A RANDOMIZED, DOUBLE-BLINDED, VEHICLE-CONTROLLED, PARALLEL-GROUP, MULTICENTER STUDY TO COMPARE PERRIGO UK FINCO'S ESTRADIOL VAGINAL CREAM 0.01% TO ESTRACE® (ESTRADIOL) VAGINAL CREAM, USP, 0.01% (WARNER CHILCOTT (US), LLC) AND BOTH ACTIVE TREATMENTS TO A VEHICLE CONTROL IN THE TREATMENT OF VULVAR AND VAGINAL ATROPHY

NCT02770365

02-08-2016

STATISTICAL ANALYSIS PLAN

PRG-NY-15-007: Estradiol Vaginal Cream 0.01%

[redacted]

A Randomized, Double-Blinded, Vehicle-Controlled, Parallel-Group Multicenter Study To Compare Perrigo UK FINCO's Estradiol Vaginal Cream 0.01% To ESTRACE® (estradiol) Vaginal Cream, USP, 0.01% (Warner Chilcott (US), LLC) And Both Active Treatments To A Vehicle Control In The Treatment Of Vulvar And Vaginal Atrophy

Perrigo New York, Inc.
1701 Bathgate Ave.
Bronx, NY 10457

[redacted]

[redacted]

[redacted]

Table of Contents

[redacted]

List of Abbreviations

AE	Adverse Event
ANOVA	Analysis of Variance
CI	Confidence Interval
CMH	Cochran–Mantel–Haenszel Test
ITT	Intent-to-Treat (Population)
LOCF	Last Observation Carried Forward
MBS	Most Bothersome Symptom
MedDRA	Medical Dictionary for Regulatory Activities
mITT	Modified Intent-to-Treat (Population)
PD	Protocol Deviation
PP	Per-Protocol (Population)
PV	Protocol Violation
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
TEAE	Treatment-Emergent Adverse Event
TESAE	Treatment-Emergent Serious Adverse Event
WHO Drug	World Health Organization Drug Dictionary
VVA	Vulvar and Vaginal Atrophy

Statistical Analysis Plan

1 Purpose of Statistical Analysis Plan

The purpose of the statistical analysis plan is to describe in detail all the data, statistical methods, and summary tables required to implement the statistical analysis of Clinical Study Protocol PRG-NY-15-007 [redacted].

2 Study Objectives

To demonstrate bioequivalence of Perrigo UK FINCO's Estradiol Vaginal Cream 0.01% compared with ESTRACE (estradiol) Vaginal Cream, USP, 0.01% (Warner Chilcott (US), LLC) in the treatment of vulvar and vaginal atrophy and to demonstrate superiority of the two active treatments over Vehicle.

3 Study Design and Sample Size

3.1 Study Design

The study will be conducted as a double-blind, randomized, parallel-group, vehicle-controlled, multicenter trial.

Each subject will be randomly assigned to one of following treatment groups [redacted]:

(1) Test: Estradiol Vaginal Cream 0.01%, [redacted]
(2) Reference: ESTRACE (estradiol) Vaginal Cream, USP, 0.01%, manufactured by Warner Chilcott (US), LLC
(3) Vehicle: Vehicle of the test product, [redacted]

Subjects will be admitted into the study only after written informed consent has been obtained and all of the inclusion and none of the exclusion criteria have been met. Randomization will be performed according to a computer generated randomization scheme where the treatment group designation has been assigned to the subject number. The treatment designation will remain blinded until after the final database is locked. An independent third party generator will generate and hold the randomization code throughout the study. Randomized subjects will apply 2 grams of the assigned treatment once daily intravaginally at bed time for 7 days at approximately the same time each day.

Subjects will come to the study site for clinical evaluations at Visit 1/Screening (Day -14 to Day -1), Visit 2/Randomization (Day 1), and Visit 3/End of Study (Day 8, +3 days) or at early discontinuation.

3.2 Sample Size

[redacted] subjects will be screened to randomize [redacted] subjects [redacted] to obtain [redacted] mITT subjects and 5 [redacted] PP subjects. [redacted]

[redacted]

4 Populations To Be Analyzed

Three subject populations are defined as follows:

(1) Intent-to-treat (ITT) population: any subject who: (a) was randomized into the study and (b) applied at least 1 dose of assigned study medication intravaginally;

(2) Modified Intent-to-treat (mITT) population: any subject who (a) was randomized in the study (b) applied at least 1 dose of assigned study medication intravaginally, (c) completed at least one post randomization assessment, and (d) met eligibility criteria;

(3) Per Protocol (PP) population; any subject who (a) was randomized in the study and met all inclusion/exclusion criteria, (b) had not taken any concomitant medications prohibited by the protocol or had any other significant protocol violations; (c) returned for Visit 3/End of Study within the designated visit window [redacted] with data on the primary efficacy endpoint and; d) had a study medication application compliance [redacted]

Study Medication is packaged in blocks of [redacted] of Test:Reference:Vehicle. Study medication assignment is blinded.

[redacted]

5 Planned Analyses

5.1 Methodological Considerations

All randomized subjects who received study medication will be evaluated for safety. The efficacy analysis will be conducted on both the PP and the mITT populations. No adjustments of p-values for multiple comparisons will be made. No interim analyses are planned. Two-sided hypothesis testing will be conducted for all the tests. Resulting p-values $\leq$ than 0.05 will be considered statistically significant unless noted otherwise. SAS software will be used for all data analyses and tabulations.

5.2 Handling of Dropouts or Missing Data

Missing efficacy data will be imputed via the LOCF method in the mITT population. For the PP population, subjects who discontinued early due to treatment failure will be considered as non-responders in the primary efficacy endpoint and a treatment failure in the secondary efficacy endpoint while missing values for other efficacy variables will not be imputed. [redacted]

For demographic and baseline characteristics and the safety profile, each variable will be analyzed using all available data. Subjects with missing data will be excluded only from the analyses for which data are not available.

5.3 Demographics and Baseline Characteristics

Baseline variables (e.g., age, ethnic origin) will be summarized descriptively by treatment group. For each continuous variable, the summary will include the mean, standard deviation, minimum and maximum. For each categorical variable, the summary will include frequencies and percentages.

5.4 Subject Accountability

A summary of subject disposition will be provided for all subjects. Descriptive summaries of subject disposition, reason for discontinuation, and analyses population will be provided by treatment group. The data will also be presented in subject data listings.

5.5 Efficacy Variables and Analyses

5.5.1 Primary Endpoint

The primary endpoint of the study is the proportion of subjects identified as responders at Visit 3/End of Study. A responder is defined as a subject with at least a 25% reduction from baseline (Visit 1/Screening) in the sum of % (percent) basal/parabasal + % (percent) intermediate cells on vaginal cytology AND vaginal pH < 5.0 with a change from Visit 1/Screening vaginal pH of at least 0.5.

Equivalent Efficacy
The compound hypothesis to be tested for clinical equivalence between test and reference 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 and p_R are the proportions of responders at Visit 3/End of Study (Day 8) for the test and reference products, respectively. The test product will be considered to be clinically equivalent to the reference product if the 90% CI on the difference in their proportions of responders, calculated by the Wald's method with Yates' continuity correction, is contained within the limits -0.20 to +0.20. Rejection of the null hypothesis supports the conclusion of therapeutic equivalence between the test and reference products for the primary efficacy variable. [redacted]

Analysis for bioequivalence will be performed based on the following SAS code (SAS Institute v.9.1.3):

P-value is chosen from Continuity Adj. Chi-Square test.

The SAS code for 90% confidence interval (trt=1 for Test, trt=2 for Reference):
proc freq data=XX ;

```
  where trt in (1,2);
  tables trt* success / alpha=0.10 riskdiff;
  output out=CIDIFF (keep= l_rdif2 u_rdif2 ) riskdiff;
run;
```

For the final 90% continuity-corrected CI, the lower limit = (l_rdif2 - yates) and the upper limit = (u_rdif2 + yates), where yates, the Yates' continuity-correction factor, is derived as (1/n1+1/n2)/2, n1=number of subjects in Test arm and n2=number of subjects in Reference arm.

Superiority

The hypotheses to be tested for superiority of the test and reference products over Vehicle are:

$H_0: p_T \leq p_V$ versus $H_A: p_T > p_V$

$H_0: p_R \leq p_V$ versus $H_A: p_R > p_V$

Where p_T, p_R and p_V are the proportions of responders at Visit 3/End of Study (Day 8) for the test, reference and Vehicle products, respectively. The tests will be conducted independently for the test product and the reference product using two-sided, $\alpha = 0.05$, continuity-corrected Z-tests. Superiority will be established if the proportion of responders in the active treatment group is greater and statistically different ($P \leq 0.05$) than that in the Vehicle. Rejection of the null hypothesis supports the conclusion of superiority of the test and reference products over the Vehicle product for the primary efficacy variable.

Superiority analyses will be conducted on both the mITT and the PP populations. [redacted]

Analysis for bioequivalence will be performed based on the following SAS code (SAS Institute v.9.1.3):

The SAS code for p-value (trt=2 for Reference vs. trt=3 for Vehicle):

```
proc freq data=XX ;
  where trt in (2,3);
  table trt * success /chisq ;
run;
```

In the mITT population, a LOCF approach will be used to impute missing efficacy data. PP subjects who discontinue due to treatment failure will be included as non-responders for the primary endpoint. No formal statistical analyses are planned to evaluate the consistency of efficacy results across the multiple centers for the primary efficacy variable. These results, however, will be tabulated and, if a center's response is obviously inconsistent with those of the other centers, this will be explored and addressed in the final study report.

5.5.2 Secondary Endpoints

The secondary endpoint will be the proportion of subjects with treatment success based on the improvement (change from baseline (Visit 2/Randomization)) of the Most Bothersome Symptom (MBS) of VVA at Visit 3/End of Study. Treatment success is defined as a change

from baseline which results in a score of 0 (none) or 1 (mild) at Visit 3 if the MBS is vaginal dryness, vaginal or vulvar irritation or itching, dysuria and vaginal pain associated with sexual activity. If the MBS is vaginal bleeding associated with sexual activity then treatment success will be absence of vaginal bleeding associated with sexual activity as reported at the end of the study. The same methods of analysis as for the primary analyses will be conducted for the secondary endpoint.

5.6 Safety Variables and Analyses

Duration of Treatment and Medication Compliance

Number of applications, days of exposure, and compliance rate will be summarized by treatment group using descriptive statistics. For each subject, the overall duration of treatment (days) will be calculated using the following formula:

(Date of last application of study medication) - (Date of first application of study medication) + 1.

Medication compliance rate (%) will be calculated for each subject as follows:

(Total number of applications used) / (Expected number of applications) *100%.

Subjects who complete the study are expected to have 7 applications. For prematurely discontinued subjects, expected number of applications will be determined based on the expected number of applications by the time of discontinuation, i.e. the overall duration of treatment. Descriptive summaries of exposure and medication compliance rate will be provided by treatment group for the ITT subjects.

Adverse Events

Adverse events (AEs) will be coded in MedDRA, version 18.1. Treatment-Emergent Adverse Event (TEAE) is defined as any AE occurs on or after application of the first dose of study drug. Number and percent of subjects reporting TEAEs will be tabulated by treatment group. Summaries will be presented by body system and preferred term for the ITT population, and further by severity and relationship to study medication. TEAEs reported by 5% or more subjects for any treatment group will also be summarized. In the summaries of incidence rates (frequencies and percentages) by severity and relationship to study drug, subjects who report more than one event that are mapped to the same preferred term will be counted only once under the strongest severity and relationship, accordingly. Chi-Square or Fisher's exact test will be used to compare the proportions of subjects of the two active treatment groups who report any TEAE.

Treatment-Emergent Serious Adverse Events (TESAEs) will be discussed within the clinical study report. TEAEs, TESAEs and TEAEs that led to treatment interruption or discontinuation will be presented in data listings.

Concomitant Medications, Laboratory Values, and Vitals Signs

Concomitant medications will be coded using the WHO Drug Dictionary, version September 2015, and will be presented in data listings. All laboratory data and vital signs data will be displayed in listings.

Safety comparisons will be performed only for the ITT population.

6 Appendices

6.1 Handling of Missing or Incomplete Dates for Adverse Events and Concomitant Medications

Adverse Events

Handling of partial dates is only considered for the start date. An adverse event with a partial start date is considered treatment emergent if:

- only the day is missing and the start month/year is the same or after the month/year of the first dose

- the day and month are missing and the start year is the same or greater than the year of the first dose date

- the start date is completely missing

Concomitant Medications

Handling of partial dates is only considered for the stop date. A medication with a partial stop date is considered concomitant if:

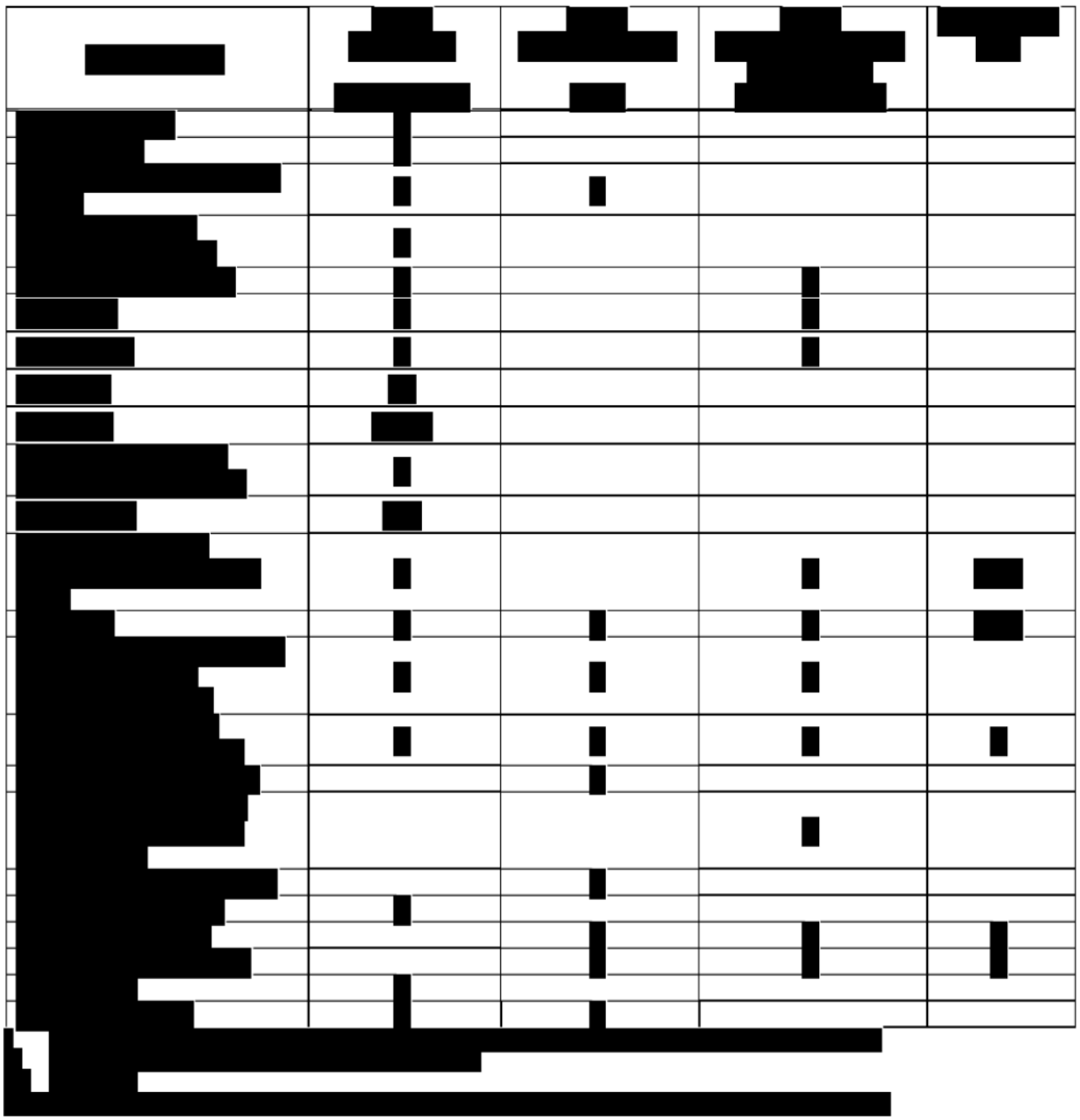
- only the day is missing and the stop month/year is the same or after the month/year of the first dose

- the day and month are missing and the stop year is the same or greater than the year of the first dose date

- the stop date is completely missing or the medication is ongoing

6.2 Summary of Assessments

The schedule of visits and procedures to be conducted at each visit are summarized in the Schedule of Study Procedures.



7 [redacted]