STATISTICAL ANALYSIS PLAN
17 May 2018

A Phase 1/2 Clinical Trial to Assess the Safety and Preliminary Efficacy of Lipoxin Analog BLXA4-ME Oral Rinse for the Treatment of Gingivitis

NIDCR Protocol Number: 14-020-E

Principal Investigator:

Hatice Hasturk, DDS, PhD
The Forsyth Institute
245 First Street
Cambridge, MA 02142
Telephone: 617-892-8499
Fax: 617-892-8505

IND Sponsor:

Thomas E. Van Dyke, DDS, PhD
The Forsyth Institute
245 First Street
Cambridge, MA 02142
Telephone: 617-892-8503
Fax: 617-892-8504

Funding Agency:

National Institute of Dental and Craniofacial Research (NIDCR)
Program Official: Jane Atkinson, DDS
Grant Number: U01DE026934-01

PREPARED BY

Robert James, MS, MStat, Senior Biostatistician III
Rho, Inc.
6330 Quadrangle Drive, Suite 500
Chapel Hill, NC 27517
Telephone: (919) 408-8000
Fax: (919) 408-0999

DOCUMENT VERSION CONTROL

Version Number	Date	Comments/Changes
0.1	22 August 2017	First Draft
0.2	27 October 2017	Clarified the SAP as to the handling of missing values at DSMB request.
1.0	02 February 2018	Changed "worst observation carried forward" to last observation carried forward" Minor changes to the General Analysis and Reporting Conventions, and the Demographic and Other Baseline Characteristics.
1.1	02 May 2018	Updated sections 8.1 and 9.1.1 to describe analyses of subjects that receive the wrong drug treatment.
2.0	03 May 2018	General formatting; Second approved version
3.0	17 May 2018	Modified analysis populations in sections 4.1 and 4.2 to describe how to handle subjects that receive the wrong drug treatment. Removed the sentences describing handling of subjects that received the wrong treatment from sections 8.1 and

	9.1.1.Section 9.3.2 added clarification that all efficacy analyses were performed using the efficacy population. Added section 11 Changes to the Analyses Planned in the Protocol

APPROVALS

Approved: _____ Date: _____

Name: Thomas Van Dyke, DDS, PhD

Title: IND Sponsor

_____ Date: _____

Name: Hatice Hasturk, DDS, PhD

Title: Principal Investigator

_____ Date: _____

Name: Robert James, MS, MStat

Title: Senior Biostatistician III

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AE	adverse event
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
BLXA4-ME	Investigational drug, a stable lipoxin analog
BOP	Bleeding on Probing
CAL	Clinical Attachment Level
CI	Confidence Interval
CS	Clinically Significant
DBP	Diastolic Blood Pressure
GCF	Gingival Crevicular Fluid
GSI	Gingivitis Severity Index
IL-1β	Interleukin-1β
LXA_4	Lipoxin A_4
MedDRA	Medical Dictionary for Regulatory Activities
MGI	Modified gingival index
OMAS	Oral Mucositis Assessment Scale
PD	Pocket depth
PI	Plaque Index
PSI	Plaque Severity Index
PT	Preferred Term
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SBP	Systolic Blood Pressure
SD	Standard Deviation
SE	Standard Error
SOC	System Organ Class
TEAE	Treatment-Emergent Adverse Event

1. PURPOSE OF THIS DOCUMENT

This statistical analysis plan (SAP) for protocol NIDCR Protocol 14-020-E is being developed before database lock and the statistical analyses of the data. The objective of this document is to (1) briefly summarize the purpose of the study and study design, (2) detail the planned statistical methodology, (3) define original and derived variables necessary to complete the analyses, and (4) delineate the most appropriate summary tables, listings, and graphical displays to be included in the study report.

This SAP will be presented to the Principal Investigator for approval. The statistical analyses and their corresponding tables, listings, and figures, as outlined in this document will be presented to the PI at the completion of the study analyses.

2. PROTOCOL SUMMARY

2.1 Study Objectives

The primary objective is to evaluate the safety of an investigational compound, BLXA4-ME, topically applied for 28 days as a daily oral rinse in adults with gingival inflammation. Safety will be assessed by the incidence of adverse events, including mucosal inflammation and irritancy, and findings from oral examinations and safety laboratory tests. Additionally, subjects will be monitored for development or progression of periodontitis and oral flora will be analyzed to detect an increase in opportunistic organisms.

The secondary objective is to assess preliminary efficacy of the oral rinse, by monitoring changes in the plaque index (PI), modified gingival index (MGI), bleeding on probing (BOP), and levels of interleukin-1β (IL-1β) in gingival crevicular fluid (GCF).

2.2 Overall Study Design and Plan

This study has a randomized parallel placebo-controlled double-blind study design with one treatment arm and two control arms (placebo oral rinse, and no rinse controls). Subjects in the treatment (N=50) and placebo rinse (N=50) arms will be double-blinded and receive oral rinse (BLXA4-ME or placebo rinse) to be applied once daily after morning teeth brushing. The no-rinse control arm (N=25) will be blinded from the investigators and apply no oral rinse. Subjects will perform their randomized treatment for 28 days. Safety parameters will be assessed before treatment and at Days 3, 7, 14, 21, and 28 after the initial treatment. Efficacy parameters will be assessed before treatment and Day 14 and Day 28 after the initial treatment. Total duration of subject participation will be 3 months (in clinic visits through 28 days, and a follow-up safety phone call at 2 months after last product use).

The Study Flow Chart and the Schedule of Events can be found in Appendices 13.1 and 13.2, respectively.

2.3 Study Population

Approximately 125 healthy adult male and female subjects with gingival inflammation, determined by mean full mouth MGI of at least 2.0 at screening and baseline, will participate in the study.

2.4 Blinding

This is a double-blind study. The PI, study coordinators, study examiners, data managers, and statisticians will be masked to the assigned treatment. Subjects receiving oral rinse will not know whether they are in the active or placebo rinse group. The no-rinse control will be blinded only from the PI, study coordinators, and study examiners since it is not possible to blind the subject from not having an oral rinse. The product code may be broken for an individual subject if he or she experiences a serious adverse event (SAE) and cannot be adequately treated without knowing the identity of the study product. The study blind will be broken after database lock.

2.5 Sample Size Determination

Sample size determination is based on the power to detect AEs (a primary safety outcome) with additional consideration given to detecting post-baseline study arm differences with respect to the MGI (a secondary efficacy outcome).

The probability of observing a specific AE for the study's sample size can be estimated. This probability depends on the sample size within a treatment group and on the true underlying probability of an individual experiencing the specific AE in that treatment group. Table 2.5.1 summarizes the probabilities of observing an AE given a sample size and true underlying event rate. For example, if the true probability of a subject experiencing a specific AE is 5%, the probability of observing 1 or more of that specific AE within a treatment group of 40 subjects is 87.1%.

The study's test oral rinse, placebo oral rinse, and no-rinse control arms are assigned treatment group sample sizes of 40, 40, and 20 subjects, respectively. For the test oral rinse and the placebo oral rinse arms, a sample size of 40 subjects per treatment group should have probabilities of 80.5%, 87.1%, 91.6%, and 96.4% of observing an AE having underlying incidence probabilities of 4%, 5%, 6%, and 8%, respectively. The no-rinse control arm is considered to be of lesser importance and, thus, to conserve the overall sample size, is assigned a sample size of 20 subjects, which should have probabilities of 55.8%, 64.2%, 71.0%, and 81.1% of observing an AE having underlying incidence probabilities of 4%, 5%, 6%, and 8%, respectively. Thus, the above sample sizes of 40, 40, and 20 subjects should give adequate power (> 80%) to detect AEs that occur at true underlying frequencies of 4% or greater in the test and placebo oral rinse study arms, and adequate power (> 80%) to detect AEs that occur at true underlying frequencies of 8% or greater in the no-rinse control arm. Conversely, if no AEs occur with a sample size of 20 and 40 subjects then the upper one-sided 95% CI of the underlying AE incidence rate would be 14% and 7.5%, respectively.

Table 2.5.1 Probability of Observing One or More Adverse Events within a Study Arm

"True" Underlying Probability of an Individual Experiencing an Adverse Event	Probability of Observing One or More Adverse Events within a Study Arm	
	No-Rinse Control Study Arm (n = 20)	Test and Placebo Oral Rinse Study Arms (n = 40)
0.01%	0.20%	0.40%
0.10%	2.0%	3.9%
0.25%	4.9%	9.5%
0.50%	9.5%	18.2%
1.00%	18.2%	33.1%
2.00%	33.2%	55.4%
3.00%	45.6%	70.4%
4.00%	55.8%	80.5%
5.00%	64.2%	87.1%
6.00%	71.0%	91.6%
8.00%	81.1%	96.4%
10.0%	87.8%	98.5%

Sample size (n) refers to the final sample size per treatment group of enrolled subjects who complete the study.

Power estimates for the secondary efficacy endpoint (MGI) are found in Table 2.5.2. In this power analysis, MGI refers to the MGI scores averaged over all tooth sites within each subject and visit. The power estimates in the table are based upon 50,000 simulations of mixed models analyses. The mixed models simulations include the MGI as the dependent variable, and fixed effects for treatment group (test oral rinse, placebo oral rinse, and no-rinse control), visit (Baseline, Day 14, and Day 28), and their treatment group-by-visit interaction. The model includes random effects for subject using a compound symmetry variance structure (i.e., exchangeable covariance). These power analyses assume an overall within-study group standard deviation of 0.3 MGI units at each visit and a difference in the mean change from baseline MGI at 28 days between the test oral rinse and the placebo rinse or between the test oral rinse and the no-rinse control of 0.2 MGI units. Correlation of MGI within subject among visits was assumed to be 0.3. The treatment group sample sizes of 40, 40, and 20 subjects for the test oral rinse, placebo oral rinse, and no-rinse control arms, respectively, are estimated to have powers of 71% and 54% to detect significant differences of 0.2 in the mean MGI scores between the test oral rinse and placebo oral rinse study arms, and between the test oral rinse and the no-rinse control study arms, respectively. Even though these sample sizes would be considered under-powered for testing efficacy in a Phase 3 clinical trial, for this Phase 1/2 protocol, with efficacy as a secondary objective, these study sample sizes can provide evidence (significant or not) of efficacy and provide efficacy estimates with their associated standard deviations needed for planning future Phase 2 or Phase 3 trials.

It is estimated that 50, 50, and 25 subjects will need to be enrolled to have final sample sizes of 40, 40, and 20 subjects for the test oral rinse, placebo oral rinse, and no-rinse control study arms, respectively, assuming a 20% subject attrition rate before completing the study.

Table 2.5.2. Mixed Models Power Analysis for MGI (Secondary Efficacy Outcome) Showing the Power for Several Sample Size Scenarios

Enrolled and Eligible Subjects			Subjects after 20% Drop-out			Power (%)	
Test (n)	Placebo (n)	Control (n)	Test (n)	Placebo (n)	Control (n)	Trt – Placebo[a]	Trt – Control[b]
50	50	25	40	40	20	70.9	53.8

Control = no-rinse control study arm; MGI = modified gingival index; placebo = placebo oral rinse study arm; test = test oral rinse study arm; trt = treatment (placebo or test oral rinse) study arm.

a Trt-Placebo: Power (probability) for finding a statistically significant difference (alpha = 0.05) in the change from baseline MGI at 28 days between the test oral rinse and the placebo oral rinse study arms.

b Trt-Control: Power (probability) for finding a statistically significant difference (alpha = 0.05) in the change from baseline MGI at 28 days between the test oral rinse and the no-rinse control study arms.

3. GENERAL ANALYSIS AND REPORTING CONVENTIONS

The following is a list of general analysis and reporting conventions to be applied for this study. If departures from these general conventions are present in the specific evaluations section of this SAP, then those conventions will take precedence over these general conventions.

- All analysis will be performed using the SAS System version 9.4 or later.
- All subject data listings and summary listings will be sorted for presentation in order of treatment group and subject number in the following order: BLXA4-ME rinse, placebo rinse, and non-rinse control.
- Missing values for both numeric and character variables will be presented as blanks in a table or listing.
- Mixed case will be used for all titles, footnotes, column headers, and programmer-supplied formats.
- All date values will be presented as DDMMMYYYY (e.g. 21AUG2008) format. A 4-digit year is required for all dates. Partial dates will dashes to represent missing days or months (e.g. - -JUL2009).
- All observed time values will be presented using a 24-hour clock HH:MM format (e.g., 01:35, or 17:26).
- Population sizes will be presented for each treatment group in the column heading as (N=XX).
- The sample size (n) shown as a summary statistic will be the number of subjects with non-missing values.
- Categorical variables will be summarized using counts (n) and percentages (%) and will be presented in the form n (%). All summaries for categorical variables will include all categories, even if none of the subjects had a response in a particular category, to ensure completeness. Continuous variables will be summarized using descriptive statistics: n, mean, median, standard deviation (SD) or standard error (SE), minimum, and maximum. The mean and median will be reported at 1 more significant digit than the precision of the data. The SD and/or SE will be reported at 2 more significant digits than the precision of the data. The minimum and maximum will be reported to the same level of precision as the original observations.
- Calculated variables will not be rounded prior to analysis. Rounding will only occur after analysis. If any values are calculated to have more significant digits than the original data, then the value will be rounded so that it is the same level of precision as the original data.
- P-values will be reported to 3 decimal places if greater than 0.001. If less than 0.001 then p-values will be reported as ‘<0.001’.

4. ANALYSIS POPULATIONS

The exclusion criterion of periodontitis was removed in October 2015 for Protocol Version 4.0. This will possibly change the study population. To explore the possible effects of including study participants with periodontitis, sensitivity analyses will be performed as described in section 9.3.2.2.

4.1 Safety Population

The Safety Population will consist of all subjects who are randomized into a study arm and complete their Baseline Visit. For the safety evaluations, subjects will be analyzed according to the treatment actually received. Any subjects that in error were administered both the BLXA4-ME and the placebo rinse at different times during the study will be analyzed in the BLXA4-ME treatment arm. Subjects that in error were administered both the non-rinse control treatment and either the BLXA4-ME or placebo rinse will be analyzed according to the rinse administered.

4.2 Efficacy Population

The Efficacy Population will consist of all subjects of the Safety Population who have a baseline and at least 1 post-baseline assessment of 1 or more of the secondary efficacy outcome measures. Subjects that dropout or withdraw prior to Day 14 will be replaced in the efficacy population. For the efficacy evaluations, subjects will be analyzed according to the treatment actually received. Any subjects that began the study administering the treatment for one of the study arms and then later in error switched treatments to another study arm will not be included in the efficacy population.

5. STUDY SUBJECTS

5.1 Disposition of Subjects

The disposition of all subjects and the study populations to which they belong will be presented in the subject listings and summarized in a table as counts by treatment allocation. The subject dispositions will include subjects randomized, completing the study, discontinuing/ withdrawing, and the number of subjects in both the safety and efficacy analysis populations. The reasons for discontinuing/ withdrawing from the study will be included in the disposition listing and table.

5.2 Protocol Deviations

Protocol deviations will be presented in a listing and summarized in a table by treatment allocation and type of deviation. Types of deviations that will be summarized as counts and percentages in the table include:

- Consent deviation
- Eligibility deviation
- Missed visit
- Procedure performed outside of visit window
- Study procedure not completed
- Study procedure deviation
- Other

Table 5.2.1 Assessment Time Windows

Study Visit	Time window (study day ± assessment window)
Visit 1 – Enrollment, Baseline	Day 0
Visit 2	Day 3 ± 1
Visit 3	Day 7 ± 1
Visit 4	Day 14 ± 1
Visit 5	Day 21 ± 1
Visit 6 – Final Study Visit	Day 28 ± 1
Visit 7 – Telephone Safety Follow-up	Day 90 ± 14

6. DEMOGRAPHIC AND OTHER BASELINE CHARACTERISTICS

Baseline data (e.g., demography, medical history, dental history, and concurrent medications) will be summarized.

Demographic baseline characteristics will be summarized descriptively for both the Safety and Efficacy Populations. The demographic variables of sex, age, ethnicity, and race will be listed and summarized in a table. The categorical variables of sex, ethnicity, and race will be summarized by frequencies. Age will be summarized as mean, SD, median, minimum, and maximum.

To assess overall health of the study participants prior to study treatment, any significant medical history will be included in the subject listings. The number and percentage of subjects with each medical history event will be presented by body system or oral structure.

Results of the baseline oral exam will be listed by area/oral structure examined and will denote an abnormal or normal result; and, if abnormal, the associated description of the abnormality will be listed.

Tooth status at baseline will be presented in a listing by subject and tooth, and will denote whether the tooth is missing, sound, decayed, filled, fractured, was crowned, and/or is a dental implant.

All prior and concomitant medications will be coded to preferred drug names and therapeutic drug class using the WHO Drug Dictionary. Use of concomitant medications will be summarized for each therapeutic drug class and each preferred drug name. Concomitant medications will include medications reported on the Concomitant Medications page of the CRF.

7. MEASUREMENTS OF TREATMENT COMPLIANCE

Compliance will be recorded at all study visits by measurement of the volume of remaining product. Expected volume returned is calculated as the volume dispensed - (number of days since last study visit x 30ml). A subject will have "over used" the rinse if the percent volume used is > 120%. A subject will have "under used" the rinse if the percent volume used is < 80%. A subject listing will capture under- and over-usage of study drug and non-compliance as determined by the volume of returned study drug and as reported by the subject queries. If more than 12 subjects are out of study drug compliance, a table will summarize the frequencies of under- and over-usage of study drug by treatment arm.

8. SAFETY EVALUATION

8.1 Overview of Safety Analysis Methods

All safety analyses will be carried out using the Safety Population.

- Adverse events (AEs)
- Unanticipated problems
- Vital signs
- Laboratory determinations (hematology, serum chemistry, and urinalysis)
- Physical examinations
- Dental examinations

With the exception of the progression of periodontitis based on the dental examinations, no statistical comparisons between the treatment groups are planned for the safety variables.

8.2 Progression of Periodontitis

Progression of periodontitis will be defined as an increase of 2 mm or more in pocket depth (PD) or clinical attachment level (CAL) from the Day 0 baseline measurement. The frequency (percent) of subjects with progression of periodontitis will be summarized by study arm, visit, and overall. Generalized linear mixed models regression based on a binomial distribution and logit link will compare the odds ratios of the overall incidence of the progression of periodontitis among the study arms at the Day 14 and Day 28 time points. The results of the generalized linear mixed model will be presented in a table and include p-values and odds ratios (with 95% CI) for the treatment arm comparisons at the two study visits. The p-values will not be adjusted for multiplicity.

The generalized linear mixed models analyses will automatically adjust for any missing values provided the missing values are missing at random; and thus, no additional imputation is necessary for missing values that are missing at random.

Missing observations are “not missing at random” if their missing-ness is possibly correlated with disease progression or efficacy. Missing PD or CAL observations due to missed visits or early study withdrawal will be deemed “not missing at random” if there were any treatment-emergent AE (TEAE) prior to or concurrent with a subject’s missed visit and the relationship between the TEAE and study drug was rated as “Possible”, ”Probable”, or ”Definite”.

If more than 5% of the PD or CAL measures are missing, or any missing PD or CAL measures are deemed “not missing are random”, then a sensitivity analysis will be performed by repeating the above generalized linear mixed model with all missing values imputed as having progression of periodontitis. The sensitivity analysis will also be triggered if for any other reason the principal investigator suspects that missing values may bias the analysis. The reasons a sensitivity analysis is triggered will be footnoted in the sensitivity analysis table.

8.3 Extent of Exposure

Summary statistics will be presented for the treatment exposure. Treatment exposure will be defined as the number of days that the subject performed their allocated oral rinse treatment and the total volume of rinse used.

8.4 Adverse Events

All AEs will be coded to a system organ class (SOC) and preferred term (PT) using the most current version of the Medical Dictionary for Regulatory Activities (MedDRA). All tabular AE summaries will be for TEAEs. TEAEs will be defined as those AEs with an onset on or after the date of first dose of study drug (study visit 1: Day 0). If an AE was recorded prior to the first dose of study drug and there was an increase in its severity, the AE will be considered a TEAE. All other AEs will be classified as non-treatment-emergent. Treatment-emergent AEs will be flagged in all AE data listings.

An overall summary table will be developed to report the number of events and the incidence of participants having at least one event in the following categories:

- TEAEs
- TEAEs indicated as serious (SAEs)
- TEAEs that lead to study drug discontinuation
- TEAEs with an outcome of death
- TEAEs that were reported as having a definite/certain or possible relation to study drug
- TEAEs reported as having a severity rating of severe

In addition, a summary table of TEAEs classified by system organ class (SOC) and preferred term (PT) will be provided for each of the following:

- TEAEs
- TEAEs by maximum severity
- TEAEs by relationship to study drug

The summary of all TEAEs will present both the number of TEAEs and the incidence of TEAEs. The summary of TEAEs by severity, relationship to study drug, and week of treatment will only report the incidence of TEAEs. When reporting the number of TEAEs, if the same TEAE occurs for a subject on multiple occasions the event will be counted once for each occurrence. When reporting the incidence, a subject will only be counted once if they ever experience an event within the SOC and ever experience the individual PT.

For the summary of TEAEs by severity, if the severity of the TEAE is not reported, then the severity of the TEAE will be counted as severe. If the same TEAE occurs for a subject on multiple occasions, the TEAE will be categorized according to the highest severity rating for that TEAE in that subject. For the summary of TEAEs by relationship to study drug, if the relationship is missing, it will be counted as definite/certain. If the same TEAE occurs for a subject on multiple occasions, the TEAE will be categorized according to the closest relationship to study drug reported for that TEAE in that subject.

8.5 Deaths, Serious Adverse Events, and Other Significant Adverse Events

In addition to these summary tables, the incidence of adverse events leading to death, the incidence of SAEs, and incidence of TEAEs that led to study drug being discontinued will be presented in summary tables and listings. The summary listings will provide all of the information reported for that AE and will include the number of days from date of first dose to date of the occurrence of the AE.

8.6 Clinical Laboratory Evaluation

Subject clinical chemistry, hematology, and urinalysis samples will be obtained at the baseline, Day 14 and Day 28 visits. All lab results will be listed by subject, laboratory parameter, and time. Out-of-range flags (H: high; or L: low) when available will be used to denote abnormal values.

All treatment-emergent abnormal values, as determined on the basis of the normal values provided by the central laboratory, will be listed with relevant subject information. In addition, all treatment-emergent values that were deemed clinically significant (CS) by the on-site medical monitor will be flagged in the listing. An additional listing will only list records whose values were deemed clinically significant.

The quantitative hematology, clinical chemistry, and urinalysis parameters and their change from baseline will be summarized in tables by study arm and study visit. Summary statistics will include n, mean, SD, median, minimum, maximum values, and the number and percent of subjects with missing values for the laboratory test. Repeat or ad hoc laboratory tests will not be summarized, but they will be included in the listings. The order of presentation will cluster the laboratories by test and time of measurement in three tables for clinical chemistry, hematology, and urinalysis data, respectively.

For hematology, clinical chemistry, and urinalysis laboratory parameters, shift tables will be presented for change from baseline at study Day 14 and Day 28. For the shift tables, the Baseline laboratory parameter values will be cross-referenced with their associated study visit Day 14 and Day 28 laboratory parameter values with respect to the n and percent of parameter values flagged as low, normal, high, or abnormal based on the laboratory provided normal ranges for each of the lab parameters. The shift tables will include an extra column for baseline values that shift to missing values due to early withdrawal or otherwise.

8.7 Microbial Profiles in Pooled Supragingival Plaque

The pooled supragingival plaque samples collected from the mesiobuccal surfaces of the Ramfjord teeth at the baseline, Day 14 and Day 28 visits will be analyzed for the presence of pathogenic organisms. The frequency and percent of the presence of the following organisms will be summarized (tables, listings and checkerboard plots) by study arm and visit: *Fusobacterium nucleatum subspecies vincentii, Campylobacter concisus, Campylobacter rectus, Tannerella forsythensis, Prevotella intermedia, Prevotella nigrescens, Porphyromonas gingivalis, Capnocytophaga sputigena, Streptococcus oralis, Actinomyces naeslundii, Actinomyces israeli, Eubacterium brachy, Eikenella corrodens,* and spirochetes. The table and figure will also capture the frequency and percent of subjects with microbial counts missing.

8.8 Vital Signs, Physical Findings, and Other Observations Related to Safety

8.8.1 Vital Signs

The vital sign parameters will include sitting pulse rate, respiration rate, blood pressure (systolic (SBP) and diastolic (DBP)), and body temperature (oral). All vital signs and weight values will be listed, by subject. Out-of-range flags (H or L) will be used to denote abnormal values. The abnormal values will be identified based upon the clinically significant criteria CS as listed in Table 9.8.1.1. In addition, all treatment-emergent CS vital sign values or weights will be listed separately, by subject, with relevant subject information.

The actual vital sign values and the change from baseline values will be summarized at all assessment times. Summary statistics will include n, mean, SD, median, minimum, maximum values, and the number and percent of subjects with missing values for the vital sign measure.

Table 8.8.1.1 CS Criteria for Vital Signs

Parameter	PCS Criteria
Vital Signs	
Blood pressure (sitting)	
Diastolic	< 50 mmHg, %, with decrease from BL ≥15 mmHg; or > 105 mmHg, with increase from BL ≥15 mmHg
Systolic	< 90 mmHg, %, with decrease from BL ≥20 mmHg; or > 180 mmHg, with increase from BL ≥20 mmHg
Heart rate (sitting)	< 50 BPM, with decrease from BL ≥10 BPM; or > 120 BPM; or Change from BL ≥ 30 BPM
Temperature (oral)	> 101°F or 38.4°C, with increase from BL ≥ 2°F or 0.75°C
Weight	Change from BL ≥ 7% body weight

BL=Baseline

8.8.2 Oral Exam

All oral exam findings will be listed by subject. Any treatment-emergent abnormal findings will be listed. Subjects will be considered to have a treatment-emergent abnormal finding if the subject has a normal assessment for a specific oral structure at the baseline visit and is assessed as having an abnormal finding at a post-baseline study visit. A table will summarize by the number and percent of subjects that have treatment-emergent abnormal findings for each oral structure. This table will also summarize for each oral structure the number and percent of subjects that improve or have no change in normal/abnormal status from the Baseline visit. The number and percent of subjects with missing values due to early withdrawal or otherwise will be summarized in the table for each oral structure.

8.8.3 Oral Mucositis Assessment Scale Ulceration and Erythema Scores

The Oral Mucositis Assessment Scale (OMAS) score for ulceration and erythema will be listed for each oral structure by subject and visit. A table will summarize by the number and percent of subjects for each oral structure with ulceration and erythema scores that worsen, improve, or stay the same from baseline. The table will also summarize the number and percent of subjects with the OMAS scores missing due to early withdrawal or otherwise.

8.8.4 Tooth Status

The tooth status will be listed for each tooth by subject and visit. A table will summarize by the number and percent of subjects and by the number and percent of teeth for each tooth attribute at each study visit that worsen, improve, stay the same from baseline, or have missing values due to early study withdrawal or otherwise . Tooth status attributes recorded include: missing, sound, decayed, filled, fractured, crown, and dental implant.

9. EFFICACY EVALUATION

9.1 Overview of Efficacy Analysis Issues

9.1.1 Handling of Dropouts and Missing Data

.

Subjects that dropout or are withdrawn prior to Day 14 will be replaced in the Efficacy Population. The mixed models analyses (described in the sections below) will automatically adjust for any missing values provided the missing values are missing at random; and thus, no additional imputation is necessary for missing values that are missing at random.

Missing observations are "not missing at random" if their missing-ness is possibly correlated with disease progression or efficacy. Missing efficacy observations due to missed visits or early study withdrawal will be deemed "not missing at random" if there were any treatment-emergent AE (TEAE) prior to or concurrent with a subject's missed visit and the relationship between the TEAE and study drug was rated as "Possible", "Probable", or "Definite".

If more than 5% of a given efficacy endpoint is missing, or any missing observations are deemed "missing not at random" then a sensitivity analysis will be performed by repeating the given efficacy analysis with all missing values imputed using "last observation carried forward". The sensitivity analysis will also be triggered if for any other reason the principle investigator suspects that missing values may bias the analysis. The reasons any sensitivity analyses are triggered will be footnoted in the sensitivity analysis tables.

9.2 Efficacy Endpoints

9.2.1 Primary Efficacy Endpoints

This is a Phase 1/2 study and there are no primary efficacy endpoints as listed in the protocol; however, the secondary endpoints of modified gingival index (MGI) followed by bleeding on probing (BOP), both at Day 28, are considered the efficacy endpoints of chief importance.

9.2.2 Secondary Efficacy Endpoints

At baseline, study Day 14, and study Day 28, the following secondary efficacy variables: modified gingival index (MGI), bleeding on probing (BOP), IL-1β, plaque index (PI), plaque severity index (PSI) and gingivitis severity index (GSI) will be recorded. PI, MGI, and BOP will be scored for each of the 6 tooth-sites for each tooth. IL-1β will be measured in the GCF collected from one tooth (assigned by algorithm) in each quadrant. All efficacy analyses will be performed on the Efficacy Population.

The time point of greatest interest would be Day 28, but analysis will also be performed for Day 14.

The exploratory analysis pocket depth (PD) and clinical attachment level (CAL) will be descriptive.

9.2.2.1 Plaque Index (PI)

Plaque Index will be assessed at 6 sites per tooth (mesiobuccal, buccal, distobuccal, distolingual, lingual, and mesiolingual surfaces of the tooth). Each tooth site will have its PI scored as follows:

0 = No plaque;
1 = Film at gingival margin;
2 = Moderate (easily visible);
3 = Abundance of material.

The PI will be summarized as the whole-mouth average for each participant and study visit as follows: First the PI tooth-site measures will be averaged for each tooth; and then, these tooth averages will be averaged over the entire mouth for a given participant.

9.2.2.2 Modified Gingival Index (MGI)

Modified Gingival Index (MGI) will be assessed at 6 sites per tooth (mesiobuccal, buccal, distobuccal, distolingual, lingual, and mesiolingual surfaces of the tooth). Each tooth-site will have its MGI scored as follows:

0 = Absence of Inflammation
1 = Mild inflammation; slight change in color, little change in texture of any portion of, but not the entire marginal or papillary gingival unit
2 = Mild inflammation; criteria as above but involving the entire marginal and 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 MGI will be summarized as the whole-mouth average for each participant and study visit as follows: First the MGI tooth-site measures will be averaged for each tooth; and then, these tooth averages will be averaged over the entire mouth for a given participant.

9.2.2.3 Bleeding on Probing (BOP)

Bleeding on probing is a dichotomous measurement that will be recorded at the same six sites per tooth as PI and MGI. The BOP will be reduced to one record per subject as percent of sites with BOP.

9.2.2.4 Plaque and Gingivitis Severity Indices

Plaque Severity Index (PSI) : This index allows for a comparison of the tooth surface sites from each group that received the most severe Silness and Loe Plaque Index scores; that is, a Silness and Loe Plaque Index score of 2 and 3. The mean Plaque Severity Index will be calculated for each subject by dividing the total number of tooth surface sites scored either 2 or 3 by the total number of tooth surface sites scored in the mouth for plaque formation (number of teeth scored multiplied by six).

Gingivitis Severity Index (GSI): This index allows for a comparison of the gingival sites from each group that received the most severe Modified Gingival Index scores; that is, a Modified Gingival Index score of 3 or 4, by the total number of sites scored in the entire mouth for gingivitis (number of teeth scored multiplied by six).

9.2.2.5 Interleukin-1β (IL-1β)

The IL-1β will be summarized as the arithmetic and geometric means of the IL-1β concentrations over the 4 GCF samples taken at each visit for a study participant. These means will be referred to as the 4-sample IL-1β means in the remainder of this document.

9.2.3 Exploratory Efficacy Endpoints

9.2.3.1 Pocket Depth

The pocket depth (PD) will be measured at each of the 6 tooth-sites for each tooth of each participant during the baseline and study Day 28 visit.

9.2.3.2 Clinical Attachment Loss

The clinical attachment level (CAL) will be calculated at each of the 6 tooth-sites for each tooth of each participant during the baseline and study Day 28 visit. CAL will be calculated as PD minus the cementoenamel junction.

9.3 Analysis Methods

9.3.1 Primary Efficacy Analyses

All efficacy analyses are considered secondary analyses.

9.3.2 Secondary Efficacy Analyses

The main objectives of the secondary efficacy analyses are to: 1) obtain some evidence of efficacy (either a statistically significant effect or a trend); 2) obtain estimates of the size and variance of possible treatment effects that can then be used to plan later clinical trials; and 3) explore which of several outcomes may best be used to measure efficacy. All efficacy analyses will be based on the efficacy population.

9.3.2.1 Secondary Efficacy Endpoints

Secondary analyses will test for treatment arm differences in the change from baseline of the efficacy outcomes: MGI, percent of tooth sites with BOP, PI, GSI, PSI, and IL-1β, at the study Day 14 and 28. These secondary endpoints will be analyzed using mixed models with change from baseline of the whole mouth average of the efficacy endpoint as the dependent variable, subject as a random effect, a covariate adjustment for baseline, and fixed effects for visit, study arm, and a study arm-by-visit interaction.

The following SAS code using mixed models will be used to analyze the secondary endpoints: MGI, PI, GSI, PSI , IL-1β (4-sample mean), and percent of sites with BOP.

```
Proc Mixed data=dataset;
   class subject visit study_arm;
   model  change  = baseline study_arm visit study_arm *visit ;
   repeated visitn / sub=subject  type=cs;
   lsmeans   study_arm*visit / pdiff;
run;
```

where: change = change in the efficacy endpoints from baseline.

These mixed models analyses assume the efficacy outcomes are approximately normal in distribution. A compound symmetry variance structure is assumed for the mixed models; however, if the Schwarz Bayesian criterion indicates that an unstructured variance structure better fits the model then the unstructured variance structure for the repeated effects will be used. Residual analysis will test the underlying assumptions of the mixed models analyses. The choice of using the arithmetic 4-sample IL-1β means or the geometric 4-sample IL-1β means will be based upon the models with the better residual fit to model assumptions. If for any efficacy endpoint, the normality assumption is found to be unreasonable, then log-transformation will be used to normalize the data. Furthermore, if log-transformation fails to normalize the data then generalized linear mixed models with the appropriate links and distributional assumptions or will be used to analyze these efficacy outcomes.

Within the each of the mixed models analysis of the efficacy endpoints, pairwise contrasts will be made to test for statistical differences among the 3 study arms with respect to the change from baseline at the Day 14 and Day 28 visits. Further pairwise contrasts will test each endpoint for statistical differences between the Day 14 and Day 28 visits within each study arm.

Since this is a safety phase 1 / 2 with safety measures as the primary endpoints, the secondary efficacy endpoints will not be adjusted for multiplicity. However, the contrast between the treatment and placebo rinse arms with respect to the MGI change from baseline at Day 28 will be considered the secondary endpoint of chief importance. The same statistical contrast but with BOP change from baseline as the dependent variable will be considered the statistical contrast next in order of importance.

Tables will summarize the p-values, contrasts, and least squares means (with standard errors) of the change from baseline and their differences among the treatment groups and visits for each of the mixed models analyses described above.

Summary statistics of the secondary endpoints and their change from baseline, including n, means, minimums, maximums, standard deviations, and standard errors, will be summarized in a table by study arm and study visit. Raw values of all the secondary efficacy endpoints, both on the tooth-site level and the subject level will be listed.

9.3.2.2 Sensitivity Analyses of Secondary Efficacy Outcomes

The exclusion criterion for periodontitis was removed in October 2015 for Protocol Version 4.0. This may change the study population by adding study participants with more advanced disease and possibly increase the variability in the effect sizes of the secondary efficacy outcomes. To explore the possible effects of changing the study population, sensitivity analyses will be performed for each of the secondary efficacy analyses. These sensitivity analyses will be performed similar to the secondary efficacy analyses but with the exclusion of subjects with periodontitis that would have been excluded in the original protocol. The results of the sensitivity analyses will then be compared to the secondary efficacy analyses to understand how removing the exclusion for periodontitis may affect the efficacy results.

9.3.3 Exploratory Efficacy Analyses

Exploratory analyses will test for treatment arm differences in the change from baseline of the exploratory efficacy outcomes: PD and CAL at study Day 28. The PD and CAL change from baseline at Day 28 will be analyzed using ANCOVA with the PD or CAL change from baseline as the dependent variable, a covariate adjustment for baseline, and fixed effects study arm.

The following SAS code using mixed models will be used to analyze the secondary endpoints: PGSS, whole-mouth averages of PI, and MGI, IL-1β 4-sample mean, and percent of sites with BOP.

```
Proc GLM data=dataset;
   class subject study_arm;
   model  change  = baseline study_arm  ;
   lsmeans   study_arm / pdiff;
run;
```

where: change = change in the efficacy endpoints at study Day 28 from baseline.

The ANCOVA analyses assume the outcomes are approximately normal in distribution with homogeneity of variance. Residual analysis will test these underlying assumptions. If these underlying assumptions are found to be unreasonable for the ANCOVA analyses then the nonparametric Wilcoxon rank sum will be used to test for pairwise differences between the 3 treatment arms with respect to the PD or CAL change from baseline at study Day 28.

Tables will summarize the p-values, and least squares means (with standard errors) of the change from baseline and their differences among the treatment groups and visits for each of the mixed models analyses. If the nonparametric Wilcoxon rank sum test is employed then p-values with their associated Hodges-Lehmann estimators will be reported along with the median, minimum, maximum, 10^{th}, and 90^{th} percentiles for the changes from baseline of the PD and CAL measures for each treatment arm.
Raw values of all the PD and CAL measures, both on the tooth-site level will be listed.

A sensitivity analysis will be performed to compare generalized estimating equations (GEE) methods to the mixed models results in the secondary efficacy analyses. These generalized

estimating equations (GEE) method with a linear link and exchangeable covariance matrix (using a within subject correlation structure) will be used to estimate the treatment group effects on the change from baseline in MGI with a covariate adjustment for the baseline MGI. These exploratory GEE models will be performed at both the whole-mouth and tooth-site levels.

10. INTERIM ANALYSES AND DATA MONITORING

No formal interim analysis is planned.

11. CHANGES TO THE ANALYSES PLANNED IN THE PROTOCOL

The safety and efficacy populations were expanded on to describe the handling of subjects that received the wrong treatment.

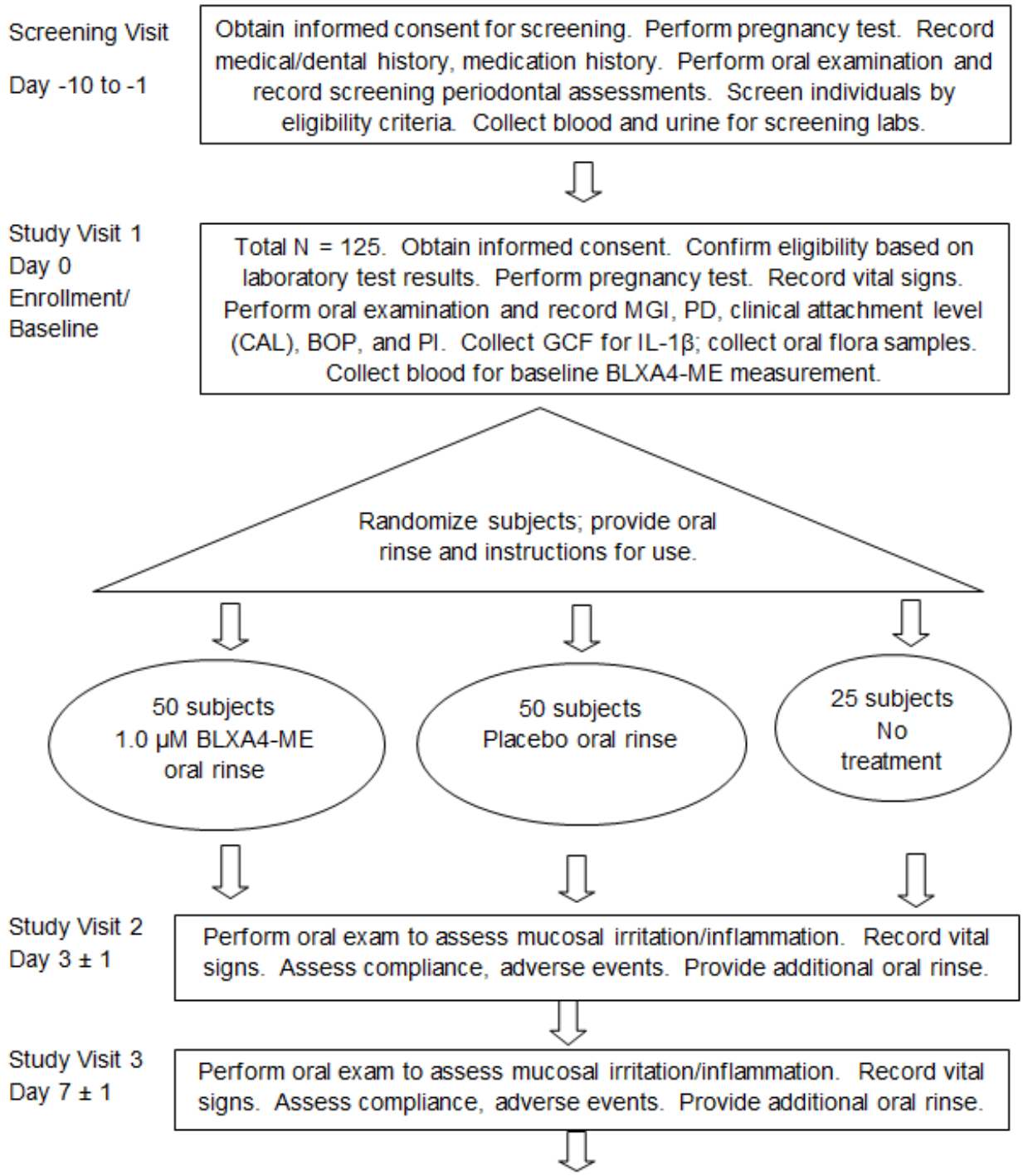
12. LIST OF PLANNED TABLES, FIGURES, AND LISTINGS

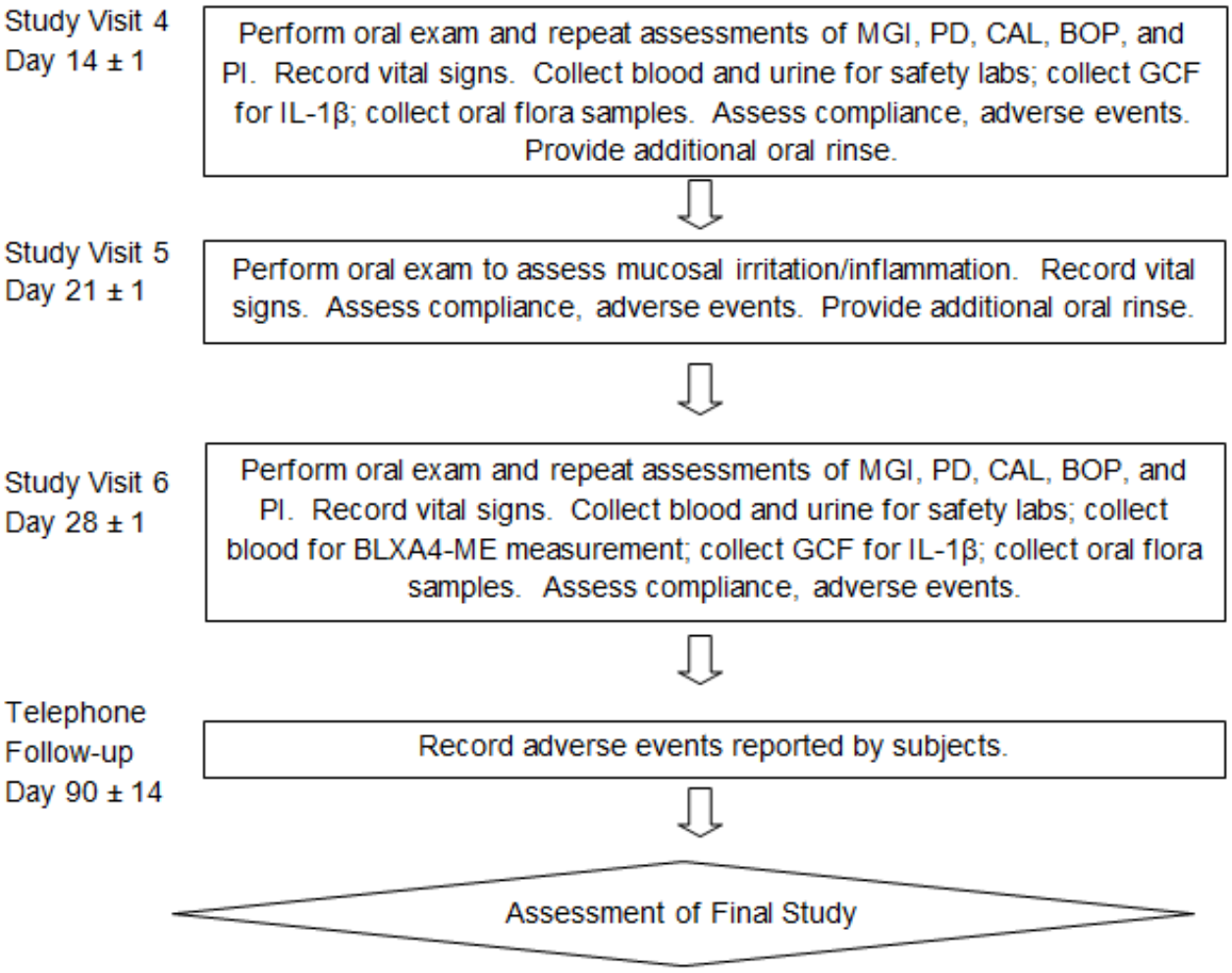
Lists of planned tables, figures, and listings will be provided in a separate document.

13. APPENDICES

13.1 Study Flow Chart

Schematic of Study Design




Study Visit 4
Day 14 ± 1
Perform oral exam and repeat assessments of MGI, PD, CAL, BOP, and PI. Record vital signs. Collect blood and urine for safety labs; collect GCF for IL-1β; collect oral flora samples. Assess compliance, adverse events. Provide additional oral rinse.
Study Visit 5
Day 21 ± 1
Perform oral exam to assess mucosal irritation/inflammation. Record vital signs. Assess compliance, adverse events. Provide additional oral rinse.
Study Visit 6
Day 28 ± 1
Perform oral exam and repeat assessments of MGI, PD, CAL, BOP, and PI. Record vital signs. Collect blood and urine for safety labs; collect blood for BLXA4-ME measurement; collect GCF for IL-1β; collect oral flora samples. Assess compliance, adverse events.
Telephone Follow-up
Day 90 ± 14
Record adverse events reported by subjects.
Assessment of Final Study

13.2 Schedule of Events

Procedures	Screen Day -10 to -1	Enrollment/ Baseline Visit 1 Day 0	Visit 2 Day 3 ± 1 day	Visit 3 Day 7 ± 1 day	Visit 4 Day 14 ± 1 day	Visit 5 Day 21 ± 1 day	Visit 6 Day 28 ± 1 day	F/U Day 90 ± 14 days
Obtain informed consent	X	X						
Medical/dental history	X[a,b]	X[a]	X	X	X	X	X	
Record concomitant medications and dental procedures	X	X	X	X	X	X	X	
Urine pregnancy test	X	X						
Vital signs[c]	X	X	X	X	X	X	X	
Intraoral and extraoral examinations	X	X	X	X	X	X	X	
Assessment of plaque and calculus	X							
Solicit adverse events			X	X	X	X	X	X
Modified gingival index	X	X			X		X	
Gingival crevicular fluid sample for IL-1β		X			X		X	
Probing depth; measure GM-CEJ to determine clinical attachment level		X			X		X	
Bleeding on probing		X			X		X	
Plaque index		X			X		X	
Plaque samples to analyze oral flora		X			X		X	
Blood for plasma BLXA4-ME level and SPM profiling in serum and plasma		X					X	
Blood for eligibility or safety laboratory tests[d]	X				X		X	

Procedures	Screen Day -10 to -1	Enrollment/ Baseline Visit 1 Day 0	Visit 2 Day 3 ± 1 day	Visit 3 Day 7 ± 1 day	Visit 4 Day 14 ± 1 day	Visit 5 Day 21 ± 1 day	Visit 6 Day 28 ± 1 day	F/U Day 90 ± 14 days
Urine for eligibility or safety laboratory tests[d]	X				X		X	
Assess compliance			X	X	X	X	X	
Provide oral rinse		X	X	X	X	X		
Hygiene instructions or referral for periodontal treatment								X

F/U = safety follow-up telephone call; GM-CEJ = distance from free gingival margin to cementoenamel junction; IL-1β = interleukin-1β; SPM = specialized pro-resolution mediator

a Includes history of alcohol and tobacco use.

b Includes demographic information.

c Pulse rate, respiratory rate, blood pressure, and oral body temperature are collected at all visits marked with "X". Weight is collected at the Screening Visit and Visit 6 (Day 28).

d The on-site medical monitor will review the results of safety laboratory tests to determine whether a laboratory test should be redrawn.