Statistical Analysis Plan for Protocol 212537

A Randomized, Examiner Blind, Clinical Study Investigating the Efficacy of a Stannous Fluoride Dentifrice in Improving Gingival Health after 3 Weeks Use

NCT04050722

GlaxoSmithKline Consumer Healthcare

980 Great West Road, Brentford,

Middlesex, TW8 9GS

United Kingdom (UK)

Amendment to Statistical Analysis Plan approval date: 4-Oct-2019

STATISTICAL REPORTING AND ANALYSIS PLAN

A Randomized, Examiner Blind, Clinical Study Investigating the Efficacy of a Stannous Fluoride Dentifrice in Improving Gingival Health after 3 Weeks Use

Protocol Number: 212537

Phase: 4

Document History

Document	Version Date	Summary of Changes (New analysis or Change in planned analysis)
Original Analysis Plan	28-Aug-2019	Not applicable
Amendment 1	04-Oct-2019	• Table 1-1 Schedule of Activities: Visit 4 Week 3 (21± 2 days post V3) updated to Visit 4 Week 3 (21± 2 days post V2) in accordance with Clinical Protocol Administrative Change Letter v1.0, dated 11-Sep-2019. • Figure 14.2.1 Mean Bleeding Index Score over Time, updated to topline in list of TLFs

Table of contents

Document History 2
Table of contents 3
List of tables 4
Abbreviations 5
1 Summary of Key Protocol Information 6
1.1 Study Design 6
1.2 Study Objectives 8
1.3 Treatments 9
1.4 Sample Size Calculation 10
2 Planned Analyses 10
2.1 Interim Analysis 10
2.2 Final Analyses 10
3 Considerations for data analyses and Data Handling Conventions 10
3.1 Baseline Definition 10
3.2 Subgroups/Stratifications 11
3.3 Centers Pools 11
3.4 Timepoints and Visit Windows 11
4 Data Analysis 11
4.1 Populations for Analysis 12
4.1.1 Subject Disposition 12
4.1.2 Protocol Deviations 12
4.1.3 Analysis Populations 13
4.2 Subject Demographics and Other Baseline Characteristics 14
4.2.1 Demographic Characteristics 14
4.2.2 Baseline Characteristics 14
4.2.3 General Medical History 14
4.3 Treatments (Study Product, Rescue Medication, other Concomitant Therapies, Compliance) 14
4.3.1 Study Product Compliance and Exposure 15
4.3.2 Prior and Concomitant Medication 15
4.4 Analysis of Efficacy 16
4.4.1 Primary Efficacy Endpoint 16
4.4.2 Secondary Efficacy Variables 18

4.4.3 Handling of Missing Values/Censoring/Discontinuations 19
4.5 Analysis of Secondary Objectives 19
4.5.1 Efficacy (Secondary) 19
4.6 Analysis of Safety 21
4.6.1 Adverse Events and Serious Adverse Events 21
4.6.2 Other Safety Variables 22
4.7 Analysis of Other Variables 22
5 Changes to the Protocol Defined Statistical Analysis Plan 23
Attachment 1: List of Data Displays 24

List of tables

Table 1-1 Schedule of Activities 7
Table 1-2 Investigational/Study Product Supplies 9
Table 4-1 Analysis Populations 13
Table 4-2 The Bleeding Index (BI) 16
Table 4-3 The Modified Gingival Index 18
Table 4-4 Turesky Plaque Index 19

Abbreviations

Abbreviation	Term
AE	Adverse Event
ANCOVA	Analysis of Covariance
BDRM	Blind Data Review Meeting
BI	Bleeding Index
CI	Confidence Interval
eCRF	electronic case report form
GSK CH	GlaxoSmithKline Consumer Healthcare
ICF	Informed Consent Form
κ	Kappa coefficient
MedDRA	Medical Dictionary for Regulatory Activities
MFC	Master Formulation Code
MGI	Modified Gingival Index
m-ITT	Modified Intent-To-Treat
NA	Not Applicable
NBS	number of bleeding sites
OHT	Oral Hard Tissue
OST	Oral Soft Tissue
PP	Per Protocol
ppm	parts per million
PT	Preferred Term
SAE	Serious Adverse Event
SD	Standard Deviation
SE	Standard Error
SMFP	sodium monofluorophosphate
SnF_2	stannous fluoride
SOC	System Organ Class
TEAE	Treatment Emergent Adverse Event
TPI	Turesky Plaque Index

The purpose of this Statistical Reporting and Analysis Plan is to describe the planned analyses and outputs to be included in the Clinical Study Report for Protocol 212537 version 1.0 dated 22-Jun-2019 and Protocol Administrative Change Letter version 1.0 dated 11-Sep-2019.

1 Summary of Key Protocol Information

This study is designed to evaluate the gingivitis efficacy of twice daily brushing with a 0.454% stannous fluoride (SnF_2) dentifrice over 3 weeks in a population with mild to moderate plaque-induced gingivitis.

The study will compare a marketed dentifrice (Sensodyne Repair and Protect) containing 0.454% SnF_2 with a negative control dentifrice.

This study will be considered successful if there is a statistically significant difference in Bleeding Index (BI) after 3 weeks of twice daily brushing with the SnF_2 dentifrice compared to the negative control dentifrice and the difference is in favour of the SnF_2 dentifrice.

1.1 Study Design

This will be a single center, controlled, single blind (examiner blind), randomized, stratified (gender and baseline mean whole mouth Modified Gingival Index [MGI] score), two-treatment arm, parallel design, clinical study. Study subjects will be aged 18-65 years, non-smokers, in good general health with generalized mild to moderate plaque-induced gingivitis and ≥ 20 natural teeth that meet all study criteria at both the Screening and Baseline visits, including ≥ 40 evaluable surfaces for MGI, BI, and Turesky Plaque Index (TPI).

Approximately 130 (n=65/group) subjects will be randomized to one of the study products.

The schedule of activities in Table 1-1 provides an overview of the subject visits and study procedures. The investigator may schedule visits (unplanned visits) in addition to those listed on the schedule of activities, in order to conduct evaluations or assessments required to protect the well-being of the subject.

Table 1-1 Schedule of Activities

Procedure/Assessment	Visit 1 Screening	Lead in Phase Minimum 1 day - Maximum 28 days	Visit 2 Day 0 Baseline	Visit 3 Week 2 (14± 2 days post V2)	Visit 4 Week 3 (21± 2 days post V2)
Informed consent	X				
Demographics	X				
Medical history	X				
Prior/current medications and treatments	X				
OST assessment	X		X	X	X
OHT assessment	X				X
Gross assessment of gingival health[1]	X				
Inclusion/Exclusion criteria	X		X		
Subject Eligibility	X		X		
Subject continuance				X	X
Concomitant medications and treatments			X	X	X
MGI assessment			X	X	X
Repeat MGI Assessment[2]			X	X	X
BI assessment			X	X	X
Disclose dental plaque			X	X	X
TPI Assessment			X	X	X
Repeat TPI Assessment[2]			X	X	X
Stratification/Randomization			X		
Dental Prophylaxis			X		
Post-prophylaxis examination to confirm whole mouth TPI=0			X		
Dispense study product, toothbrush and diary			X		
Supervised brushing with study product			X	X	
Subject brings study product, toothbrush and completed diary to site for compliance check[3]				X	X
Subject returns study product and diary to site					X
End of study dental prophylaxis (optional)					X
Adverse events/incidents[4]	X		X	X	X
Study conclusion					X

Abbreviations: OST = Oral soft tissue, OHT = Oral hard tissue, MGI = Modified gingival index, BI = Bleeding index, TPI = Turesky plaque index

Footnotes:

1. Subjects with generalized mild-moderate gingivitis, as determined by visual assessment of gingival health, will continue in the study.

2. At least 2 repeatability assessments should be performed each day (≥ 1 in the morning; ≥ 1in the afternoon)
3. Visits 2-4: Adherence to Lifestyle Guidelines/Medication Requirements;
   Visits 3-4: Compliance with use of study product (visual check if returned study supplies/review of diary)
4. Adverse events (AEs)/incidents will be recorded from the signing of informed consent until 5 days after the last dose of study product.

1.2 Study Objectives

The study objectives are as follows:

Objective(s)	Endpoint(s)
Primary	
To evaluate the clinical efficacy of a 0.454% SnF_2 dentifrice in reducing gingivitis (following dental prophylaxis), as measured by the Bleeding Index (BI), compared to a negative control dentifrice, when used twice daily for 3 weeks.	Mean BI at Week 3
Secondary	
To evaluate the clinical efficacy of a 0.454% SnF_2 dentifrice in reducing gingivitis (following dental prophylaxis), as measured by the BI compared to a negative control dentifrice, when used twice daily for 2 weeks.	Mean BI at Week 2
To evaluate the clinical efficacy of a 0.454% SnF_2 dentifrice in reducing gingivitis (following dental prophylaxis), as measured by the number of bleeding sites (NBS), compared to a negative control dentifrice, when used twice daily for 2 and 3 weeks.	Mean NBS at Week 2 and Week 3
To evaluate the clinical efficacy of a 0.454% SnF_2 dentifrice in reducing gingival inflammation (following dental prophylaxis), as measured by the Modified Gingival Index (MGI), compared to a negative control dentifrice, when used twice daily for 2 and 3 weeks.	Mean MGI at Week 2 and Week 3
To evaluate the clinical efficacy of a 0.454% SnF_2 dentifrice in reducing supra-gingival plaque formation (following dental prophylaxis) as measured by the Turesky modification of the Quigley & Hein plaque index (TPI), compared to a negative control dentifrice, when used twice daily for 2 and 3 weeks.	Mean TPI (overall and interproximal) at Week 2 and Week 3
Safety	
To assess the safety and tolerability of a 0.454% SnF_2 dentifrice when used twice daily over 3 weeks.	Treatment emergent adverse events

1.3 Treatments

There are two study product groups in this study:

- Test dentifrice: Sensodyne Repair and Protect containing 0.454% SnF_2 (1100 parts per million (ppm) fluoride)
- Negative control dentifrice: Colgate Cavity Protection (1100ppm fluoride as sodium monofluorophosphate [SMFP]).

The following study products will be supplied by the Clinical Supplies Department, GlaxoSmithKline Consumer Healthcare (GSK CH):

Table 1-2 Investigational/Study Product Supplies

	Investigational Products	
	Test Product	**Negative Control Dentifrice**
Product Name	Sensodyne Repair and Protect	Colgate Cavity Protection
Pack Design	Carton of 2 over-wrapped tubes	Carton of 2 over-wrapped tubes
Dispensing Details	1 carton of 2 tubes- Baseline visit	1 carton of 2 tubes- Baseline visit
Product Master Formulation Code (MFC)	Commercial Product (USA Market place product) (CCI [redacted])	Commercial Product (USA Market place product)
Dose/Application	Full ribbon of toothpaste on head of toothbrush provided	Full ribbon of toothpaste on head of toothbrush provided
Route of Administration	Oral topical	Oral topical
Usage Instructions	Subjects will brush their teeth for one timed minute twice a day (morning and evening) in their usual manner	Subjects will brush their teeth for one timed minute twice a day (morning and evening) in their usual manner
Return Requirements	All used/unused samples to be returned	All used/unused samples to be returned

1.4 Sample Size Calculation

A sufficient number of healthy subjects will be screened to randomize approximately 130 subjects (65 per study product group) to ensure 60 evaluable subjects per treatment complete the entire study.

The primary objective for this study is to evaluate the clinical efficacy of a 0.454% SnF_2 dentifrice in reducing gingivitis (following dental prophylaxis), as measured by the BI, compared to a negative control dentifrice, when used twice daily for 3 weeks.

With 60 evaluable subjects per study product group, it should be possible to detect a 0.1 difference in BI (sd=0.17) [Parkinson et al, 2018b] with 90% power and a two sided 5% significance level. This difference would approximately represent a 23% difference between study product groups.

This estimate will allow detection of a 5% difference in a key secondary variable MGI with 90% power.

2 Planned Analyses

2.1 Interim Analysis

No interim analysis is planned for this study.

2.2 Final Analyses

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

- All subjects have completed the study as defined in the protocol.
- All required database cleaning activities have been completed and database has been locked.
- 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 latest non-missing assessment prior to first study product use (Visit 2, Day 0).

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

3.2 Subgroups/Stratifications

Subjects will be stratified by gender (Male/Female) and baseline mean whole mouth MGI score (Low: ≤2.0/High: >2.0).

Based on above stratification factors, there will 4 strata as follows:

- Stratum 1: Male, Low MGI Score (Baseline MGI ≤ 2.00)
- Stratum 2: Male, High MGI Score (Baseline MGI > 2.00)
- Stratum 3: Female, Low MGI Score (Baseline MGI ≤ 2.00)
- Stratum 4: Female, High MGI Score (Baseline MGI > 2.00)

All eligible subjects will be stratified based on above 4 strata, to ensure a balance of gender and gingivitis across both study product groups and then randomized to study product.

The primary and secondary analysis will include gender, and baseline MGI stratification as factors, with the exception of the analysis of MGI, for which the stratification factor of MGI will not be included as the baseline value of MGI will be included as a covariate.

No subgroup analyses are planned for this study.

3.3 Centers Pools

Since this is single center study, pooling of centres 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 case-by-case basis at the Blinded 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 Syneos Health. 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.

Unless otherwise described, 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 will also be displayed. Percentages for screen failure subjects will be based on the total number of subjects screened.

The number and percentage of subjects who complete and discontinue the study, broken down by reason for discontinuation, will be presented by study product group and overall in Table 14.1.1. 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 by study product group and overall. Percentages will be based on the number of subjects randomized in the relevant study product group or overall.

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

Subject disposition information will be listed for non-randomized subjects (Listing 16.2.1.2), including demographic information (age, sex and race), screening date, reason for screen failure and any further details of reason for screen failure.

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 devitions not leading to exclusion from the PP population.

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

- Violation of inclusion or exclusion criteria
- Non-compliance with study product use
- Use of prohibited treatment or medication before or during the study
- Violation of visit windows

The specific details of the important protocol deviations and how these will be assesed will be specified in the Blind Data Review Plan and subjects with important protocol devitions will be identified at the BDRM.

The number and percentage of subjects with at least one important protocol deviation, important protocol deviations not leading to exclusion from 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 group and overall (Table 14.1.2) and listed in Listing 16.2.2.1.

All protocol deviations collected on the protocol deviation case report form 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

The analysis populations defined for this study are as follows:

Table 4-1 Analysis Populations

Population	Definition/Criteria	Analyses Evaluated
Safety	• All randomized subjects who receive at least one dose of the study product. This population will be based on the study product the subject actually received.	Safety
Modified Intent-To-Treat (m-ITT)	• All randomized subjects who received at least one dose of the study product and provided at least one post-baseline assessment of efficacy (BI, MGI or TPI). All m-ITT population summaries and analyses will be presented according to the study product randomized.	Efficacy
Per Protocol (PP)	• All subjects included in the m-ITT population who have at least one assessment of efficacy (BI, MGI or TPI) considered unaffected by protocol violations. Subjects with a protocol violation that is deemed to affect efficacy for only some (but not all) of the efficacy assessments will be part of the PP population. Protocol deviations that would exclude subjects from the PP population are defined in Section 4.1.2 (Protocol Deviations).	Efficacy analyses for BI score
Repeatability	• Comprise of all subjects who have a repeat clinical assessment of efficacy (R) at any visit. There will be a separate population for repeat MGI assessment and repeat TPI assessment: • MGI Repeatability population: Subjects with at least one initial and repeat assessment of MGI at any visit.	Repeatability analyses

Population	Definition/Criteria	Analyses Evaluated
	• TPI Repeatability population: Subjects with at least one initial and repeat assessment of TPI at any visit.	

The primary population for assessment of efficacy will be the m-ITT Population. A PP analysis will be performed only on the primary variable (BI score) if more than 10% of subjects in the m-ITT population are excluded from the PP Population.

Any repeat clinical data collected for the repeatability assessment will only be used to assess repeatability. The main assessment of efficacy will be based on the initial assessment.

Subjects excluded from any of the analysis populations will be listed in Listing 16.2.3.1.

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, frequency count (n) and percentage (%) of subjects for categorical variables] will be presented for demographic characteristics by study product group and overall. These variables include gender, race, ethnicity and age (years), and will be presented for the Safety population (Table 14.1.3.1), the m-ITT population (Table 14.1.3.2) and if applicable, for the PP population (Table 14.1.3.3).

Demographic information will be listed for all randomized subjects in Listing 16.2.4.1.1.

4.2.2 Baseline Characteristics

Baseline characteristics presented for the Safety population (Table 14.1.4.1) include baseline mean MGI score, category (≤2.00/>2.00) and the the number and percentage of subjects included in each stratum as described in Section 3.2 (Subgroups/Stratifications). Baseline characteristics will also be presented for the m-ITT population (Table 14.1.4.2) and the PP Population (Table 14.1.4.3).

Baseline characteristics will be listed for all randomized subjects in Listing 16.2.4.1.2.

4.2.3 General Medical History

Medical history data will be listed (Listing 16.2.4.2) with start date and end date or ongoing at the start of the study.

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

Randomization details will be listed, including the planned study product, the actual study product the subject was randomized to and the randomization date (Listing 16.1.7).

4.3.1 Study Product Compliance and Exposure

Product compliance will be recorded as number of missed brushings since last visit and number of additional brushings since last visit.

Exposure to study product will be calculated for each visit interval and overall study product duration as follows:

- Exposure (Days) = Date of $Visit_n$ – Date of $Visit_{n-1}$.

Compliance will be calculated for each visit interval and overall as:

- Compliance (%) = (Actual number of brushings / Expected number of brushings) x 100.

Where:

- Expected number of brushings = (Date of $Visit_n$ – Date of $Visit_{n-1}$) x 2.
- Actual number of brushings = Expected number – missed brushings + additional brushings.

Overall exposure and compliance will be calculated from Visit 2 (Day 0) to Visit 4 (Week 3) [or the last available visit for subjects who discontinue prior to Visit 4 (Week 3)].

The number of missed brushings, additional brushings and percentage compliance will be summarized by study product group using descriptive statistics (n, mean, SD, median, minimum and maximum) for each visit interval and overall study product duration, in Table 14.2.1.1.1 for the Safety population and Table 14.2.1.1.2 for the m-ITT population. In addition, the number and percentage of subjects <80%, between 80% - 120% and >120% compliant will also be presented.

Product exposure (days) will be summarized by study product group using descriptive statistics (n, mean, SD, median, minimum and maximum), for each visit interval and overall study product duration, in Table 14.2.1.2.1 for the Safety population and Table 14.2.1.2.2 for the m-ITT population.

Compliance and exposure data will be listed for all randomized subjects (Listing 16.2.5.1)

4.3.2 Prior and Concomitant Medication

Prior or concomitant medication taken by or administered to a subject will be recorded in the case report form. The prior and concomitant medications will be coded using an internal validated medication dictionary, GSK Drug.

Prior medication will be listed by subject, with drug name, GSK drug synonym, dose, frequency, route, start date, study day relative to first use of study product and end date (Listing 16.2.4.3). Prior medications are defined as those which stopped before the date of first use of the study product. If the stop date is unknown or incomplete and the medication cannot be considered as stopped prior to the first use of study product then the medication will be considered as a concomitant medication.

Concomitant medications and significant non-drug therapy will be listed similarly (Listing 16.2.4.4) with either ongoing or end date displayed. Concomitant medications are defined as medications that started or stopped on or after the date of 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 analysis of primary and secondary efficacy variables will be based on the initial assessment.

4.4.1 Primary Efficacy Endpoint

4.4.1.1 Primary Efficacy Endpoint Definition

The primary efficacy endpoint is the mean whole mouth BI score at Week 3.

The BI will be assessed for the facial and lingual/palatal gingival surfaces of all evaluable teeth, six sites per tooth (mesiobuccal, buccal, distobuccal, mesiolingual/palatal, lingual/palatal and distolingual/palatal). Gingival bleeding is assessed 30 seconds after probing. Assessments will be performed one quadrant at a time; BI scores will be recorded before moving to the next quadrant. The BI will be assessed by the same examiner on all evaluable teeth at Baseline (Visit 2), Week 2 (Visit 3) and Week 3 (Visit 4). The BI scoring system will be as follows:

Table 4-2 The Bleeding Index (BI)

Score	Description
0	Absence of bleeding on probing
1	Bleeding observed within 30 seconds of probing
2	Bleeding observed immediately on probing

The mean whole mouth BI score for each subject will be derived from the total BI score divided by the number of tooth sites scored.

4.4.1.2 Statistical Hypothesis, Model, and Method of Analysis

Descriptive statistics (n, mean, SD, SE, median, minimum, maximum) of the primary variable will be provided at Baseline, Week 2 and Week 3 visits by study product group for the m-ITT population (Table 14.2.2.1).

The primary analysis is a comparison of the mean whole mouth BI score between the test dentifrice and negative control dentifrice at Week 3 for subjects eligible for the m-ITT population.

The null hypothesis for the primary endpoint is that there is no difference in the mean BI score between the 2 study product groups.

H_0: $\mu_1 = \mu_2$

The alternative hypothesis is that there is a difference in the mean BI score between the 2 study product groups.

H_1: $\mu_1 \neq \mu_2$

The mean BI score will be analyzed using analysis of covariance (ANCOVA) with study product group, gender, and baseline MGI stratification as factors and baseline mean BI score as a covariate. Adjusted means and their SEs, for each study product group will be displayed with the study product group difference, SE, 95% confidence interval (CI) of the difference, between-product p-value and percent difference between study product groups in Table 14.2.2.2 for m-ITT population.

Percent difference will be calculated as:

- Percent Difference = (Adjusted Mean Difference/Adjusted Mean of Negative Control Dentifrice)*100.

All statistical tests of hypotheses will be two-sided and will employ a level of significance of $\alpha = 0.05$.

This study will be considered successful if a statistically significant difference between the adjusted mean BI scores of the two study product groups at Week 3 is observed to be in favor of the test dentifrice (SnF_2 dentifrice).

The assumption of normality and homogeneity of variance in the ANCOVA model will be investigated. Violation of this assumption will be overcome using a suitable data transformation or a non-parametric technique (e.g. Van Elteren test).

The raw mean and SE of the BI score for each study product group will be presented graphically over time for the m-ITT population in Figure 14.2.1.

The BI obtained for each tooth, surface and site will be listed by subject and visit in Listing 16.2.6.1 for all randomized subjects. The mean whole mouth BI score will be listed by subject and visit in Listing 16.2.6.4 for all randomized subjects.

4.4.1.3 Supportive Analyses

If there is more than 10% difference in the overall number of subjects between PP and m-ITT populations, a summary of the primary efficacy variable will be presented for all subjects in the PP population (Table 14.2.2.3) and the same ANCOVA model applied to the primary analysis will be performed on the PP population (Table 14.2.2.4).

4.4.2 Secondary Efficacy Variables

4.4.2.1 Mean BI at Week 2

The mean BI score at Week 2 will be calculated in the same way as stated in Section 4.4.1.1.

4.4.2.2 Number of Bleeding Sites at Week 2 and Week 3

The number of bleeding sites (NBS) will be derived from BI , where a bleeding site is a site scored as 1 or 2. The BI scoring system is described in Section 4.4.1.1.

4.4.2.3 Modified Gingival Index at Week 2 and Week 3

The MGI will be assessed for the facial and lingual/palatal gingiva of all evaluable teeth, four sites per tooth (facial surface - papilla and margin; lingual/palatal surface - papilla and margin).

The MGI scoring system will be as decribed in Table 4-3 and will be assessed by the same examiner on all evaluable teeth at Baseline (Visit 2), Week 2 (Visit 3) and Week 3 (Visit 4).

Table 4-3 The Modified Gingival Index

Score	Description
0	Absence of inflammation
1	Mild inflammation: slight change in colour, little change in texture of any portion of the marginal or papillary gingival unit
2	Mild inflammation: criteria as [1] 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

The mean whole mouth MGI score for each subject will be derived from the total MGI score divided by the number of tooth sites scored.

4.4.2.4 Turesky Plaque Index (Overall and Interproximal) at Week 2 and Week 3

Supra-gingival plaque will be assessed on the facial and lingual surfaces of the teeth using the TPI. Each tooth surface will be divided into three areas; three scores will be recorded facially (mesiofacial, facial, distofacial) and three scores lingually (mesiolingual, lingual and distolingual) generating a total of six scores per tooth. The plaque will be disclosed and scored for each site as follows.

Table 4-4 Turesky Plaque Index

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

The mean Overall TPI score for each subject will be derived from the total TPI score over all tooth sites divided by the number of tooth sites scored.

The mean Interproximal TPI score for each subject will be derived from the total TPI score over all interproximal tooth sites divided by the number of interproximal tooth sites scored.

4.4.3 Handling of Missing Values/Censoring/Discontinuations

Missing data will not be replaced or imputed. Subjects who withdraw from the study early will be included in the statistical analysis up to the point of withdrawal. Subjects who withdraw will not be replaced.

4.5 Analysis of Secondary Objectives

Each secondary variable will be analyzed separately as per the primary variable (Section 4.4.1.2). All analyses will be conducted on the m-ITT population only. For the analysis of MGI the stratification factor of MGI will not be included as the baseline value of MGI is included as a covariate.

4.5.1 Efficacy (Secondary)

4.5.1.1 Mean BI at Week 2

The mean BI score at Week 2 will be analyzed in the same way as stated in Section 4.4.1.2.

4.5.1.2 Number of Bleeding Sites at Week 2 and Week 3

Descriptive statistics (n, mean, SD, SE, median, minimum, maximum) of the NBS will be provided by visit and randomized study product group for the m-ITT population (Table 14.2.3.1).

The NBS will be analyzed using ANCOVA with study product group, gender, and baseline MGI stratification as factors and baseline NBS as a covariate. Adjusted means and their SEs for each study product group will be displayed along with the study product group difference, SE, 95% CI of the difference, between-product p-value and percent difference between study product groups in Table 14.2.3.2 for m-ITT population.

The assumption of normality and homogeneity of variance in the ANCOVA model will be investigated. Violation of this assumption will be overcome using a suitable data transformation or a non-parametric technique (e.g. Van Elteren test).

The raw mean and SE of the NBS for each study product group will be presented graphically over time for the m-ITT population in Figure 14.2.2.

Bleeding sites will be listed in Listing 16.2.6.1 for all randomized subjects. The NBS will be listed by subject and visit in Listing 16.2.6.4 for all randomized subjects.

4.5.1.3 Mean MGI at Week 2 and Week 3

Descriptive statistics (n, mean, SD, SE, median, minimum, maximum) of the mean MGI score will be provided by visit and randomized study product group for the m-ITT population (Table 14.2.4.1).

The mean MGI score will be analyzed using ANCOVA with study product group and gender as factors and baseline mean MGI score as a covariate. Adjusted means and their SEs for each study product group will be displayed with the study product group difference, SE, 95% CI of the difference, between-product p-value and percent difference between study product groups in Table 14.2.4.2 for m-ITT population.

The assumption of normality and homogeneity of variance in the ANCOVA model will be investigated. Violation of this assumption will be overcome using a suitable data transformation or a non-parametric technique (e.g. Van Elteren test).

The raw mean and SE of the MGI score will be presented graphically over time by study product group for the m-ITT population in Figure 14.2.3.

The MGI obtained for each tooth, surface and site will be listed by subject and visit in Listing 16.2.6.2 for all randomized subjects. The mean whole mouth MGI score will be listed by subject and visit in Listing 16.2.6.4 for all randomized subjects.

4.5.1.4 Mean TPI (Overall) at Week 2 and Week 3

Descriptive statistics (n, mean, SD, SE, median, minimum, maximum) of the mean Overall TPI score will be provided by visit and randomized study product group for the m-ITT population (Table 14.2.5.1).

The mean Overall TPI will be analyzed using ANCOVA with study product group, gender, and baseline MGI stratification as factors and baseline mean Overall TPI score as a covariate. Adjusted means and their SEs for each study product group will be displayed with the study product group difference, SE, 95% CI of the difference, between-product p-value and percent difference between study product groups in Table 14.2.5.2 for m-ITT population.

The assumption of normality and homogeneity of variance in the ANCOVA model will be investigated. Violation of this assumption will be overcome using a suitable data transformation or a non-parametric technique (e.g. Van Elteren test).

The raw mean and SE of the Overall TPI score for each study product group will be presented graphically over time for the m-ITT population in Figure 14.2.4.

The TPI obtained for each tooth, surface and site will be listed by subject and visit in Listing 16.2.6.3 for all randomized subjects. The mean Overall TPI score will be listed by subject and visit in Listing 16.2.6.4 for all randomized subjects.

4.5.1.5 Mean TPI (Interproximal) at Week 2 and Week 3

Descriptive statistics (n, mean, SD, SE, median, minimum, maximum) of the mean Interproximal TPI score will be provided by visit and randomized study product group for the m-ITT population (Table 14.2.6.1).

The mean Interproximal TPI will be analyzed using ANCOVA with study product group, gender, and baseline MGI stratification as factors and baseline mean Interproximal TPI score as a covariate. Adjusted means and their SEs for each study product group will be displayed with the study product group difference, SE, 95% CI of the difference, between-product p-value and percent difference between study product groups in Table 14.2.6.2 for m-ITT population.

The assumption of normality and homogeneity of variance in the ANCOVA model will be investigated. Violation of this assumption will be overcome using a suitable data transformation or a non-parametric technique (e.g. Van Elteren test).

The mean Interproximal TPI score will be listed by subject and visit in Listing 16.2.6.4 for all randomized subjects.

4.6 Analysis of Safety

All safety data will be reported for the Safety population as per actual study product received. The safety profile of the study products will be assessed with respect to Exposure, AEs, incidents and OST.

Exposure is described in Section 4.3.1 (Study Product Compliance and Exposure).

4.6.1 Adverse Events and Serious Adverse Events

All AEs will be reviewed by the Clinical Research Scientist or Designee prior to database lock 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 the electronic case report form (eCRF).

Treatment emergent adverse events (TEAEs) are defined as AEs with an onset date/time on or after the date/time of first study product use. Adverse events 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 treatment-emergent AEs by system organ class (SOC) and Preferred Term (PT) (Table 14.3.1.1.1). Summary of the number and percentage of subjects with at least one AE, total number of AEs, number and percentage of AEs within each SOC and PT will be displayed.
- Table of treatment-emergent AEs by Oral/Non-Oral and PT (Table 14.3.1.1.2)
- Table of related treatment-emergent AEs by Oral/Non-Oral and PT (Table 14.3.1.1.3)
- Listing of all AEs (Listing 16.2.7.1 for all randomized subjects; Listing 16.2.7.2 for non-randomized subjects)
- Listing of incidents (Listing 16.2.7.3)
- Listing of deaths (Listing 14.3.2.1)
- Listing of non-fatal SAEs (Listing 14.3.2.2)
- Listing of treatment-emergent AEs leading to study or drug discontinuation (Listing 14.3.2.3)
- Listing of treatment-emergent AEs classified as oral (Listing 14.3.2.4)

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

4.6.2 Other Safety Variables

A shift table of OST by examination will be provided, for each of the study products (Table 14.3.4.1) comparing baseline results to post-baseline results at Week 2 and Week 3, for the Safety population.

Oral soft tissue (OST) and oral hard tissue examination (OHT) data will be listed (Listing 16.2.9.1 and Listing 16.2.9.2 respectively) for all randomized subjects.

4.7 Analysis of Other Variables

Repeat MGI and TPI assessments will be performed by the clinical examiner at Baseline (Visit 2), Week 2 (Visit 3) and Week 3 (Visit 4). At least 2 repeat assessments should be performed for each index on each clinical assessment day (≥1 in the morning; ≥1 in the afternoon). 'Repeat' subjects will be selected at random from those in attendance. Different subjects can be used for repeat MGI and TPI assessments.

The repeat dental assessments (MGI and TPI) will be compared to the original assessments and will be used to investigate intra-examiner variability. The repeat assessments will not be used in any efficacy analyses..

The first and repeat assessments for each tooth site will be cross tabulated for MGI (Table 14.2.7.1) and for TPI (Table 14.2.7.2).

A weighted Kappa coefficient (κ), along with the 95% CI will be calculated to assess the intra-examiner reliability. Fleiss-Cohen weighted kappa will be calculated for the repeatability analysis. Reliability will be deemed:

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

This analysis will be conducted for MGI using the MGI Repeatability population and for TPI using the TPI Repeatability population.

5 Changes to the Protocol Defined Statistical Analysis Plan

There were no changes or deviations to the originally planned statistical analysis specified in the protocol, dated 22-Jun-2019 and protocol administrative change letter, dated 11-Sep-2019.

Attachment 1: List of Data Displays

212537 (CCI
Gingivitis) List of TFLs