Protocol GDC-229-002

A Multicenter, Randomized, Double-Blind, Vehicle Controlled Study Evaluating the Therapeutic Equivalence and Safety of GDC-229 (Investigational Metronidazole 0.75% Vaginal Gel) and Metronidazole 0.75% Vaginal Gel in the Treatment of Bacterial Vaginosis

Statistical Analysis Plan

Prepared for:

Gage Development Company, LLC
9450 W. Bryn Mawr Avenue
Suite 640
Rosemont, IL 60018

STATISTICAL ANALYSIS PLAN APPROVAL

[redacted]	DATE (DDMMMYYYY)
	16Feb2018
	16 FEB2018
	16 Feb 2018
	16 Feb 2018

TABLE OF CONTENTS

W W

AE	Adverse Event
BMI	Body Mass Index
BV	Bacterial Vaginosis
CI	Confidence Interval
CSR	Clinical Study Report
DA	Drug Accountability
eCRF	Electronic Case Report Form
HEENT	Head, Eyes, Ears, Nose and Throat
ITT	Intent-To-Treat
KOH	Potassium hydroxide
LOCF	Last Observation Carried Forward
mITT	modified Intent-To-Treat
MedDRA	Medical Dictionary for Regulatory Activities
OGD	(FDA) Office of Generic Drugs
PP	Per Protocol
PT	Preferred Term
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SD	Standard Deviation
SOC	System Organ Class
STI	Sexual Transmitted Infection
TOC	Test Of Cure
WHO	World Health Organization

1.0 INTRODUCTION

The purpose of this document is to detail the planned statistical analysis and data presentations that will be performed to support the Clinical Study Report (CSR) for Gage, LLC study protocol GDC-229-002. The analyses detailed herein are based on the study protocol (Version 01.00, dated 23JAN2017) and electronic case report forms (eCRFs, dated 16FEB2017). Any post-hoc or unplanned analyses not identified in this SAP will be clearly identified in the CSR.

2.0 PROTOCOL SUMMARY

2.1 Background

Bacterial vaginosis (BV) is a common bacterial infection in women. BV is typically diagnosed using Amsel's criteria (Amsel 1983). The four Amsel's criteria used to determine a clinical diagnosis of BV are as follows: (1) vaginal pH > 4.5; (2) positive "whiff" test ("fishy" odor); (3) homogenous, non-viscous, off-white (milky or gray) discharge; (4) the presence of ≥ 20% clue cells.

Oral and topical antibiotics are used to treat BV. Metronidazole was approved for use in the United States in 1963 and is currently a first-line treatment for BV (Flagyl (metronidazole) NDA 012623, Approval Date 18JUL1963). Gage Development Company is developing a drug product candidate that is designed to be therapeutically equivalent to the marketed drug metronidazole 0.75% vaginal gel for the treatment of BV.

2.2 Objectives

The primary objective of this study is to assess the therapeutic equivalence of an experimental formulation of a metronidazole 0.75% vaginal gel (GDC-229; test product) versus metronidazole 0.75% vaginal gel (Oceanside Pharmaceuticals) in subjects with BV. A vehicle (placebo) arm is included for assay sensitivity purpose.

The secondary objective of this study is to evaluate the safety and tolerability of GDC-229 versus metronidazole 0.75% vaginal gel (Oceanside Pharmaceuticals) and vehicle (placebo) vaginal gel in subjects with BV.

2.3 Trial Design

This is a multicenter (with approximately 40 study sites), randomized, double-blind (for both subjects and investigators), vehicle-controlled bioequivalence study with a clinical endpoint. Following completion of the screening procedures, eligible subjects will be randomized in a blinded manner and in a 1:1:1 ratio to one of three treatment arms:

- GDC-229, the test product
- Reference product (metronidazole 0.75% vaginal gel by Oceanside Pharmaceuticals)
- Vehicle (placebo) vaginal gel.

There are three planned study visits as follows:

1. Study entry Visit (Day 1, screening and randomization);

2. Post-Treatment Assessment (Study days 8-15, i.e., 7 to 14 days after randomization), and
3. Test of Cure (Study days 22-31, i.e., 21 to 30 days after randomization).

Though early termination can occur anytime after enrollment, the study generally is up to 31 days in duration for each subject. The study treatment is to be administered intravaginally once daily at bedtime during Study Days 1 to 5.

All procedures and assessments associated with each study visit are shown in the Appendix 5.1 of this SAP. Including a recruitment period of about seven months, the study duration is expected to be approximately eight months.

2.4 Study Endpoints

2.4.1 Primary Endpoint

The primary study endpoint is "Clinical Cure", defined as the resolution at the Test of Cure visit of the following three abnormal BV signs included in Amsel's criteria:

- Return to normal physiological vaginal discharge as determined by the investigator
- Negative whiff test
- The presence of clue cells at <20% of the total epithelial cells on microscopic examination of the saline wet mount.

2.4.2 Secondary Endpoints

The secondary efficacy endpoints are the "Bacteriological Cure" and the "Therapeutic Cure" at the Test of Cure visit (Visit 3). Bacteriological Cure is defined as Nugent's score < 4 (via Gram stain central laboratory analysis). A Nugent's score ≥ 4 is considered abnormal and represents an imbalance in vaginal flora that permits BV to manifest (Nugent 1991). Subjects who are responders are expected to achieve a normal Nugent's score (i.e., 0 to 3) following the study treatment.

Therapeutic Cure is defined as having achieved Clinical Cure, vaginal pH of < 4.7, and Bacteriological Cure (Nugent's score < 4) at Visit 3.

2.5 Sample Size Consideration

Assuming the test and reference products' Clinical Cure rates are within 4% of each other, the placebo clinical cure rate is 20%, the test and reference clinical cure rates are at least 36.0% and 12.4% of randomized subjects do not qualify for the modified Intent-to-Treat (mITT) population (defined in Section 3.4 of this SAP), then 603 subjects will be required to achieve at least 90% power for superiority and therapeutic equivalence (Schwebke 2015; Chavoustie 2015). However, the previous studies that were used to estimate effect size had significant differences in study design and definition of primary endpoint. Taking these differences into consideration, approximately 738 subjects were planned to be randomized. The assumption of a 36% test and reference clinical cure rate was derived after review of the FDA's Statistical Review of Metronidazole vaginal gel 1.3% (Center for Drug Evaluation and Research 2014).

3.0 STATISTICAL METHODS

3.1 Statistical Handling Policy

3.1.1 Interim Safety Review

No formal interim analysis will be conducted. A blinded Medical Monitor will periodically perform a blinded review of adverse event data.

3.1.2 Analysis Conventions

This section details general approaches to be used for the statistical analyses. Departures from these general approaches may be outlined in the specific detailed sections of this statistical analysis plan, and will take precedence over the general approaches. The following approaches will be applied to all data presentations and analyses.

- Data listings will be provided for all CRF data with one listing for one CRF panel. All data listings will be sorted for presentation in the order of site identifier (ID), subject ID, and date of procedure or event.
- Statistical summary tables will be provided for most CRF data, and generated for each of the 3 treatment groups, and for overall subjects (i.e., 3 groups combined) when appropriate.
- Summary statistics will consist of the number and percentage of responses in each category for discrete (categorical) variables, and the number of non-missing observations (n), mean, median, standard deviation (SD), minimum, and maximum (abbreviated as "6-number statistics") for continuous variables.
- All mean values will be formatted to one more decimal places than the measured value, and standard deviation values will be formatted to two more decimal places than the measured value. For median values, the same decimal place as the measured values will be reported, if that is feasible without losing accuracy; otherwise, median values will be formatted to one more decimal place than the measured values.
- All percentages will be rounded to one decimal place. The number and percentage of categorical responses will be presented in the form XX (XX.X), where the percentage is in the parentheses. The denominator for percentage calculations will be the number of non-missing observations (n) when this number is shown, or will be the total sample size of the relevant treatment group for tables like adverse events, medical histories, or other tables where this number of non-missing observations (n) is not presented.
- All statistical tests will use a significance level of $\alpha = 0.05$. Two-tailed tests will be performed for all analyses that use statistical testing.
- All p-values will be rounded to 3 decimal places. All p-values that round to 0.000 will be presented as '<0.001', and p-values that round to 1.000 will be presented as '>0.999'. Any p-value ≤ 0.05 will be considered statistically significant and will be marked with one asterisk (e.g., 0.025*).

- All analysis and summary tables will have the population sample size in the column heading.
- Baseline is defined as the last data point on or before the day of randomization.
- Calculating change from baseline to a visit will be done as follows: Change from Baseline = Observed value at the visit – Baseline value.
- Subjects who have only baseline (without any post-baseline measurements) for a parameter at a visit, or only have post-baseline value for a parameter at a visit, will be excluded from the summary of change from baseline by visit for that parameter.
- Version 9.4 of SAS® or higher will be the statistical software package used to produce all statistical analysis tables and data listings.

3.2 Subject Disposition

Subject disposition will be summarized as follows: for the All Screened population with the following data:

- Number of inform consented subjects in each of the following 4 groups:
  - Screen failures
  - Placebo
  - GDC-229
  - Metronidazole
- Summary of screen failures by the reason for ineligibility for screen failed subjects
- The number and percentage of randomized subjects who completed or discontinued prematurely from the study, the number and percentage of subjects who discontinued by each reason for all subjects.
- A listing of randomized subjects who discontinued prematurely from the study. The listing will include information about study site ID number, subject ID number, age, number of doses used, and reason for discontinuation.
- The number and percentage of randomized subjects at each study visit.
- The number of randomized subjects who were enrolled, completed or discontinued prematurely at each study site.

The Disposition page of the CRF will be used to determine those subjects that discontinued prematurely from the study.

3.3 Protocol Deviations

Protocol deviations will be summarized by type of deviations listed on the CRF. Both the number of deviation events and distinct number and percentage of subjects will be presented for each type of deviation by treatment group and for all Intent-to-Treat subjects.

3.4 Analysis Populations

All Screened Population: all subjects who signed Consent Form and screened at the study sites.

Intent-to-Treat (ITT) Population: All randomized subjects.

Safety Population: All randomized subjects who received at least 1 dose of study drug. If the drug accountability is unknown (e.g., due to being lost to follow-up after being randomized) then subjects will be assumed to have taken study drug and included in the safety population. Subjects reporting 0 doses will be excluded.

Modified Intent-to-Treat (mITT) Population: a subset of the Safety Population who also met all 4 Amsel's criteria at baseline, but excluding those who subsequently demonstrate a positive or indeterminate test result for other concomitant vaginal or cervical infections at baseline (e.g. C. trachomatis, N. gonorrhoeae) which may interfere with the efficacy assessment for BV, or who have a baseline Nugent's score < 4. If baseline STI value is missing, use the Post-Treatment Assessment Visit infection test result. Subjects without a Nugent's score at baseline will be excluded from the mITT population.

Per Protocol (PP) Population: All mITT subjects except for those who had at least one of the protocol deviations listed below. If lack of efficacy (LOE) is the discontinuation reason and the subject took at least 3 doses, then the subject will be included in PP as a treatment failure and then the 6 protocol deviations below will not exclude those subjects:

1. Used topical or systemic antimicrobial drugs after randomization, this includes the following drug classes: antimycotics, antibacterials, antiprotozoals, gynecological anti-infectives, antiseptics and antifungals.
2. Used non-study vaginal products during the study
3. Did not start treatment within 2 days of randomization
4. Took less than 5 doses of study drug
5. Test of cure visit (V3) was not in the Study Day 22 – 31 window, i.e., Visit 3 was not 21 – 30 days after randomization
6. Test of cure visit (V3) not done

Other important violations that represent material non-compliance with study drug or protocol that would impact the primary and secondary efficacy endpoints will be reviewed prior to database lock (i.e., before unblinding). These will be documented in the Evaluability Memo which will be included in the clinical study report.

Subjects in the mITT population who discontinue early from the study, and do not have an assessment of Clinical Cure or Bacteriological Cure will be included in the mITT population as treatment failures. These subjects without actual outcomes, but are classified as treatment failures will not be included in the PP analysis population.

sm W W l W

Subject demographics and pre-treatment characteristics will be summarized for the ITT analysis population, and some of these summaries may also be done for other populations when deemed important.

smp

A summary of demographics and subject characteristics will be created for the ITT, Safety, mITT, and PP populations, and include age, race, ethnicity, height, weight, and body mass index (BMI) at the study entry visit with:

- 6-number statistics (number of non-missing observations, mean, median, SD, minimum, and maximum) for age, height, baseline weight, baseline BMI (to be calculated, see formula in Section 3.8.4) by treatment group and for all subjects.
- Age will be calculated (with the INTCK function in SAS®) using the study entry visit date and the date of birth. Subject age (years) will also be categorized as 18-20, 21-30, 31-40, 41-50, 51+, and then summarized categorically along with race, ethnicity.
- BMI at the study entry visit will also be categorized as: Underweight (< 18.5), Normal (18.5 – 24.9), Overweight (25.0 – 29.9), and Obese (≥ 30.0). The number and percentage of subjects in each of these categories will be presented by treatment group and for all subjects.

3.5.2 Medical History

Medical history data will be coded by using the Medical Dictionary for Regulatory Activities (MedDRA) Version 20.0. The codings will be reviewed by qualified medical personnel, and then summarized by system organ class (SOC) and preferred term (PT) and by the medical history status "Ongoing", or "Resolved" by treatment group and for all subjects. If a subject reports more than one SOC/PT, that subject will be counted only once toward the status "Ongoing" for the summary of that SOC/PT.

3.5.3 Gynecological History

Gynecological history data will be summarized for key parameters collected in the CRF, with categorical (qualitative) parameters as:

- Is there a history of Bacterial Vaginosis?
- Is there a history of Sexually Transmitted Infections?

and continuous (quantitative) parameters as:

- Approximate number of times (per month) you have vaginal intercourse
- Current total number of sexual partners

The number and percentage of ITT subjects will be provided by treatment group and for all subjects for each response to the question for categorical data, and 6-number statistics for continuous data.

3.5.4 Pregnancy History

Pregnancy history data includes only 2 parameters:

- Number of times subject has been pregnant
- Date last pregnancy ended.

The first parameter will be summarized by 6-number statistics and also categorically (as 0, 1, 2, 3, etc.) for ITT subjects. The second one's Year part (e.g., 1998) will be used to get "Number of years since the last pregnancy ended", which is to be approximated as 2017 – Year of last pregnancy ended. Then the "Number of years since the last pregnancy ended" will be summarized by 6-number statistics.

3.5.5 Substance Use

Substance use include tobacco, alcohol, and illicit drugs. Categorical responses, e.g., "Never", "Current", or "Former", will be summarized by number and percentage of ITT subjects in each response/category by treatment group and for all subjects.

3.5.6 Prior Medications

All medications taken by the subject for one month prior to randomization (pre-study) and during the study will be recorded on the prior and concomitant medication CRF page. Prior medications are medications that were taken up to one month before randomization, and concomitant medications are medications taken on or after the subject was randomized.

Medications will be coded to the therapeutic drug classes and generic drug names by using the World Health Organization (WHO) Drug Dictionary, March 2016 version. The codings will be reviewed by qualified medical personnel, and then summarized by treatment group and for all subjects with the number and percentage of ITT subjects who had prior medications that were coded to each generic drug name and therapeutic drug class, as well as the number and percentage of subjects who had at least one prior medication. Subjects reporting more than one drug in each drug class/generic name are only counted once to that drug class/generic name.

Prior medications will be also be summarized by the descending order of frequencies (total number of ITT subjects) for drug class (i.e., no generic drug names provided).

For the purpose of assessing prior and concomitant medications, the following rules will be used for partial or missing medication start and stop dates.

D=Day,Y=Year,M=Month

Parameter	Missing	Comments	Imputation
Start date	D only	M and Y same as M and Y of first dose of study drug	Date of first dose of study drug
		M and/or Y not same as date of first dose of study drug	First day of month
	M and D	Y same as Y of first dose of study drug	Date of first dose of study drug
		Y not same as Y of first dose of study drug	Use Jan 01 of Y
	M, D, and Y	None - date completely missing	Day prior to date of first dose of study drug

Stop date	D only	M and Y same as M and Y of last dose of study drug	Date of last dose of study drug
		M and/or Y not same as date of last dose of study drug	Last day of month
	M and D	Y same as Y of last dose of study drug	Date of last dose of study drug
		Y not same as Y of last dose of study drug	Use Dec 31 of Y
	M, D, and Y	None - date completely missing and NOT ongoing	Date of last dose of study drug

3.6 Concomitant Medications

As defined in the above Section 3.5.6, concomitant medications are medications taken on or after the subject was randomized to study treatment. Similar to prior medications, concomitant medications will be summarized for Safety populations by the number and percentage of subjects in each coded drug name and drug class. They will also be summarized by the descending order of frequencies (total number of Safety subjects) for drug class (i.e., no generic drug names provided).

A medication may be both "prior" and "concomitant" if a subject took it prior to the randomization (pre-study) and also used it again on/after randomization. Another summary table with unique number and percentage of subjects for each treatment group and all subjects will be presented for medications belonging to both "prior" and "concomitant" (i.e., a medication will not be included in the table unless a subject took it before randomization as well as on or after randomization).

3.7 Efficacy Analyses

The mITT population will be used to compare both the test and reference products to placebo. The PP population will be used to evaluate therapeutic equivalence between the test and reference product.

3.7.1 Primary Efficacy Analyses

Primary efficacy endpoint is the "Clinical Cure" at the Test of Cure (TOC) visit (Visit 3) and includes the following 3 parameters from the "Study Procedure" CRF page:

- Resolution of the abnormal vaginal discharge to a normal physiologic discharge, which is from the variable *SPCDBV* (*Was there clinical diagnosis of bacterial vaginosis?*), with value of "No"
- Negative whiff test, which is from the variable *SPWT* (*Was there a positive whiff test after adding a drop of 10% KOH to vaginal discharge?*), with value of "No"

- Presence of clue cells at < 20% of the total epithelial cells on microscopic examination of the saline wet mount, which is from the variable *SPSWM*, with value of "No".

To investigate therapeutic equivalence, the 90% confidence interval (CI) for difference, [test - reference], in proportion of subjects achieving "Clinical Cure" will be calculated, using the PP analysis population. Adopted from FDA Office of Generic Drugs Draft Product Specific Guidance for Generic Drug Development on Metronidazole (Recommended April 2013) document, the compound hypothesis to be tested here is:

H_0: $p_T - p_R < -0.20$ or $p_T - p_R > 0.20$

versus

H_A: $-0.20 \leq p_T - p_R \leq 0.20$

where p_T = cure rate of test treatment and p_R = cure rate of reference treatment.

Let

n_T – sample size of test treatment group

cn_T = number of cured subjects in test treatment group

n_R = sample size of reference treatment group

cn_R = number of cured subjects in reference treatment group

$$\hat{p}_T = cn_T / n_T, \quad \hat{p}_R = cn_R / n_R,$$

$$\text{and se} = (\hat{p}_T(1 - \hat{p}_T)/n_T + \hat{p}_R(1 - \hat{p}_R)/n_R)^{1/2}$$

The 90% confidence interval for the difference in proportions between test and reference was calculated as follows, using Yates' correction

$$L = (\hat{p}_T - \hat{p}_R) - 1.645\ \text{se} - (1/n_T + 1/n_R)/2$$

$$U = (\hat{p}_T - \hat{p}_R) + 1.645\ \text{se} + (1/n_T + 1/n_R)/2$$

If the lower limit $L \geq -0.20$ and upper limit $U \leq 0.20$, we will reject null hypothesis H_0 and claim the bioequivalence ("Therapeutic Equivalence") of test treatment GDC-229 and reference treatment.

Two-sided 90% confidence interval (CI), with continuity correction, for the "Clinical Cure" rate of each treatment arm will also be presented, with the following formula:

$$\hat{p} \pm (z_{\alpha/2} \times \mathrm{se}(\hat{p}) + (1/2n))$$

,

where p is the percentage (proportion) of subjects with "Clinical Cure", $z_{\alpha/2}$ is the standard normal distribution's 5th percentile (with α = 0.1), and n is the analysis population size in the relevant treatment group, with standard error for the proportion.

$$\mathrm{se}(\hat{p}) = \sqrt{\hat{p}\,(1-\hat{p})\,/\,n}$$

.

PROC FREQ (SAS 9.4) will be employed for the analysis. Proportion of subjects with "Clinical Cure" in test product GDC-229 and reference drug will also be compared against that in placebo group, and p values from chi-square test, or Fisher's exact test - if any cells have expected frequency of < 5, will be presented. As a parameter for determining adequate study sensitivity, both test product and the reference drug need to have higher "Clinical Cure" rates than placebo and with p values < 0.05, using the mITT analysis population and Last Observation Carried Forward (LOCF), which means if no assessments are available at End of Trial visit (either Visit 3 for completers, or Early Termination (visit 4) for discontinued subjects), then the Visit 2 assessment results will be used. A 95% CI for the "Clinical Cure" rate of each treatment arm will be presented for mITT subjects. The formula for this 95% CI will be the same as above 90% CI except for replacing the α value with 0.05.

3.7.2 Secondary Efficacy Analyses

The endpoints for secondary efficacy analyses are

- Bacteriological Cure, i.e., Nugent's score < 4
- Therapeutic Cure, defined as follows:
  - Clinical cure as defined in the Primary Endpoint
  - Vaginal pH < 4.7
  - Nugent's score < 4 (i.e., Bacteriological Cure)

Total number and proportion of subjects achieving "Bacteriological Cure", "Therapeutic Cure" (also each of the 3 components) will be presented for mITT population as well as PP population by study visit and treatment arm.

Nugent's score will also be summarized categorically by its raw score (0, 1, 2, …,10). All results from the Amsel's 4 criteria will also be summarized categorically ("Yes"/ "No") by study visit and treatment arm for both mITT population and PP population.

3.7.3 Other Outcome Measurements

The other assessments as follows (each with "Yes", or "No" answer) on "Study Procedure" form will be presented in data listing:

- Was there a vaginal secretion with pH greater than 4.5?
- Was the saline wet mount assessed for evidence of any other vulvovaginitis other than BV?

- Did the wet mount contain any evidence for candidiasis, Trichomonas vaginalis, Chlamydia trachomatis, Neisseria gonorrhoeae, active herpes simplex, or active HPV?
- Was the urine drug screen performed?
- Were any of the results positive?
- In the last 48 hours have you used any spermicides, tampons, douches, diaphragms, condoms or other intravaginal product?
- Since your last visit have you used any spermicides, tampons, douches, diaphragms, condoms or other intravaginal product?
- Have you had intercourse within the 48 hours prior to a visit?

3.8 Safety Analyses

All safety analyses will be performed for the Safety population unless noted otherwise.

3.8.1 Treatment Exposure and Compliance

Exposure to study treatment will be summarized categorically by

- the number of doses (1, 2, 3, etc.) actually used
- the length of time (# of days) over which the product is used.

The length of time over which the product is used will be calculated as Treatment Stop Date – Treatment Start Date +1.

Treatment Compliance for each subject in Safety population will be calculated as:

Number of doses actually used / 5 x 100%,

where denominator 5 is the total number of doses supposed to be taken by a randomized subject in this study. The following information for each treatment group will be presented for Treatment Compliance:

- Six-number statistics (sample size, mean, median, SD, minimum, and maximum)
- The number and percentage of subjects in each of the following 4 categories of treatment compliance: <60%, 60% - <100%, 100%, and >100%.

All drug accountability (DA) data on DA form will be presented in a data listing.

3.8.2 Adverse Events and Serious Adverse Events

A treatment-emergent adverse event, (TEAE) is any symptom, sign, illness, or experience that develops or worsens in severity and/or frequency after the study treatment started. A serious adverse event (SAE) is any adverse event that results in any of the following outcomes: Death; immediate threat to life; inpatient hospitalization or prolongation of an existing hospitalization; persistent or significant disability/incapacity; congenital anomaly/birth defect; other serious (medical important) AEs. All AEs will be coded with system organ class (SOC) and preferred term (PT) from the MedDRA Version 20.0. The codings will be reviewed by qualified medical personnel.

Incidence of AEs and SAEs will be summarized by SOC and PT with the number and percentage of subjects with each SOC and PT by treatment arm. The following tables will be presented for Safety population:

- AEs by SOC and PT
- AEs that led to study discontinuation or death by SOC and PT
- AEs by SOC and PT by severity
- AEs by SOC and PT by relatedness to the study treatment
- AEs in descending order of SOC frequencies (no PT displayed)
- AEs in descending order of PT frequencies (no SOC displayed)

All the above tables (except for where specified) are counting number of distinct subjects in each SOC and PT. Subjects reporting more than one AE/SAE in each PT will be counted only once to that PT, using the most severe intensity, for unique number of subjects counting. The only exception to this will be for the summary by relatedness to the study treatment, where subjects will be counted only once using the strongest relatedness to the study treatment for the purpose of counting distinct number of subjects. The same principle also will be applied to the summary at the SOC level.

An overall AE high-level summary table will also be provided for Safety populations with total number and percentage of distinct subjects (but no SOC or PT) for

- AE severity (Mild, Moderate, Severe)
- AE relatedness (Not Related, Possibly Related, Related)
- SAE
- Death, or discontinued study due to any AE

Another high-level summary table for Safety population will present the counts of events (not distinct subjects) as the following chart. Note that each percentage, if provided, is based on the number on the one where "(100.0)" is displayed. This AE summary table, done for each treatment group and all Safety subjects, provides the cross reference between every level of relatedness and severity.

Relatedness (to the study treatment)	AE Severity			
	Mild	Moderate	Severe	Total # (%) of Events
Not Related	xx	xx	xx	xx (x.x)
Possibly Related	xx	xx	xx	xx (x.x)
Related	xx	xx	xx	xx (x.x)
Total # (%) of Events	xx	xx	xx	xx (100.0)

Two listing-style tables, one is for SAEs, and the other is for AEs that led to premature study discontinuation, will also be presented, with the details about the event onset date, resolved date, study day of onset since the study treatment (calculated as Onset Date – Randomization Date +1), severity, outcome, medical intervention needed for the AE, and relatedness to the study treatment, as well as other supportive data such as the subject's age and number of days on the study.

3.8.3 Local Site Reaction

Local site reactions include any reaction that occurs within the treatment area of subjects. Local site reactions are prespecified in the protocol and will be collected independently of

adverse events. Local site reactions that require medical intervention or extend beyond the treatment area will be documented as adverse events. At Visits 2 (Post-Treatment Assessment) and 3 (Test of Cure), subjects will be assessed at treatment area for the following 7 symptoms:

- Erythema
- Petechiae
- Erosion/ulceration
- Edema
- Burning/stinging
- Pain
- Pruritus (itching)

Subjects need to rate them as

- 1 = Absent,
- 2 = Mild (slight, barely perceptible),
- 3 = Moderate (distinct presence), or
- 4 = Severe (marked, intense).

These data, along with the following 2 questions on the same CRF, will be summarized categorically by visit.

- Did the local site reaction extend beyond the treatment area and was recorded as an AE? (“Yes”, “No”)
- Did the local site reaction require medical intervention? (“Yes”, “No”)

3.8.4 Vital Signs and Weight

Vital signs measurements, done at all study visits, include heart rate, pulse rate, respiratory rate, body temperature, systolic and diastolic blood pressures, in addition to height and weight. BMI (unit: kg/m^2) will be calculated from height (inch) and weight (lb):

$$BMI = Weight\ (lb) / (Height\ (inches))^2 \times 703$$

Then it will be categorized as: Underweight (< 18.5), Normal (18.5 – 24.9), Overweight (25.0 – 29.9), and Obese (≥30.0). Both numerical and categorical summaries for BMI will be presented. The 6-number summary statistics for continuous variables and their changes from baseline, as well as the number and percentage of subjects in each BMI category will be presented by study visit and treatment arm.

3.8.5 Physical Examination

Physical examination will be performed at all study visits in the following body areas:

- Dermatological
- Respiratory
- H.E.E.N.T
- Cardiovascular
- Gastrointestinal
- Hepatic/Biliary
- Genitourinary/Reproductive

- Renal
- Neurological/Psychological
- Musculoskeletal
- Metabolic
- Hematological
- Immunological/Allergies
- Recent Trauma (past 3 months)
- Prior Surgery
- Congenital
- Other

The examination results ("Normal", "Abnormal"), and changes from baseline ("Improved" – if "Abnormal" at baseline and "Normal" at post-baseline, "No Change", "Worsened" – if "Normal" at baseline and "Abnormal" at post-baseline), will be summarized by presenting the number and percentage of subjects in each category by body area and treatment at each study visit.

3.8.6 Gynecological Exam

Gynecological exam will be done at each study visit in the following areas:

- Breasts
- Adnexae
- Vulva
- Vaginal wall
- Cervix
- Uterus
- Rectal

The examination results ("Normal", "Abnormal"), and changes from baseline ("Improved" – if "Abnormal" at baseline and "Normal" at post-baseline, "No Change", "Worsened" – if "Normal" at baseline and "Abnormal" at post-baseline), will be summarized categorically by visit.

For "Abnormal" results, a further question "Was the result clinically significant?" will also be summarized with number of "Yes" subjects.

3.8.7 Other Safety Assessments

Other study assessments, such as pregnancy test (and pregnancy outcome, if known), and urine drug screen will be provided in data listings.

4.0 REFERENCES

1. Amsel R, Totten PA, Spiegel CA, et al. Nonspecific vaginitis: diagnostic criteria and microbial and epidemiologic associations. Am J Med. 1983;74(1):14-22.
2. Center for Drug Evaluation and Research. Application Number 205223Orig1s000. Statistical Reviews. Statistical Review and Evaluation of metronidazole vaginal gel 1.3%. Review date January 18, 2014.

3. Chavoustie SE, Jacobs M, Reisman HA, et al. Metronidazole vaginal gel 1.3% in the treatment of bacterial vaginosis: a dose-ranging study. J Lower Genital Tract Disease. 2015;19(2):129-134.
4. Korn AP, Bolan G, Padian N, et al. Plasma cell endometritis in women with symptomatic bacterial vaginosis. Obstet Gynecol. 1995;85(3):387–90.
5. Nugent RP, Krohn MA, and Hillier SL. Reliability of diagnosing bacterial vaginosis is improved by a standardized method of gram stain interpretation. J Clin Micro. 1991;29(2):297-30.
6. Schwebke JR, Marrazzo J, Beelen AP, et al. A Phase 3, multicenter, randomized, double-blind, vehicle-controlled study evaluating the safety of efficacy of metronidazole vaginal gel 1.3% in the treatment of bacterial vaginosis. Sexually Transmitted Diseases. 2015;42(7):376-381.

5.0 APPENDICES

5.1 Study Schedule of Assessments

Visit	Visit 1	Visit 2	Visit 3	
Visit Name	**Entry Visit**	**Post-Treatment Assessment**	**Test of Cure**	**Early Termination[a]**
Study Day(s)	**1**	**8 to 15 days**	**22 to 31 days**	
Informed consent	X			
Inclusion/exclusion criteria	X			
Demographics	X			
Medical and gynecological history	X			
Physical examination	X		X	X
Pelvic examination	X	X	X	X
Height	X			
Weight	X		X	X
Vital signs (after sitting for 5 - 10 mins)	X	X	X	X
Concomitant medications	X	X	X	X
Randomization[b]	X			
Dispense study medication[c] and dosing instructions	X			
Collect study medication		X		X[d]
Assessment of local site reactions		X	X	X
Adverse events	X	X	X	X
Urine pregnancy test	X	X	X	X
Urine drug screen	X			
Amsel's criteria evaluations:				
-Assessment of vaginal discharge	X	X	X	X
-KOH Whiff Test	X	X	X	X
-Saline Wet Mount	X	X	X	X
-pH	X	X	X	X
Collection of samples for Nugent's score based on Gram stain (send to central laboratory)	X		X	X
Collection of samples for STI assessment[e] (send to central laboratory)	X			

a. At any time a subject decides to discontinue her participation in the study or is discontinued from the study for any other reason, she will be advised to complete the Early Termination procedures.
b. Randomization to occur if the results of the full set of Amsel's criteria are positive.
c. Study medication will be dispensed at Visit 1. The subject will be instructed to self-administer study medication intravaginally once daily at bedtime for 5 consecutive days.
d. Unless collected at previous visit.

e. Central laboratory analysis for *Chlamydia trachomatis* and *Neisseria gonorrhoeae.*

5.2 Table of Contents for Data Display

Following is the list of planned tables and listings. Tables will be numbered according to the nomenclature used to support the final CSR. The datasets in the OGD Guidance will be created and used in our programming of the tables in this SAP.

The format of each unique table is provided in a separate document of "Table Shells", but no data listing shells are provided. Final outputs may be slightly different in layout from that of illustrated in "Table Shells". The table shells will not be amended to match the actual tables in such cases.

Table Number	Table Title	Analysis Population
14.1.1	Summary of Subject Disposition and Reasons for Discontinuation	All Screened
14.1.1.1	Screen Failure Reason Summary	Screen Failure
14.1.2	List of Subjects Who Prematurely Discontinued the Study	ITT
14.1.3	Summary of Subject Disposition by Study Site	ITT
14.1.4	Number and Percentage of Subjects at Each Visit	ITT
14.1.5	Summary of Protocol Deviations	ITT
14.1.6	Summary of Analysis Populations	ITT
14.1.7	Demographics	ITT
14.1.7S	Demographics	Safety
14.1.7M	Demographics	mITT
14.1.7P	Demographics	PP
14.1.8.1	Medical History	ITT
14.1.8.2	Gynecological History	ITT
14.1.8.3	Pregnancy History	ITT
14.1.8.4	Substance Use	ITT
14.1.9.1C	Number and Percentage of Subjects with Prior Medication Use by Drug Class in Descending Order of Frequencies	ITT
14.1.9.1G	Number and Percentage of Subjects with Prior Medication Use by Drug Generic Name in Descending Order of Frequencies	ITT
14.1.9.1	Number and Percentage of Subjects with Prior Medication Use by Drug Class and Generic Name	ITT
14.1.9.2C	Number and Percentage of Subjects with Concomitant Medication Use by Drug Class in Descending Order of Frequencies	Safety

14.1.9.2G	Number and Percentage of Subjects with Concomitant Medication Use by Drug Generic Name in Descending Order of Frequencies	Safety
14.1.9.2	Number and Percentage of Subjects with Concomitant Medication Use by Drug Class and Generic Name	Safety
14.1.9.3C	Number and Percentage of Subjects with Prior Medications Also Used After the Randomization by Drug Class in Descending Order of Frequencies	Safety
14.1.9.3G	Number and Percentage of Subjects with Prior Medications Also Used After the Randomization by Drug Name in Descending Order of Frequencies	Safety
14.1.9.3	Number and Percentage of Subjects with Prior Medications Also Used After the Randomization by Drug Class and Generic Name	Safety
14.2.1	Analysis of Therapeutic Equivalence	PP
14.2.2	Analysis of Study Sensitivity	mITT
14.2.3M	Summary of Bacteriological Cure and Therapeutic Cure	mITT
14.2.3P	Summary of Bacteriological Cure and Therapeutic Cure	PP
14.2.4M	Nugent's Score and Amsel's Criteria Summary by Study Visit	mITT
14.2.4P	Nugent's Score and Amsel's Criteria Summary by Study Visit	PP
14.3.0	Treatment Exposure and Compliance	Safety
14.3.1.1	Overall Number and Percentage of Subjects with Adverse Events	Safety
14.3.1.2	Overall Number of Adverse Events by Severity and Relatedness to the Study Treatment	Safety
14.3.1.3C	Summary of Adverse Events by System Organ Class in Descending Order of Frequencies	Safety
14.3.1.3P	Summary of Adverse Events by Preferred Term in Descending Order of Frequencies	Safety
14.3.1.4	Summary of Adverse Events by System Organ Class and Preferred Term	Safety
14.3.1.5C	Summary of Adverse Events by System Organ Class and by Severity	Safety
14.3.1.5	Summary of Adverse Events by System Organ Class and Preferred Term and by Severity	Safety
14.3.1.6C	Summary of Adverse Events by System Organ Class and by Relatedness to the Study Treatment	Safety

14.3.1.6	Summary of Adverse Events by System Organ Class and Preferred Term and by Relatedness to the Study Treatment	Safety
14.3.2.1	List of Serious Adverse Events	Safety
14.3.2.2	List of Adverse Events that Led to Premature Study Discontinuation	Safety
14.3.5	Local Site Reaction	Safety
14.3.6	Vital Signs and Weight	Safety
14.3.7	Physical Examination	Safety
14.3.8	Gynecological Examination	Safety

Listing Number	**Listing Title**	**Analysis Population**
16.2.1.1	Subject Information	All Screened
16.2.1.2	Inclusion / Exclusion Criteria Not Met	All Screened
16.2.1.3	Subject Disposition	ITT
16.2.1.4	Subject Study Visits	ITT
16.2.2	Protocol Deviations	ITT
16.2.3	Subjects Excluded from the Efficacy Analysis	ITT
16.2.4.1	Demographics	ITT
16.2.4.2	Medical History	ITT
16.2.4.3	Gynecological History	ITT
16.2.4.4	Pregnancy History	ITT
16.2.4.5	Prior and Concomitant Medications	ITT
16.2.4.6	Substance Use	ITT
16.2.5.1	Drug Accountability	ITT
16.2.5.2	Treatment Exposure	ITT
16.2.6.1	Study Procedures	ITT
16.2.7.1A	Adverse Events – Part A	ITT
16.2.7.1B	Adverse Events – Part B	ITT
16.2.7.2	Local Site Reaction	ITT
16.2.8.1	Nugent's Scores	ITT
16.2.8.2	Sexual Transmitted Infection Tests	ITT
16.2.9.1	Vital Signs	ITT
16.2.9.2	Physical Examination	ITT
16.2.9.3	Gynecological Examination	ITT
16.2.9.4	Pregnancy Test	ITT

16.2.9.5	Pregnancy Notification	ITT
16.2.9.6	Pregnancy Outcome	ITT