingivitis, as measured by the number of bleeding sites after 12 weeks twice daily toothbrushing.	Number of bleeding sites at 12 weeks, compared to baseline
Secondary Objectives	**Secondary Endpoints**
Efficacy	
To evaluate the efficacy of a dentifrice containing 67% w/w sodium bicarbonate and 0.310% w/w sodium fluoride, compared to negative control, for the assessment of gingivitis, as measured by the number of bleeding sites after 12 weeks twice daily toothbrushing.	Number of bleeding sites at 12 weeks, compared to negative control
To evaluate the efficacy of a dentifrice containing 67% w/w sodium bicarbonate and 0.310% w/w sodium fluoride for the assessment of gingivitis, as measured by the number of bleeding sites after 3 and 6 weeks twice daily toothbrushing.	Number of bleeding sites at 3 and 6 weeks, compared to baseline and negative control
To evaluate the efficacy of a dentifrice containing 67% w/w sodium bicarbonate and 0.310% w/w sodium fluoride for the assessment of gingivitis, as measured by the Bleeding Index (BI) after 3, 6 and 12 weeks twice daily toothbrushing.	Mean BI at 3, 6 and 12 weeks, compared to baseline and negative control
To evaluate the efficacy of a dentifrice containing 67% w/w sodium bicarbonate and 0.310% w/w sodium fluoride for the assessment of gingivitis, as measured by the Modified Gingival Index (MGI) after 3, 6 and 12 weeks twice daily toothbrushing.	Mean MGI at 3, 6 and 12 weeks, compared to baseline and negative control
To evaluate the efficacy of a dentifrice containing 67% w/w sodium bicarbonate and 0.310% w/w sodium fluoride for the assessment of plaque accumulation, as measured by the Turesky Plaque Index (TPI, overall and interproximal) after 3, 6, and 12 weeks twice daily brushing.	Mean TPI (overall and interproximal) at 3, 6 and 12 weeks, compared to baseline and negative control
Safety	
To assess the oral tolerability of a dentifrice containing 67% w/w sodium bicarbonate and 0.310% w/w sodium fluoride, over 12 weeks twice daily use.	Treatment-emergent adverse events over 12 weeks

1.3 Treatments

Table 1-3 presents the study products.

Table 1-3 Investigational/Study Product Supplies

Product Description	Test Product	Reference Product
Product Name	Corsodyl Original Toothpaste	Colgate Cavity Protection Toothpaste
Pack Design	Carton of 6 over-wrapped tubes	
Dispensing Details	One carton – baseline visit	
Product Master Formulation Code (MFC)	CCI (Commercial Product)	Commercial Product
Dose Application	Full ribbon of toothpaste on head of toothbrush provided	
Route of Administration	Oral	
Usage Instructions	Subjects will brush their teeth for at least a minute twice a day (morning and evening)	
Return Requirements	All used/unused samples to be returned	

Detailed instructions for the return of study product/study supplies for the accountability checks and subsequent destruction which will be provided by the sponsor during the study in time for study close out visit.

1.4 Sample Size Calculation

A sample size of approximately 100 subjects per treatment group (before presumed 6% dropout rate) is needed to provide at least 90% power to achieve statistical significance (at the 5%, two-sided level) for the primary outcome (change from baseline at 12 weeks in the number of bleeding sites in the sodium bicarbonate group). The assumed treatment efficacy estimates for this calculation comes from the sponsor clinical study (CCI 2012) report where the change (SD) from baseline at 12 weeks (without pre-prophylaxis) was -14.7 (9.8) bleeding sites.

For the secondary objective (change from baseline at 12 weeks in the number of bleeding sites in the bicarbonate group compared to the negative control group), data from the sponsor clinical study (CCI 2012) focusing on the non-smoking strata and using a t-test comparing their respective mean changes (SD of change) of 15.5 (10.1) *vs.* 11.6 (9.1) indicates power will exceed 80%. (PASS 2019 Power Analysis and Sample Size Software (2019). NCSS, LLC. Kaysville, Utah, USA, ncss.com/software/pass). The assumed (from the sponsor clinical study CCI 2012) +0.7 correlation of the number of bleeding sites at baseline to week 12 helps attain that power.

2 Planned Analyses

2.1 Interim Analysis

No interim analysis is planned.

2.2 Final Analyses

The final planned primary analyses will be performed after the completion of the following sequential steps:

1. All subjects have completed the study as defined in the protocol.
2. All required database cleaning activities including any external data reconciliation have been completed and the database has been locked.
3. All criteria for unblinding the randomization codes have been met and the randomization codes have been distributed.

3 Considerations for data analyses and Data Handling Conventions

3.1 Baseline Definition

For all endpoints, the baseline value will be the Day 0 (Visit 2) pre-brushing assessment with a non-missing value.

Unless otherwise stated, if baseline data is missing no derivation will be performed and the baseline value will be set to missing.

3.2 Subgroups/Stratifications

No subgroups or stratification factors are defined in this study.

3.3 Centers Pools

Since this is a single center study, pooling of centers is not applicable.

3.4 Timepoints and Visit Windows

The timepoints and visits for this study are defined in Table 1-1 “Schedule of Activities”. Any deviation from the study schedule may be reviewed on a case-by-case basis at the Blind Data Review Meeting (BDRM) to determine whether the data should be excluded from the Per Protocol (PP) population.

4 Data Analysis

Data analysis will be performed by CCI with oversight from Haleon. The statistical analysis software used will be SAS version 9.4 or higher.

Prior to database closure a BDRM will be conducted in which various aspects of the trial will be discussed and agreed.

One aspect that will be considered prior to or during the BDRM is the assessment of the number of subjects who have dropped or discontinued from the study due to pandemic related events (e.g. Coronavirus Disease of 2019 [COVID-19]) and the potential need of a sensitivity analysis. Any major changes to planned analyses will need an amendment to the SAP.

Except as described below, all listings will be produced for all randomized subjects.

4.1 Populations for Analysis

4.1.1 Subject Disposition

Screen failures are defined as subjects who consent to participate in the clinical study but are not subsequently randomized. An enrolled subject is a subject who has signed informed consent and is eligible to proceed beyond the screening visit.

The number of subjects screened, enrolled, and randomized will be presented in Table 14.1.1.

The number and percentage of screen failure subjects (subjects not randomized) with reasons why subjects are not randomized including percentage of subjects not randomized due to the COVID-19 pandemic will be displayed. Percentages for screen failure subjects will be based on the total number of subjects screened.

The number and percentage of randomized subjects who complete and discontinue the study, broken down by reason for discontinuation including percentage of subjects discontinued/withdrew from study due to COVID-19 pandemic, by study product and overall will also be displayed. The percentages will be based on the number of subjects randomized.

Table 14.1.1 will also present the number and percentage of subjects in each of the defined analysis populations (as defined in Section 4.1.3) by study product and overall. Percentages will be based on the number of subjects randomized.

Subject disposition including demographic data (age, sex, race and ethnicity), screening date, study product start date and time, subject status (completer, Yes/No), study completion /withdrawal date, duration (in days) in the study (defined as [(date of completion or withdrawal – start date of study product use) + 1], and the primary reason for withdrawal will be listed (Listing 16.2.1.1) by study product.

Subject disposition information will be listed for non-randomized subjects (Listing 16.2.1.2), displaying subject number, demographic information (age, sex, race and ethnicity), enrolled

(Yes/No), screening date, reason for screen failure, any further details of reason for screen failure and discontinuation status due to COVID-19 pandemic.

4.1.2 Protocol Deviations

Protocol deviations will be tracked by the study team throughout the conduct of the study. Data will be reviewed prior to unblinding and closure of the database to ensure all important deviations are captured and categorized. Subjects with important protocol deviations liable to influence the efficacy outcomes will be excluded from the PP population. Subjects may also be identified as having important protocol deviations not leading to exclusion from the PP population.

Important deviations of the protocol procedures may include, but will not necessarily be limited to the following:

- Consent procedures
- Inclusion/Exclusion criteria
- Concomitant medication/therapy
- Laboratory assessments
- Study procedures
- Randomization procedures
- Study drug dosing/study product administration/study product compliance
- Visit schedule/interval
- Other

The specific details of the important protocol deviations will be listed in Protocol Deviation Management Plan, and the assessment process will be specified in the Blind Data Review Plan. Subjects with important protocol deviations will be identified at the BDRM.

The number and percentage of subjects with at least one important protocol deviation, subjects with important protocol deviations not leading to exclusion from the PP population with reasons for deviations and subjects with important protocol deviations leading to exclusion from the PP population with reasons for deviations will be presented by study product and overall for all randomized subjects (Table 14.1.2) and listed in Listing 16.2.2.1.

All protocol deviations collected on the protocol deviation case report form page will be listed in Listing 16.2.2.2. The listing will present date of deviation, type of deviation, and deviation description.

4.1.3 Analysis Populations

Five analysis populations are defined.

Population	Definition / Criteria	Analyses Evaluated
Safety	Comprise of all randomized subjects who receive at least one dose of study product. This population will be based on the product the subject actually received.	• Demographics • Safety
Modified Intent-to-Treat (mITT)	Comprise of all randomized subjects who complete at least one dose of study product and have at least one post baseline clinical performance assessment. This population will be based on the study product to which the subject was randomized. Any subject who receives a randomization number will be considered to have been randomized.	• Demographics • Compliance • Efficacy
Per-Protocol (PP)	Comprise of all randomized subjects who do not have any protocol deviations that could confound the interpretation of analyses conducted on the mITT. This population will be based on the study product to which the subject was randomized. Protocol deviations that may exclude subjects from the PP population are defined in Section 4.1.2 (Protocol Deviations).	• Efficacy
MGI Repeatability	Comprise of all subjects who have at least one repeat MGI clinical assessment at any visit.	• Repeatability for MGI
TPI Repeatability	Comprise of all subjects who have at least one repeat TPI clinical assessment at any visit.	• Repeatability for TPI

NOTES:
Please refer to Appendix 1: List of Data Displays, which details the population to be used for each display being generated.

The numbers of subjects included in each of the analysis populations will be summarized for all screened subjects (Table 14.1.1). Subjects excluded from any of the analysis populations will be listed in Listing 16.2.3.1.

The primary population for assessment of efficacy will be the mITT Population. A PP analysis will be performed on the primary endpoint only if more than 10% of mITT subjects are excluded from the PP Population. A decision on whether a PP analysis will be performed will be made prior to study unblinding (release of the randomization codes).

4.2 Subject Demographics and Other Baseline Characteristics

4.2.1 Demographic Characteristics

Descriptive statistics (number of subjects [n], mean, standard deviation [SD], median, minimum, and maximum for continuous variables and frequency count [n] and percentage [%] of subjects for categorical variables) will be presented for demographic variables by study

product and overall. These variables include age, gender, race, and ethnicity and will be presented for the Safety population (Table 14.1.3.1) and the mITT population (Table 14.1.3.2).

Demographic information will be listed (Listing 16.2.4.1) for all randomized subjects.

4.2.2 General Medical History

Medical and surgical history (in the last year) including allergies or drug sensitivity will be listed in Listing 16.2.4.2, with start date and end date or ongoing at the start of study product.

4.3 Treatments (Study Product, Rescue Medication, other Concomitant Therapies, Compliance)

Randomization details will be listed, including the randomization number, the planned study product, the actual study product the subject was randomized to and the randomization date (Listing 16.1.7.1).

The study product kit allocations will be listed (Listing 16.1.7.2), including kit number and study product information.

4.3.1 Study Product Compliance and Exposure

Compliance data will be summarized for the mITT population and will be assessed by number of brushings.

Number of brushings, brushing compliance (%), number of missed brushings and number of additional brushings will be summarized using descriptive statistics by cumulative visit and by study product. The number and percentage of subjects <80%, between 80% - 120% and >120% compliant in each study product will also be presented by cumulative visit and by study product (Table 14.2.1).

Number of brushings is defined as: [(date of Visit *x* – date of Visit 2) multiplied by 2 – number of missed brushings + number of additional brushings], where *x* = 3, 4 and 5, respectively.

Brushing compliance (%) is defined as: [100 x (Actual number of brushings / Expected number of brushings)], where expected number of brushings is defined as: [(date of Visit *x* – date of Visit 2) multiplied by 2].

Study product compliance (number of brushings, brushing compliance [%], number of missed brushings and number of additional brushings) will be listed in Listing 16.2.5.1 for all randomized subjects by study product.

Supervised study product application (subject number, visit, date and time of the supervised procedure) will be listed (Listing 16.2.5.2) for all randomized subjects.

4.3.2 Prior and Concomitant Medication

Prior medications/treatments, including prescription and non-prescription drugs, dietary supplements and herbal remedies, taken in the last 30 days and prior to signing the informed consent form, will be documented in the case report form (CRF). The prior and concomitant medications will be coded using a validated medication dictionary, World Health Organization Drug Dictionary (WHODD).

Prior medications and prior non-drug treatments will be listed by subject, with drug name, WHODD Drug Synonym, reason, route, dose, frequency, start date and end date both relative to study product start date (Listing 16.2.4.3) for all randomized subjects. Prior medications are defined as those which stopped before the first use of the study product.

Concomitant medications and concomitant non-drug treatments/significant non-drug therapies taken during treatment will be listed similarly (Listing 16.2.4.4) for all randomized subjects with either ongoing or end date displayed. Concomitant medications are defined as medications that started or stopped on or after the first use of the study product, or are ongoing.

Unknown dates will not be imputed. However, if the start or stop date is unknown, then it will be assumed to be concomitant medication unless the partial start date or stop date indicates differently.

4.4 Analysis of Efficacy

The primary population for assessment of efficacy will be the mITT population. No repeat assessment data is to be used in any efficacy analyses.

Statistical analyses of the endpoints will be performed using the change from baseline, in order to easily present direct estimates for the within-treatment mean changes from baseline (e.g. the primary endpoint).

Descriptive summary statistics will be presented for the actual and change from baseline values at each time point.

All p-values presented will be two-sided and assessed at the 5% significance level. A sequential testing strategy will be used to adjust for multiplicity for the comparison between test and reference products in the number of bleeding sites at Week 12. This will only be assessed for confirmatory evidence if the primary endpoint (change from baseline within the test product at Week 12) achieves a statistically significant reduction. There will be no further adjustments for multiplicity for any other secondary endpoints.

4.4.1 Primary Efficacy Endpoint

4.4.1.1 Primary Efficacy Endpoint Definition

The primary endpoint for the study is the change from baseline number of bleeding sites at 12 weeks in the Test Product group in the mITT population.

The BI will be assessed on the facial and lingual gingival surfaces of each scorable tooth (7-7 in each arch). Three scores (according to the scale in Table 4-1 below) should be recorded buccally/labially (distal, body and mesial sites) and three scores lingually/palatally (distal, body and mesial sites). At each visit, each subject therefore has up to 168 evaluable tooth sites (i.e. 6 measurements on up to 28 scorable teeth).

Table 4-1 Bleeding Index Scoring System

Score	Description
0	No bleeding after 30 seconds
1	Bleeding observed within 30 seconds of probing
2	Bleeding observed immediately on probing
9	Missing tooth or not qualified tooth

The number of bleeding sites for each subject at each visit is calculated as the number of evaluable tooth sites with a BI score of either 1 or 2, to give a count between 0 and 168.

Descriptive statistics (n, mean SD, standard error [SE], median, minimum, and maximum) will be presented for the number of bleeding sites at each time point (absolute value, change from baseline and percent change from baseline) in Table 14.2.2.1.1 for all subjects in the mITT population by study product. Raw means (± SE) of the number of bleeding sites at each time point will be plotted by study product in Figure 14.2.2.1.5 for all subjects in the mITT population.

Additionally, the number and percentage of subjects whose number of bleeding sites has reduced (improved), remained the same and increased (worsened), respectively compared to baseline, will be presented by study product at each time point in Table 14.2.2.1.3 for the mITT population.

Individual data for the number of bleeding sites will be listed for each subject by study product group and visit in Listing 16.2.6.1 for all randomized subjects.

4.4.1.2 Statistical Hypothesis, Model, and Method of Analysis

The following null and alternative hypotheses will be evaluated:

- H_0: There is no difference in the number of bleeding sites from baseline to Week 12 in the Test Product group.
- H_1: There is a treatment effect on the number of bleeding sites at Week 12, compared to baseline, in the Test Product group.

A Mixed Model for Repeated Measures (MMRM) will be used, with change from baseline in number of bleeding sites as the dependent variable, the categorical time point (Week 3, Week 6, or Week 12), treatment group (Test Product or Reference Product) and time point × treatment group interaction as fixed effects, and baseline number of bleeding sites as a covariate. Subject will be included as a repeated measure with an unstructured covariance matrix. The Kenward

Roger degrees of freedom approach will be applied. If unstructured covariance matrix does not converge, the following covariance structures will be applied in order until convergence is achieved: 1) Toeplitz, 2) compound symmetry (CS).

Using the above model, the LS mean change from baseline number of bleeding sites will be presented for the Test Product at Week 12, along with 95% CI and p-value, testing for a non-zero change from baseline (Table 14.2.2.2.1).

The assumption of normality and homogeneity of variance in the MMRM will be investigated. In case of violation of these assumptions, a Wilcoxon Signed Rank Test (within product comparisons with baseline) will be performed to support the MMRM results.

4.4.1.3 Supportive Analyses

A PP analysis will be performed on the number of bleeding sites if there is more than a 10% difference in the number of subjects between the PP and mITT populations. A decision on whether a PP analysis will be performed will be made prior to study unblinding (release of randomization codes).

A summary of the mean number of bleeding sites at each time point will be presented for the PP population in Table 14.2.2.1.2. A summary of the number and percentage of subjects at each time point whose number of bleeding sites has improved/remained the same/worsened compared to baseline will be presented for the PP population in Table 14.2.2.1.4. Statistical analysis of the mean number of bleeding sites will be presented for the PP population in Table 14.2.2.2.2. In addition, the mean number of bleeding sites at each time point will be plotted by study product for all subjects in the PP population (Figure 14.2.2.1.6).

In order to explore the sensitivity of results to the effect of missing data, the MMRM model described in Section 4.4.1.2 will be repeated using the LOCF approach for the mITT population (Table 14.2.2.2.3) and the PP population (Table 14.2.2.2.4).

4.4.2 Secondary Efficacy Endpoints

All secondary endpoint analyses will be performed on the mITT population. Other than the sequential testing strategy for comparing the change from baseline in the number of bleeding sites between products (see Section 4.4), there will be no adjustment for multiplicity performed for any other secondary endpoints detailed in this section.

4.4.2.1 Change from Baseline in Number of Bleeding Sites

The number of bleeding sites for each subject at each visit is calculated using the BI score, and descriptive statistics and frequency counts will be presented for the mITT population as detailed in Section 4.4.1.1.

4.4.2.1.1 Statistical Hypothesis, Model and Method of Analysis

For comparisons to baseline within each treatment group, the following null and alternative hypotheses will be evaluated for post-baseline time points at Week 3 and Week 6 (for the Test Product group and the Reference Product group), and at Week 12 (for the Reference Product group) respectively:

- H_0: There is no difference in the number of bleeding sites from baseline.
- H_1: There is a treatment effect on the number of bleeding sites compared to baseline.

For comparisons to negative control, the following null and alternative hypotheses will be evaluated for post-baseline time points at Week 3, Week 6, and Week 12, respectively:

- H_0: There is no difference in change from baseline in the number of bleeding sites in the Test Product group compared to the Reference Product group.
- H_1: There is a difference in change from baseline in the number of bleeding sites in the Test Product group compared to the Reference Product group.

The results from the MMRM described in Section 4.4.1.2 will be used to test the above hypotheses.

The LS mean change from baseline number of bleeding sites will be presented for each study product at each post-baseline time point with corresponding 95% CIs and p-values to test for a non-zero change from baseline.

The difference in LS mean change from baseline number of bleeding sites (Test Product – Reference Product) will also be presented at each post-baseline time point with corresponding 95% CIs and p-values to test for a non-zero difference (Table 14.2.2.2.1).

The assumption of normality and homogeneity of variance in the MMRM will be investigated. In case of violation of these assumptions, a Wilcoxon Rank Sum test (between products comparisons) and Wilcoxon Signed Rank Test (within product comparisons with baseline) will be performed to support the MMRM results.

The MMRM model will be repeated using the LOCF approach for the mITT population (Table 14.2.2.2.3) and the PP population (Table 14.2.2.2.4).

4.4.2.2 Change from Baseline in Bleeding Index (BI)

The BI will be assessed at each evaluable tooth site for each subject at each visit as described in Section 4.4.1.1, according to the scoring system in Table 4-1. The BI score for each subject at each visit is calculated as the sum of index values over all evaluable tooth sites, divided by the number of evaluable tooth sites. Evaluable tooth sites are those scored as 0, 1 or 2, so the BI score has a range of 0 to 2.

Descriptive statistics (n, mean, SD, SE, median, minimum, and maximum) will be presented for the BI score at each time point (absolute value, change from baseline and percent change

from baseline) in Table 14.2.3.1.1. Raw means (± SE) of the BI score at each time point will be plotted by study product in Figure 14.2.3.1.3 for all subjects in the mITT population.

Additionally, the number and percentage of subjects whose BI score has reduced (improved), remained the same and increased (worsened), respectively compared to baseline, will be presented by study product at each time point in Table 14.2.3.1.2 for the mITT population.

Individual data for the BI assessment at each evaluable tooth site will be listed for each subject in Listing 16.2.6.2.1, and individual data for the derived BI score will be listed for each subject in Listing 16.2.6.2.2, by study product group and visit for all randomized subjects.

4.4.2.2.1 Statistical Hypothesis, Model and Method of Analysis

For comparisons to baseline within each treatment group, the following null and alternative hypotheses will be evaluated for post-baseline time points at Week 3, Week 6 and Week 12, respectively:

- H_0: There is no difference in the BI score from baseline.
- H_1: There is a treatment effect on the BI score compared to baseline.

For comparisons to negative control, the following null and alternative hypotheses will be evaluated for post-baseline time points at Week 3, Week 6, and Week 12, respectively:

- H_0: There is no difference in change from baseline in the BI score in the Test Product group compared to the Reference Product group.
- H_1: There is a difference in change from baseline in the BI score in the Test Product group compared to the Reference Product group.

A MMRM will be used as described in Section 4.4.1.2, with change from baseline in BI score as the dependent variable and baseline BI score as a covariate instead of baseline number of bleeding sites.

Using the above model, the LS mean change from baseline BI score will be presented for each study product at each post-baseline time point with corresponding 95% CIs and p-values to test for a non-zero change from baseline.

The difference in LS mean change from baseline BI score (Test Product – Reference Product) will also be presented at each post-baseline time point with corresponding 95% CIs and p-values to test for a non-zero difference (Table 14.2.3.2.1).

The assumption of normality and homogeneity of variance in the MMRM will be investigated. In case of violation of these assumptions, a Wilcoxon Rank Sum test (between products comparisons) and Wilcoxon Signed Rank Test (within product comparisons with baseline) will be performed to support the MMRM results.

The MMRM model will be repeated using the LOCF approach for the mITT population (Table 14.2.3.2.2).

4.4.2.3 Change from Baseline in Modified Gingival Index (MGI)

The MGI will be assessed on the facial and lingual surfaces of each scorable tooth (7-7 in each arch). Two scores (according to the scale in Table 4-2 below) should be recorded buccally/labially (papilla and margin) and two scores lingually/palatally (papilla and margin). At each visit, each subject therefore has up to 112 evaluable tooth sites (i.e. 4 measurements on up to 28 scorable teeth).

Table 4-2 Modified Gingival Index Scoring System

Score	Description
0	Absence of inflammation
1	Mild inflammation; slight change in colour, little change in colour; little change in texture of any portion of the marginal or papillary gingival unit
2	Mild inflammation; criteria as above but involving the entire marginal or papillary gingival unit
3	Moderate inflammation; glazing, redness, edema, and/ or hypertrophy of the marginal or papillary gingival unit
4	Severe inflammation; marked redness, edema and/ or hypertrophy of the marginal or papillary gingival unit, spontaneous bleeding, congestion, or ulceration
9	Missing tooth or not qualified tooth

The MGI score for each subject at each visit is calculated as the sum of index values over all evaluable tooth sites, divided by the number of evaluable tooth sites. Evaluable tooth sites are those scored as 0, 1, 2, 3 or 4 so the MGI score has a range of 0 to 4.

Descriptive statistics (n, mean, SD, SE, median, minimum, and maximum) will be presented for the MGI score at each time point (absolute value, change from baseline and percent change from baseline) in Table 14.2.4.1.1. Raw means (± SE) of the MGI score at each time point will be plotted by study product in Figure 14.2.4.1.3 for all subjects in the mITT population.

Additionally, the number and percentage of subjects whose MGI score has reduced (improved), remained the same and increased (worsened), respectively compared to baseline, will be presented by study product at each time point in Table 14.2.4.1.2 for the mITT population.

Individual data for the MGI assessment at each evaluable tooth site will be listed for each subject in Listing 16.2.6.3.1, and individual data for the derived MGI score will be listed for each subject in Listing 16.2.6.3.2, by study product group and visit for all randomized subjects.

4.4.2.3.1 Statistical Hypothesis, Model and Method of Analysis

For comparisons to baseline within each treatment group, the following null and alternative hypotheses will be evaluated for post-baseline time points at Week 3, Week 6, and Week 12, respectively:

- H_0: There is no difference in the MGI score from baseline.
- H_1: There is a treatment effect on the MGI score compared to baseline.

For comparisons to negative control, the following null and alternative hypotheses will be evaluated for post-baseline time points at Week 3, Week 6, and Week 12, respectively:

- H_0: There is no difference in change from baseline in the MGI score in the Test Product group compared to the Reference Product group.
- H_1: There is a difference in change from baseline in the MGI score in the Test Product group compared to the Reference Product group.

A MMRM will be used as described in Section 4.4.1.2, with change from baseline in MGI score as the dependent variable and baseline MGI score as a covariate instead of baseline number of bleeding sites.

Using the above model, the LS mean change from baseline MGI score will be presented for each study product at each post-baseline time point with corresponding 95% CIs and p-values to test for a non-zero change from baseline.

The difference in LS mean change from baseline MGI score (Test Product – Reference Product) will also be presented at each post-baseline time point with corresponding 95% CIs and p-values to test for a non-zero difference (Table 14.2.4.2.1).

The assumption of normality and homogeneity of variance in the MMRM will be investigated. In case of violation of these assumptions, a Wilcoxon Rank Sum test (between products comparisons) and Wilcoxon Signed Rank Test (within product comparisons with baseline) will be performed to support the MMRM results.

The MMRM model will be repeated using the LOCF approach for the mITT population (Table 14.2.4.2.2).

4.4.2.4 Change from Baseline in Turesky Plaque Index (TPI)

The TPI will be used to assess plaque on all gradable teeth meeting the inclusion/exclusion criteria (natural teeth only). The TPI will be assessed on the facial and lingual surfaces of each scorable tooth (7-7 in each arch). Three scores (according to the scale in Table 4-3 below) should be recorded buccally/labially (distal, body and mesial sites) and three scores lingually/palatally (distal, body and mesial sites). At each visit, each subject therefore has up to 168 evaluable tooth sites (i.e. 6 measurements on up to 28 scorable teeth).

Table 4-3 Turesky Plaque Index Scoring System

Score	Description
0	No plaque
1	Separate flecks of plaque at the cervical margin
2	Thin continuous band of plaque (up to 1 mm) at the cervical margin
3	Band of plaque wider than 1 mm but covering < 1/3 of the tooth surface
4	Plaque covering ≥ 1/3 but < 2/3 of the tooth surface
5	Plaque covering ≥ 2/3 of the tooth surface

Score	Description
9	Missing tooth or not qualified tooth

The overall TPI score for each subject at each visit is calculated as the sum of index values over all evaluable tooth sites, divided by the number of evaluable tooth sites. Evaluable tooth sites are those scored as 0, 1, 2, 3, 4 or 5 so the overall TPI score has a range of 0 to 5.

The interproximal TPI score for each subject at each visit is calculated as the sum of index values over all evaluable distal and mesial tooth sites, divided by the number of evaluable distal and mesial tooth sites. Similar to the overall TPI score, the interproximal TPI score has a range of 0 to 5.

Descriptive statistics (n, mean, SD, SE, median, minimum, and maximum) will be presented at each time point (absolute value, change from baseline and percent change from baseline) in Table 14.2.5.1.1 for overall TPI score and in Table 14.2.5.1.3 for interproximal TPI score. Raw means (± SE) at each time point will be plotted by study product in Figure 14.2.5.1.5 for overall TPI score and in Figure 14.2.5.1.6 for interproximal TPI score, for the mITT population.

Additionally, the number and percentage of subjects whose score has reduced (improved), remained the same and increased (worsened), respectively compared to baseline, will be presented by study product at each time point in Table 14.2.5.1.2 (overall TPI score) and Table 14.2.5.1.4 (interproximal TPI score) for the mITT population.

Individual data for the TPI assessment at each evaluable tooth site will be listed for each subject in Listing 16.2.6.4.1, and individual data for the derived TPI score (overall and interproximal, respectively), will be listed for each subject in Listing 16.2.6.4.2, by study product group and visit for all randomized subjects.

4.4.2.4.1 Statistical Hypothesis, Model and Method of Analysis

For comparisons to baseline within each treatment group, the following null and alternative hypotheses will be evaluated for post-baseline time points at Week 3, Week 6, and Week 12, for the overall TPI score and interproximal TPI score, respectively:

- H_0: There is no difference in the TPI score from baseline.
- H_1: There is a treatment effect on the TPI score compared to baseline.

For comparisons to negative control, the following null and alternative hypotheses will be evaluated for post-baseline time points at Week 3, Week 6 and Week 12, for the overall TPI score and interproximal TPI score, respectively:

- H_0: There is no difference in change from baseline the TPI score in the Test Product group compared to the Reference Product group.
- H_1: There is a difference in change from baseline in the TPI score in the Test Product group

Separate MMRMs will be used as described in Section 4.4.1.2, with change from baseline in TPI score (overall and interproximal, respectively) as the dependent variable and baseline TPI score (overall and interproximal, respectively) as a covariate instead of baseline number of bleeding sites.

Using the above model, the LS mean change from baseline TPI score will be presented for each study product at each post-baseline time point with corresponding 95% CIs and p-values to test for a non-zero change from baseline.

The difference in LS mean change from baseline TPI score (Test Product – Reference Product) will also be presented at each post-baseline time point with corresponding 95% CIs and p-values to test for a non-zero difference (Table 14.2.5.2.1 for overall TPI, and Table 14.2.5.2.3 for interproximal TPI).

The assumption of normality and homogeneity of variance in the MMRM will be investigated. In case of violation of these assumptions, a Wilcoxon Rank Sum test (between products comparisons) and Wilcoxon Signed Rank Test (within product comparisons with baseline) will be performed to support the MMRM results.

The MMRM model will be repeated using the LOCF approach for the mITT population (Table 14.2.5.2.2 for overall TPI, and Table 14.2.5.2.4 for interproximal TPI).

4.4.3 Handling of Missing Values/Censoring/Discontinuations

For the primary analyses, missing data will not be replaced or imputed. Subjects who withdraw from the study prematurely will be included in the statistical analyses up to the point of discontinuation. Subjects who withdraw will not be replaced. In the event that a subject has no individual BI, MGI or TPI assessments available at a given visit, the number of bleeding sites, BI score, MGI score or TPI score will also be considered as missing for that visit.

Experience in previous, similar studies suggest missing data will not be prevalent. The SAS Proc Mixed implementation of the repeated measures analysis of covariance (ANCOVA) that will be used in almost all analyses, implicitly imputes any missing data based on the non-missing observations. The basic theory on which Proc Mixed is based holds even with missing data as long as the missing data are missing at random.

An additional analysis will explore the sensitivity of results to the effect of missing data on the primary and secondary efficacy endpoints. The sensitivity analysis will use the same MMRM model, but with LOCF for missing data. Only post-baseline results may be carried forward.

4.5 Analysis of Safety

The safety profile of the study products will be assessed with respect to adverse events (AEs), incidents and OST/OHT abnormalities in this oral health study.

4.5.1 Adverse Events and Serious Adverse Events

All AEs will be reviewed by the Clinical Research Scientist or Designee prior to database lock and unblinding, and will be coded to a system organ class (SOC) and preferred term (PT) using the Medical Dictionary for Regulatory Activities (MedDRA).

AEs will be classified as oral and non-oral on the AE page of eCRF.

Treatment emergent adverse events (TEAEs) are defined as new AEs that occur on or after the first study product use (if this date is missing, a suitable alternative will be used e.g. date of randomization). AEs with an onset date/time prior to the first study product use will be considered as non-treatment emergent.

The following summary tables and listings will be presented by study product group.

- Table of TEAEs by SOC and PT (Table 14.3.1.1)
- Table of TEAEs by Oral/Non-Oral and PT (Table 14.3.1.2)
- Table of treatment related TEAEs by SOC and PT (Table 14.3.1.3)
- Table of treatment related TEAEs by Oral/Non-Oral and PT (Table 14.3.1.4)
- Table of AEs related to COVID-19 by SOC and PT (Table 14.3.1.5)
- Table of non-serious TEAEs by SOC and PT (Table 14.3.1.6) [only produced if there are more than 5 AEs]
- Table of serious TEAEs by SOC and PT (Table 14.3.1.7) [only produced if there are more than 5 serious TEAEs]
- Listing of all AEs (Listing 16.2.7.1 for all randomized subjects; Listing 16.2.7.2 for non-randomized subjects)
- Listing of deaths (Listing 14.3.2.1)
- Listing of non-fatal SAEs (Listing 14.3.2.2)
- Listing of TEAEs leading to study or product withdrawal (Listing 14.3.2.3)
- Listing of TEAEs classified as Oral (Listing 14.3.2.4)
- Listing of all AEs related to COVID-19 (Listing 16.2.7.3 for all screened subjects)

All incidents will be listed in Listing 16.2.7.4.

In the event that there is nothing to report, a null table or listing will be produced.

4.5.2 Other Safety Variables

Other safety variables are listed below:

- OST examination
- OHT examination

4.5.2.1 OST Examination

Where possible, this procedure should be conducted by a single dental examiner or clinically qualified designee. The OST examination will be accomplished throughout the study by direct observation and palpation with retraction aids as appropriate. The examiner will include examination of the labial mucosa (including lips), buccal mucosa, mucogingival folds, gingival mucosa, hard palate, soft palate, tonsillar area, pharyngeal area, tongue, sublingual area, submandibular area, and salivary glands.

The results of the examination will be recorded in the CRF as either normal or abnormal, with details of any abnormalities. Any post-treatment soft tissue abnormality, or worsening of a preexisting condition, observed by the examiner, or reported by the subject will be recorded on the CRF. Any abnormalities or worsening of a preexisting condition observed by the clinical examiner or reported by the subject from the OST examination carried out at Screening will be recorded as an AE.

OST will be summarized (number of subjects and percentages with abnormalities, without abnormalities, or OST not examined) by visit and study product in Table 14.3.4.1 for all subjects in the Safety Population. OST examination will be listed (Listing 16.2.8.1) for all randomized subjects.

4.5.2.2 OHT Examination

Where possible, this procedure should be conducted by a single dental examiner or clinically qualified designee for all subjects.

The OHT examination will assess grossly carious lesions or signs of erosive wear, enamel irregularities, tooth fracture, gross decay, decalcification, faulty restorations, and implants. The examination will be performed by direct observation.

Observations will be listed as absent or present, and conditions noted as present will be described. Examination findings will be described and documented in the CRF. Any abnormalities or worsening of a pre-existing condition observed by the clinical examiner or reported by the subject from the OHT examination carried out at Screening will be recorded as an AE.

OHT examination will be listed (Listing 16.2.8.2) for all randomized subjects.

4.6 Analysis of Other Variables

4.6.1 Repeatability of the Examiner

Repeat MGI and TPI assessments will be performed by the clinical examiner at Day 0 (Visit 2; Baseline), Week 3 (Visit 3), Week 6 (Visit 4) and Week 12 (Visit 5). At least 1 repeat assessment should be performed for each index on each clinical assessment day, that is, at least one MGI and one TPI repeat assessment on each assessment day. ‘Repeat’ subjects will be

selected at random from those in attendance. Different subjects can be used for repeat MGI and TPI assessments. In the event that the same subject is selected at random for repeatability assessments at multiple visits, the pairs of observations will be considered as independent for the purposes of repeatability analysis; that is, any potential with-subject correlation will be ignored.

There should be a delay of at least 10 minutes between original and repeat assessments for a given subject and ideally separated by another subject. No other procedure on the subject should be carried out between the first and the repeat assessment.

Scores from the first assessment must not be visible to the examiner/scribe when the repeat assessment is carried out.

The repeat MGI and TPI assessments will be compared to the respective original assessments (inclusive of any tooth surfaces assessed as missing or not qualified). The repeat assessments will not be used in any efficacy analysis.

The first and second assessments of each index will be analyzed with a Fleiss-Cohen weighted kappa coefficient (κ), along with the 95% CI, to assess the intra-examiner reliability. Reliability will be deemed:

- Excellent if $\kappa > 0.75$
- Fair to good if $0.4 \leq \kappa \leq 0.75$
- Poor if $\kappa < 0.4$

4.6.1.1 MGI Repeatability

The first and repeat values of the MGI for each tooth site will be combined and cross-tabulated for the MGI Repeatability population (Table 14.2.6.1). Only subjects with both first and repeat MGI assessments available for a given tooth site at the same visit will be presented and analyzed.

4.6.1.2 TPI Repeatability

The first and repeat values of the TPI for each tooth site will be combined and cross-tabulated for the TPI Repeatability population (Table 14.2.6.2). Only subjects with both first and repeat TPI assessments available for a given tooth site at the same visit will be presented and analyzed.

5 Changes to the Protocol Defined Statistical Analysis Plan

Changes from the originally planned statistical analysis specified in the protocol are outlined in Table 5-1 below.

Table 5-1 Changes to Protocol Defined Statistical Analysis Plan

Protocol	Statistical Analysis Plan	
Statistical Analysis Section	**Statistical Analysis Plan**	**Rationale for Changes**
No mention of a suitable non-parametric test of the primary efficacy endpoint.	Section 4.4.1.2: The assumption of normality and homogeneity of variance in the MMRM will be investigated. In case of violation of these assumptions, a Wilcoxon Signed Rank Test (within product comparisons with baseline) will be performed to support the MMRM results.	For consistency with previous studies, where Wilcoxon Signed Rank p-values were added for inclusion in the Clinical Study Report.
No mention of a suitable non-parametric test of within-group comparisons to baseline for the secondary efficacy endpoints.	Sections 4.4.2.1.1, 4.4.2.2.1, 4.4.2.3.1 and 4.4.2.4.1: The assumption of normality and homogeneity of variance in the MMRM will be investigated. In case of violation of these assumptions, a Wilcoxon Rank Sum test (between products comparisons) and Wilcoxon Signed Rank Test (within product comparisons with baseline) will be performed to support the MMRM results.	For consistency with previous studies, where Wilcoxon Signed Rank p-values were added for inclusion in the Clinical Study Report.

Appendix 1: List of Data Displays

CSR Section	TLF	Number	Title	Population	Template	Topline
14.1 Demographic Data Summary Tables and Figures						
	Table	14.1.1	Subject Disposition	All Screened Subjects	14.1.1	Yes
	Table	14.1.2	Incidence of Important Protocol Deviations	All Randomized Subjects	14.1.2	
	Table	14.1.3.1	Demographic Characteristics	Safety Population	14.1.3.1	
	Table	14.1.3.2	Demographic Characteristics	mITT Population	14.1.3.1	Yes
14.2 Efficacy Data Summary Tables and Figures						
	Table	14.2.1	Summary of Brushing Compliance	mITT Population	14.2.1	
	Table	14.2.2.1.1	Summary of Number of Bleeding Sites	mITT Population	14.2.2.1.1	Yes
	Table	14.2.2.1.2	Summary of Number of Bleeding Sites	PP Population	14.2.2.1.1	
	Table	14.2.2.1.3	Summary of Change in Number of Bleeding Sites	mITT Population	14.2.2.1.3	
	Table	14.2.2.1.4	Summary of Change in Number of Bleeding Sites	PP Population	14.2.2.1.3	
	Figure	14.2.2.1.5	Number of Bleeding Sites (± SE) Plot Over Time by Study Product	mITT Population	14.2.2.1.5	

CSR Section	TLF	Number	Title	Population	Template	Topline
	Figure	14.2.2.1.6	Number of Bleeding Sites (± SE) Plot Over Time by Study Product	PP Population	14.2.2.1.5	
	Table	14.2.2.2.1	Statistical Analysis of Change from Baseline in Number of Bleeding Sites	mITT Population	14.2.2.2.1	Yes
	Table	14.2.2.2.2	Statistical Analysis of Change from Baseline in Number of Bleeding Sites	PP Population	14.2.2.2.1	
	Table	14.2.2.2.3	Statistical Analysis of Change from Baseline in Number of Bleeding Sites (LOCF)	mITT Population	14.2.2.2.1	
	Table	14.2.2.2.4	Statistical Analysis of Change from Baseline in Number of Bleeding Sites (LOCF)	PP Population	14.2.2.2.1	
	Table	14.2.3.1.1	Summary of Bleeding Index (BI)	mITT Population	14.2.2.1.1	
	Table	14.2.3.1.2	Summary of Change in Bleeding Index (BI)	mITT Population	14.2.2.1.3	
	Figure	14.2.3.1.3	Bleeding Index (BI) (± SE) Plot Over Time by Study Product	mITT Population	14.2.2.1.5	
	Table	14.2.3.2.1	Statistical Analysis of Change from Baseline in Bleeding Index (BI)	mITT Population	14.2.2.2.1	
	Table	14.2.3.2.2	Statistical Analysis of Change from Baseline in Bleeding Index (BI) (LOCF)	mITT Population	14.2.2.2.1	
	Table	14.2.4.1.1	Summary of Modified Gingival Index (MGI)	mITT Population	14.2.2.1.1	
	Table	14.2.4.1.2	Summary of Change in Modified Gingival Index (MGI)	mITT Population	14.2.2.1.3	
	Figure	14.2.4.1.3	Modified Gingival Index (MGI) (± SE) Plot Over Time by Study Product	mITT Population	14.2.2.1.5	

CSR Section	TLF	Number	Title	Population	Template	Topline
	Table	14.2.4.2.1	Statistical Analysis of Change from Baseline in Modified Gingival Index (MGI)	mITT Population	14.2.2.2.1	
	Table	14.2.4.2.2	Statistical Analysis of Change from Baseline in Modified Gingival Index (MGI) (LOCF)	mITT Population	14.2.2.2.1	
	Table	14.2.5.1.1	Summary of Overall Turesky Plaque Index (TPI)	mITT Population	14.2.2.1.1	
	Table	14.2.5.1.2	Summary of Change in Overall Turesky Plaque Index (TPI)	mITT Population	14.2.2.1.3	
	Table	14.2.5.1.3	Summary of Interproximal Turesky Plaque Index (TPI)	mITT Population	14.2.2.1.1	
	Table	14.2.5.1.4	Summary of Change in Interproximal Turesky Plaque Index (TPI)	mITT Population	14.2.2.1.3	
	Figure	14.2.5.1.5	Overall Turesky Plaque Index (TPI) (± SE) Plot Over Time by Study Product	mITT Population	14.2.2.1.5	
	Figure	14.2.5.1.6	Interproximal Turesky Plaque Index (TPI) (± SE) Plot Over Time by Study Product	mITT Population	14.2.2.1.5	
	Table	14.2.5.2.1	Statistical Analysis of Change from Baseline in Overall Turesky Plaque Index (TPI)	mITT Population	14.2.2.2.1	
	Table	14.2.5.2.2	Statistical Analysis of Change from Baseline in Overall Turesky Plaque Index (TPI) (LOCF)	mITT Population	14.2.2.2.1	
	Table	14.2.5.2.3	Statistical Analysis of Change from Baseline in Interproximal Turesky Plaque Index (TPI)	mITT Population	14.2.2.2.1	

CSR Section	TLF	Number	Title	Population	Template	Topline
	Table	14.2.5.2.4	Statistical Analysis of Change from Baseline in Interproximal Turesky Plaque Index (TPI) (LOCF)	mITT Population	14.2.2.2.1	
	Table	14.2.6.1	Intra-Examiner Repeatability Analysis of Modified Gingival Index (MGI)	MGI Repeatability Population	14.2.6.1	
	Table	14.2.6.2	Intra-Examiner Repeatability Analysis of Turesky Plaque Index (TPI)	TPI Repeatability Population	14.2.6.2	
14.3 Safety Data Summary Tables and Figures						
14.3.1 Displays of Adverse Events						
	Table	14.3.1.1	Treatment Emergent Adverse Events by System Organ Class and Preferred Term	Safety Population	14.3.1.1	Yes
	Table	14.3.1.2	Treatment Emergent Adverse Events by Oral/Non-Oral and Preferred Term	Safety Population	14.3.1.2	
	Table	14.3.1.3	Treatment Related Treatment Emergent Adverse Events by System Organ Class and Preferred Term	Safety Population	14.3.1.1	
	Table	14.3.1.4	Treatment Related Treatment Emergent Adverse Events by Oral/Non-Oral and Preferred Term	Safety Population	14.3.1.2	
	Table	14.3.1.5	Adverse Events Related to COVID-19 by System Organ Class and Preferred Term	Safety Population	14.3.1.1	

CSR Section	TLF	Number	Title	Population	Template	Topline
	Table	14.3.1.6	Non-Serious Treatment Emergent Adverse Events by System Organ Class and Preferred Term	Safety Population	14.3.1.1	
	Table	14.3.1.7	Serious Treatment Emergent Adverse Events by System Organ Class and Preferred Term	Safety Population	14.3.1.1	
14.3.2 Listings of Deaths, Other Serious and Significant Adverse Events						
	Listing	14.3.2.1	Deaths	All Randomized Subjects	16.2.7.1	
	Listing	14.3.2.2	Non-Fatal Serious Adverse Events	All Randomized Subjects	16.2.7.1	
	Listing	14.3.2.3	Treatment Emergent Adverse Events Leading to Study or Product Discontinuation	All Randomized Subjects	16.2.7.1	
	Listing	14.3.2.4	Treatment Emergent Adverse Events Classified as Oral	All Randomized Subjects	16.2.7.1	
14.3.3 Narratives of Deaths, Other Serious and Certain Other Significant Adverse Events						
	NA					
14.3.4 Other Observations Related to Safety and Abnormal Laboratory Values						
	Table	14.3.4.1	Summary of Oral Soft Tissue Examination	Safety Population	14.3.4.1	
APPENDIX						
16.1.6 Listing of Subjects Receiving Test Drug(s)/Investigational Product(s) from Specific Batches, where more than one batch was used						
	NA					
16.1.7 Randomization Scheme and Codes (Subject identification and treatment assigned)						

CSR Section	TLF	Number	Title	Population	Template	Topline
	Listing	16.1.7.1	Randomization Information	All Randomized Subjects	16.1.7.1	
	Listing	16.1.7.2	Kit List Allocation	All Randomized Subjects	16.1.7.2	
16.1.9 Documentation of Statistical Methods						
	Raw output	16.1.9.1.1	Statistical Analysis of Change from Baseline in Number of Bleeding Sites (Reference: Table 14.2.2.2.1)	mITT Population	SAS Output	Yes
	Raw output	16.1.9.1.2	Statistical Analysis of Change from Baseline in Number of Bleeding Sites (Reference: Table 14.2.2.2.2)	PP Population	SAS Output	
	Raw output	16.1.9.1.3	Statistical Analysis of Change from Baseline in Number of Bleeding Sites (LOCF) (Reference: Table 14.2.2.2.3)	mITT Population	SAS Output	
	Raw output	16.1.9.1.4	Statistical Analysis of Change from Baseline in Number of Bleeding Sites (LOCF) (Reference: Table 14.2.2.2.4)	PP Population	SAS Output	
	Raw output	16.1.9.2.1	Statistical Analysis of Change from Baseline in Bleeding Index (BI) (Reference: Table 14.2.3.2.1)	mITT Population	SAS Output	
	Raw output	16.1.9.2.2	Statistical Analysis of Change from Baseline in Bleeding Index (BI) (LOCF) (Reference: Table 14.2.3.2.2)	mITT Population	SAS Output	

CSR Section	TLF	Number	Title	Population	Template	Topline
	Raw output	16.1.9.3.1	Statistical Analysis of Change from Baseline in Modified Gingival Index (MGI) (Reference: Table 14.2.4.2.1)	mITT Population	SAS Output	
	Raw output	16.1.9.3.2	Statistical Analysis of Change from Baseline in Modified Gingival Index (MGI) (LOCF) (Reference: Table 14.2.4.2.2)	mITT Population	SAS Output	
	Raw output	16.1.9.4.1	Statistical Analysis of Change from Baseline in Overall Turesky Plaque Index (TPI) (Reference: Table 14.2.5.2.1)	mITT Population	SAS Output	
	Raw output	16.1.9.4.2	Statistical Analysis of Change from Baseline in Overall Turesky Plaque Index (TPI) (LOCF) (Reference: Table 14.2.5.2.2)	mITT Population	SAS Output	
	Raw output	16.1.9.4.3	Statistical Analysis of Change from Baseline in Interproximal Turesky Plaque Index (TPI) (Reference: Table 14.2.5.2.3)	mITT Population	SAS Output	
	Raw output	16.1.9.4.4	Statistical Analysis of Change from Baseline in Interproximal Turesky Plaque Index (TPI) (LOCF) (Reference: Table 14.2.5.2.4)	mITT Population	SAS Output	
16.2 Subject Data Listings						
16.2.1 Discontinued Subjects						
	Listing	16.2.1.1	Subject Disposition	All Randomized Subjects	16.2.1.1	

CSR Section	TLF	Number	Title	Population	Template	Topline
	Listing	16.2.1.2	Subject Disposition	Non-Randomized Subjects	16.2.1.2	
16.2.2 Protocol Deviations						
	Listing	16.2.2.1	Important Protocol Deviations	All Randomized Subjects	16.2.2.1	
	Listing	16.2.2.2	All Protocol Deviations	All Randomized Subjects	16.2.2.2	
16.2.3 Patients Excluded from the Efficacy Analysis						
	Listing	16.2.3.1	Exclusions from Analysis Populations	All Randomized Subjects	16.2.3.1	
16.2.4 Demographic Data						
	Listing	16.2.4.1	Demographic Characteristics	All Randomized Subjects	16.2.4.1	
	Listing	16.2.4.2	Medical History and Current Medical Conditions	All Randomized Subjects	16.2.4.2	
	Listing	16.2.4.3	Prior Medications	All Randomized Subjects	16.2.4.3	
	Listing	16.2.4.4	Concomitant Medications and Significant Non-Drug Therapies	All Randomized Subjects	16.2.4.4	
16.2.5 Compliance and/or Drug Concentration Data (if available)						
	Listing	16.2.5.1	Brushing Compliance	All Randomized Subjects	16.2.5.1	

CSR Section	TLF	Number	Title	Population	Template	Topline
	Listing	16.2.5.2	Supervised Brushing	All Randomized Subjects	16.2.5.2	
16.2.6 Individual Efficacy Response Data						
	Listing	16.2.6.1	Number of Bleeding Sites	All Randomized Subjects	16.2.6.1	
	Listing	16.2.6.2.1	Bleeding Index (BI) Individual Scores	All Randomized Subjects	16.2.6.2.1	
	Listing	16.2.6.2.2	Bleeding Index (BI) Derived Scores	All Randomized Subjects	16.2.6.2.2	
	Listing	16.2.6.3.1	Modified Gingival Index (MGI) Individual Scores	All Randomized Subjects	16.2.6.3.1	
	Listing	16.2.6.3.2	Modified Gingival Index (MGI) Derived Scores	All Randomized Subjects	16.2.6.3.2	
	Listing	16.2.6.4.1	Turesky Plaque Index (TPI) Individual Scores	All Randomized Subjects	16.2.6.4.1	
	Listing	16.2.6.4.2	Turesky Plaque Index (TPI) Derived Scores	All Randomized Subjects	16.2.6.4.2	
16.2.7 Adverse Event Listings						
	Listing	16.2.7.1	All Adverse Events	All Randomized Subjects	16.2.7.1	Yes
	Listing	16.2.7.2	All Adverse Events	Non-Randomized Subjects	16.2.7.1	

CSR Section	TLF	Number	Title	Population	Template	Topline
	Listing	16.2.7.3	All Adverse Events Related to COVID-19	All Screened Subjects	16.2.7.1	
	Listing	16.2.7.4	Incidents	All Randomized Subjects	16.2.7.4	Yes
16.2.8 Other Listings and Listing of Laboratory Measurements, when required by regulatory authorities (if applicable)						
	Listing	16.2.8.1	Oral Soft Tissue Examination	All Randomized Subjects	16.2.8.1	
	Listing	16.2.8.2	Oral Hard Tissue Examination	All Randomized Subjects	16.2.8.2	
16.4 Individual Subject Data Listings						
	NA					