STATISTICAL ANALYSIS PLAN

ITF Research Pharma SLU
c/ San Rafael 3
28108 Madrid, Spain

A PHASE 2, DOSE-RANGING, 12-WEEK RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, PARALLEL-GROUP STUDY EVALUATING THE EFFICACY AND SAFETY OF THREE FORMULATIONS OF ULTRA-LOW DOSE ESTRIOL VAGINAL GEL (0.005% ESTRIOL VAGINAL GEL, 0.002% ESTRIOL VAGINAL GEL, 0.0008% ESTRIOL VAGINAL GEL) FOR THE TREATMENT OF VAGINAL DRYNESS IN POSTMENOPAUSAL WOMEN WITH VULVOVAGINAL ATROPHY

Protocol No: ITFE-2092-C1
Protocol v5.0 dated 07Jun2017
IND No: 104785 / EUDRACT No: 2015-005787-42

Prepared by:

STATISTICAL ANALYSIS PLAN

SIGNATURE PAGE

ITF Research Pharma SLU
c/ San Rafael 3
28108 Madrid, Spain

A PHASE 2, DOSE-RANGING, 12-WEEK RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, PARALLEL-GROUP STUDY EVALUATING THE EFFICACY AND SAFETY OF THREE FORMULATIONS OF ULTRA-LOW DOSE ESTRIOL VAGINAL GEL (0.005% ESTRIOL VAGINAL GEL, 0.002% ESTRIOL VAGINAL GEL, 0.0008% ESTRIOL VAGINAL GEL) FOR THE TREATMENT OF VAGINAL DRYNESS IN POSTMENOPAUSAL WOMEN WITH VULVOVAGINAL ATROPHY

Protocol No: ITFE-2092-C1
Protocol v5.0 dated 07Jun2017

[redacted]

STATISTICAL ANALYSIS PLAN
REVISION HISTORY

Version	Date	Summary of Revisions	Revised By
Final 1.0	22Jun2017	NA – Original version	

TABLE OF CONTENTS

LIST OF STATISTICAL ANALYSIS PLAN IN-TEXT TABLES

LIST OF ABBREVIATIONS AND DEFINITION OF TERMS

Abbreviation	Definition
AE	adverse event
ANCOVA	analysis of covariance
ATC	Anatomical Therapeutic Chemical
BMI	body mass index
cm	Centimeter
CI	Confidence Interval
CRA	Clinical Research Associate
CSR	Clinical Study Report
E1	estrone
E2	17β-estradiol
E3	estriol
ECG	electrocardiogram
eCRF	electronic case report form
FDA	Food and Drug Administration
FSH	Follicle Stimulating Hormone
GCP	Good Clinical Practice
HEENT	Head, Ears, Eyes, Nose and Throat
ICF	Informed Consent Form
ICH	International Conference on Harmonisation
ITT	Intent-to-Treat
IWRS	Interactive Web Response System
kg	Kilogram
LH	Luteinizing Hormone
LOCF	Last observation carried forward
LS Mean	Least Squares Mean
MedDRA	Medical Dictionary for Regulatory Activities
mcg/g	Microgram/gram
mg	Milligram
mmHg	Millimeters of mercury
MMRM	mixed effects model for repeated measures
msec	Millisecond
PD	Protocol Deviation
PP	Per Protocol
QTcB	QT interval, corrected according to Bazett
QTcF	QT interval, corrected according to Fridericia
SAE	serious adverse event
SAP	statistical analysis plan
SD	standard deviation
SE	standard error
SOC	system organ class
US	United States
WHO DDE	World Health Organization Drug Dictionary Enhanced

1. INTRODUCTION

Vulvovaginal atrophy is a natural consequence of the progressive estrogen deficiency that occurs in menopause. Epidemiological data have indicated that about 50% of otherwise healthy women over 60 years of age experience symptoms related to urogenital atrophy such as vaginal dryness, dyspareunia, burning, itching, as well as urinary complaints or infections of the lower urinary tract. As these alterations frequently affect the quality of life of postmenopausal women, it is important for doctors to detect their presence and offer treatment options.

Estrogen therapy is the most effective treatment of moderate to severe symptoms of vulvar and vaginal atrophy. One advantage of local treatment with estrogen is avoidance of first-pass liver metabolism, making it possible to use lower doses of estrogen compared with oral therapy; the local route also minimizes systemic adverse effects.

The primary risks of taking an estrogen product after menopause is that it can cause or exacerbate certain types of tumors, such as breast or endometrium. Taking low dose estrogens vaginally decreases the risk of the occurrence of these tumors, as a consequence of the very low doses employed that permit a local vaginal effect while the potential of absorption is minimized. It has been shown that lower doses of vaginal estrogen therapy than previously used, even with less frequent administration, often yield satisfactory results. The lowest effective dose of estrogen consistent with treatment goals, benefits, and risks for the individual woman should be the therapeutic goal (The 2012 Hormone Therapy Position Statement of the North American Menopause Society, 2012).

Estriol (E3) is one of the most important estrogens produced in the human body, along with 17β-estradiol (E2) and estrone (E1). Estriol is normally produced in the liver by irreversible transformation of estradiol and estrone. This process is considered to be a metabolic detoxification mechanism, by which potent estrogens are transformed into less active metabolites such as estriol. Also, estriol is produced in significant amounts during pregnancy as it is synthesized by the placenta. In all instances, there is no conversion from estriol to estradiol or estrone.

Estriol formulations have been marketed since 1958 for the treatment of vulvovaginal atrophy in more than 50 countries worldwide, including all European countries, but not the United States. Topical (vaginal) formulations are currently available as creams (containing 1 mg estriol/1 g cream) or as ovules (containing 0.5 mg estriol). Local vaginal estrogen therapy can also be administered in the form of estradiol vaginal tablets or vaginal creams, estradiol-releasing vaginal rings and conjugated estrogen-based vaginal creams.

The search for therapeutic alternatives, which may present improvements in relation to the current products, has been encouraged. In response to this need, ITF Research Pharma SLU has developed a 0.005% estriol vaginal gel; this estriol-based formulation for intravaginal application uses a dose of the hormone ten times lower than that included in currently marketed vaginal estriol products. It proved to be an effective treatment for postmenopausal vaginal atrophy, as the majority of symptoms and signs of postmenopausal vaginal atrophy were improved after

treatment, with improvement evidenced in as little as 3 weeks after treatment initiation. A favorable safety profile was demonstrated, as systemic exposure to estriol was shown to be extremely low after repeated administration and no differences in the adverse event (AE) profile (type and severity) were demonstrated when compared to placebo treatment (Cano A, 2012). While maintaining efficacy, the reduced estrogen exposure results in a more favorable risk/benefit ratio for the patient. This new formulation consists of a muco-adhesive gel with pleasant pharmaceutical properties (freshness, cleanliness, odorless) and easy application. These aspects are considered relevant as specialties in the market for the treatment of vaginal atrophy are often perceived as unpleasant to use by some patients and may result in discontinuation of treatment. The gel also has a hydrating effect, which may additionally contribute to the relief of some of the characteristic symptoms of atrophy.

[redacted]

2. STUDY OBJECTIVES

2.1 Primary Objective

The primary objective of this study is to evaluate the efficacy of 0.005%, 0.002%, and 0.0008% estriol vaginal gel and to determine the minimal effective dose for the treatment of postmenopausal vaginal atrophy in women that refer moderate to severe vaginal dryness as the most bothersome symptom.

2.2 Secondary Objectives

The secondary objectives of this study are to:

- Evaluate the efficacy of the three formulations of estriol vaginal gel in the improvement of other symptoms and signs of vulvovaginal atrophy.
- Evaluate the safety and tolerability of the three formulations of estriol vaginal gel.

2.3 Exploratory Objectives

The exploratory objectives of this study are to:

- Evaluate the final subjective global perception of efficacy of the three formulations of estriol vaginal gel.
- Evaluate the subject's acceptability of the three formulations of estriol vaginal gel.

3. STUDY DESIGN

This study is a double-blind, randomized, dose-ranging, placebo-controlled multicenter study with four parallel groups. Approximately 25 sites in five European countries (Czech Republic, Hungary, Italy, Spain and Sweden) will be included in the study. The study setting will include academic hospitals and gynecological clinics. The study population is postmenopausal women who report moderate to severe vaginal dryness as their most bothersome symptom in addition to other symptoms and/or signs of postmenopausal vulvovaginal atrophy.

After providing informed consent, potential subjects will be evaluated during the screening period (maximum 30 days) to determine if they are eligible to participate in the study. A screening identification number will be assigned to each eligible subject according to their order of arrival into the study at the Screening visit. Subjects who meet all of the inclusion criteria and none of the exclusion criteria (see Sections 4.1.1 and 4.1.2 of the study protocol) will be randomized into the study at the Baseline visit. A total of 280 subjects will be enrolled and randomly allocated to one of the four treatment groups (0.005%, 0.002%, or 0.0008% estriol vaginal gel or placebo vaginal gel); the allocation to treatment group will be 1:1:1:1. A blocked randomization across the entire study (i.e. all sites as a whole) will be used. Subjects will be assigned a randomization code according to their order of arrival into the study at the Baseline Visit. Initial treatment consists of daily vaginal administration of 1 gram of randomly-assigned estriol gel or placebo gel using the supplied applicator. Placebo gel will be identical in appearance, texture and odor to the estriol vaginal gel products. After 21 days of daily administration, treatment will continue with twice-weekly administration up to Week 12 (i.e., Mondays and Thursdays).

An Interactive Web Response System (IWRS) will be used in the study to assign subjects a subject number, a randomization code, to screen fail a subject, to assign drug (following randomization) at the Baseline, Week 3 and Week 8 visits, to confirm a subject has completed the study and to withdraw/discontinue a subject from the study. Subjects will be assigned a subject number during the Screening visit after the subject has provided a signed Informed Consent. The subject number will remain the same for a subject throughout the duration of the study. Subject numbers will be different from the actual randomization codes. Screen failures (subjects not randomized into the study) may be re-screened in cases where the unmet selection criteria have been resolved. A new signed Informed Consent will be needed for re-screened subjects. The subject number assigned at screening will be modified to include the number of re-screenings completed for the individual. For example, if subject 8880001 is re-screened for the first time,

they will be assigned the subject number 8881001; if re-screened for a second time, they will be assigned subject number 8882001. Only invasive procedures (endometrial biopsy and Factor V Leiden) will not be repeated if subjects are re-screened within one month the previous failure. Randomized subjects prematurely withdrawn from the study will not be replaced.

The randomization list will be kept strictly confidential, accessible only to authorized personnel until the time of unblinding. The randomization codes will be unblinded and made available for the project statistician to perform the data analysis when the study has been finalized, the database has been verified, and all protocol violations have been determined. All personnel involved directly in the study shall be unaware of the treatment assigned to the subjects until the database has been closed.

The duration of trial participation for each subject is expected to be approximately 20 weeks (from Screening to End-of-Study phone call) although the time on study treatment will be approximately 12 weeks. Subject completion is defined as the completion of all phases of the study, including the End-of-Study phone call to be performed for safety purposes at Week 16.

An overview of the study visits and procedures is given in the Schedule of Events (**Table 1 in Section 4**). Details of specific procedures are provided in **Section 4**.

4. SCHEDULE OF ASSESSMENTS

4.1 Study Visits

4.1.1 Screening Visit

Screening is defined as the period prior to the administration of study treatment. In this study, the Screening period will cover a maximum of 30 days before randomization on the day of the Baseline visit.

A subject will enter the Screening period if basic study criteria are matched with the potential participant's medical condition and profile. In case of agreement to participate, the subject will be requested to sign an Informed Consent Form (ICF).

During the Screening visit, the following procedures will be conducted for each subject. Further details are provided in Section 6.1.1 of the study protocol.

- Register the subject in IWRS to obtain unique subject identifier.
- Record demographic details: Year of birth, age and race.
- Record recent (within the past year) and relevant gynecological, medical and surgical history and any chronic or ongoing medical conditions for which the subject received or is receiving medical treatment.

- Record use of prohibited medications during the periods defined in the exclusion criteria (see Appendices 12.3 and 12.4 of the study protocol).

- Record prior (including the 4 weeks prior to study entry) and concomitant medications taken or being taken.

- Perform physical examination: General appearance; Head, eyes, ears, nose and throat (HEENT); Cardiovascular; Respiratory; Gastrointestinal; Dermatological; Neurological; Musculoskeletal; Lymphatic and any 'Other' body systems, noting any abnormal findings.

- Measurement of vital signs:
  - Height and weight
  - Seated for at least 5 minutes: blood pressure and heart rate

- Perform two 12-lead electrocardiogram (ECG) examinations, approximately 5 minutes apart from one another.

- Perform gynecological examination:
  - Visual inspection of external genitalia
  - Speculum examination of vagina
  - Speculum examination of cervix (if present)
  - Bimanual palpation of uterus and ovaries (if present)
  - Breast palpation
  - Collection of papanicolau smear

- Perform mammogram, if documented negative mammogram within 9 months prior to randomization is not available.

- Evaluation of presence and intensity of the symptom vaginal dryness.

- Collection of vaginal smear for evaluation of vaginal cytology: 2 smears to be taken, from the bottom of the right and left vaginal walls respectively.

- Collection of vaginal pH

- Perform transvaginal ultrasound (in subjects with a uterus): Endometrial thickness; Pelvic pathology.

- Perform endometrial biopsy (in subjects with a uterus).

- Collection of safety laboratory parameters: blood chemistry and hematology, including coagulation and lipids, urinalysis and serum hormone levels. These samples should be

performed under fasting conditions. A Factor V Leiden sample will also be taken at Screening if the requisite informed consent is obtained.

- Record AEs.

Before the Baseline visit, the investigator will confirm that the subject meets all the inclusion and exclusion criteria, including the endometrial histological assessment.

Subjects with signs and symptoms suggestive of infection of the genital or urinary tract requiring treatment at the start of the study will be excluded. However, once treated, the symptoms have resolved (or there are negative microbiology findings) and the required wash out period has passed according to the treatment used, subjects can be re-screened into the study. Refer to Section 4.2.3 of the study protocol for more details.

4.1.2 Baseline Visit

The Baseline visit will be performed on Day 1. The following procedures will be conducted for each subject. Further details are provided in Section 6.1.2 of the study protocol.

- Randomization: IWRS will be used to allocate the subject's randomization number and 3 study medication kit numbers.
- Evaluation of presence and intensity of the symptoms (other than vaginal dryness) and signs of vulvovaginal atrophy.
- Record AEs since the Screening visit.
- Record concomitant medications.
- Administration of the first dose of study treatment, either by the investigator or by the subject in the presence of the investigator.
- Dispensing of dosing instructions and adequate medication supplies for 3 weeks: three medication kits each containing one 10g tube of gel and 10 single-use applicators, to cover 21 expected daily applications between the Baseline visit and the Week 3 visit.
- Perform endometrial biopsy (only if the subject presents with vaginal bleeding).

4.1.3 Week 3 Visit

The Week 3 visit will be scheduled on the day after having completed 21 days of study treatment. For organizational reasons, a deviation of 3 days before or after the scheduled date is allowed.

The following procedures will be conducted for each subject. Further details are provided in Section 6.1.3 of the study protocol.

- Evaluation of presence and intensity of symptoms (including vaginal dryness) and signs of vulvovaginal atrophy.
- Collection of vaginal pH
- Collection of vaginal smear for evaluation of vaginal cytology: two smears to be taken, from the bottom of the right and left vaginal walls respectively.
- Record AEs.
- Record concomitant medications.
- Reconciliation of returned medication.
- Dispensing of adequate medication supplies: two medication kits, to cover 10 expected twice-weekly applications between the Week 3 visit and the Week 8 visit. IWRS will be used to allocate two study medication kit numbers.
- Perform endometrial biopsy (only if the subject presents with vaginal bleeding).

4.1.4 Week 8 Visit

The Week 8 visit will be scheduled on the day after having completed 8 weeks of study participation. For organizational reasons, a deviation of 5 days before or after the scheduled visit date is allowed. The following procedures will be conducted for each subject. Further details are provided in Section 6.1.4 of the study protocol.

- Record AEs.
- Record concomitant medications.
- Reconciliation of returned medication.
- Dispensing of adequate medication supplies: two medication kits, to cover eight expected twice-weekly applications between the Week 8 visit and the Week 12 visit. IWRS will be used to allocate 2 study medication kit numbers.
- Perform endometrial biopsy (only if the subject presents with vaginal bleeding).

4.1.5 Week 12 / Early Termination Visit

The Week 12 visit will be scheduled on the day after having completed 12 weeks of study participation, (or earlier if the subject terminates study participation before Week 12, in which case this status should be registered in IWRS). For organizational reasons, a deviation of 5 days before or after the estimated ideal date is allowed. The following procedures will be conducted for each subject. Further details are provided in Section 6.1.5 of the study protocol.

- Evaluation of presence and intensity of symptoms (including vaginal dryness) and signs of vulvovaginal atrophy.
- Collection of vaginal pH
- Collection of vaginal smear for evaluation of vaginal cytology: two smears to be taken, from the bottom of the right and left vaginal walls respectively.
- Evaluation of global efficacy.
- Evaluation of acceptability.
- Perform transvaginal ultrasound (in subjects with a uterus): Endometrial thickness; Pelvic pathology.
- Perform endometrial biopsy (in subjects with a uterus).
- Collection of safety laboratory parameters: blood chemistry and hematology, including coagulation and lipids, urinalysis and serum hormone levels. These samples should be performed under fasting conditions. A Factor V Leiden sample will also be taken at the Week 12 / Early Termination visit if not already obtained at Screening.
- Record AEs.
- Measurement of vital signs:
  - Weight
  - Seated for at least 5 minutes: blood pressure and heart rate
- Physical examination: General appearance; Head, eyes, ears, nose and throat; Cardiovascular; Respiratory; Gastrointestinal; Dermatological; Neurological; Musculoskeletal; Lymphatic and any 'Other' body systems, noting any abnormal findings.
- Perform two 12-lead electrocardiogram (ECG) examinations, approximately 5 minutes apart from one another.
- Perform gynecological examination:

- Visual inspection of external genitalia
- Speculum examination of vagina
- Speculum examination of cervix (if present)
- Bimanual palpation of uterus and ovaries (if present)
- Breast palpation

- Record concomitant medications.

- Reconciliation of returned medication.

4.1.6 Week 16 – End of Study Phone Call

All subjects who complete the full 12 weeks of study treatment will have a phone call scheduled for 4 weeks after the last dose of study treatment to record any AEs that may have appeared after treatment discontinuation, and to document changes in concomitant medication use. For organizational reasons, a deviation of 5 days before or after the estimated ideal date is allowed.

When a subject has completed the study, the completed status should be registered in IWRS.

An overview of the study visits and procedures is given in the Schedule of Events (**Table 1**).

Table 1: Schedule of Events

Procedure	Screening Visit	Baseline	Week 3	Week 8	Week 12 / Early Termination	Week 16 / End of Study Phone Call
Visit Window (days)	**-30 to -1**	**1**	**21 ± 3**	**56 ± 5**	**84 ± 5**	**112 ± 5**
Informed consent	X					
Randomization		X				
Medical history, medication history and demographics	X					
Physical examination [a]	X				X	
12-lead electrocardiogram [b]	X				X	
Gynecological examination [c]	X				X	
Transvaginal ultrasound (women who have a uterus)	X				X	
Mammogram (if no negative mammogram documented within 9 months prior to randomization) [d]						
Clinical laboratory tests [e]	X				X	
Endometrial biopsy (women who have a uterus)	X	*	*	*	X	
Adverse event reporting [f]	X	X	X	X	X	X
Concomitant medication reporting	X	X	X	X	X	X
Evaluation of vaginal dryness [g]	X		X		X	
Evaluation of symptoms and signs of vulvovaginal atrophy [g, h]		X	X		X	
Vaginal pH	X		X		X	
Vaginal smear for evaluation of vaginal cytology	X		X		X	
Evaluation of global efficacy [i]					X	
Evaluation of acceptability [j]					X	
Administration of the first dose of the study medication [k]		X				
Study medication supply and dispensing [l]		X	X	X		
Reconciliation of returned study medication			X	X	X	

a Physical examination to include vital signs measurements (height [Screening visit only], weight, blood pressure and heart rate).
b Two ECG exams will be performed in each of these two visits, approximately 5 minutes apart from one another.
c Gynecological examination to include visual inspection of external genitalia, speculum examination of vagina and cervix (if present), bimanual palpation of uterus and ovaries (if present) and breast palpation. Papanicolau examination will be performed at the Screening visit only.

Procedure	Screening Visit	Baseline	Week 3	Week 8	Week 12 / Early Termination	Week 16 / End of Study Phone Call
Visit Window (days)	**-30 to -1**	**1**	**21 ± 3**	**56 ± 5**	**84 ± 5**	**112 ± 5**

d Neither mammograms performed as study specific procedure nor negative mammogram documented within 9 months prior to randomization will be recorded in the eCRF but archived in subject's clinical records.

e Laboratory testing performed under fasting conditions at Screening and Week 12 includes: standard hematology panel plus prothrombin time, activated thromboplastin partial time, fibrinogen, antithrombin III/protein C/protein S; standard chemistry panel (glucose, cholesterol, high-density lipoprotein, low-density lipoprotein, triglycerides albumin, alanine aminotransferase, aspartate aminotransferase, creatine phosphokinase, bilirubin, creatinine, total protein, uric acid, blood urea nitrogen); hormones (estradiol, estriol, estrone, follicle-stimulating hormone, luteinizing hormone); and standard urinalysis with microbiology. F actor V Leiden testing will be performed only at the Screening visit (or at the Week 12 / Early Termination visit if not obtained at Screening).

f AEs and SAEs will be recorded from Screening (post ICF signature) and subsequently at every study visit and up to 30 days post last study treatment administration. Tolerability of study treatment will be evaluated at Week 3, Week 8, Week 12 and Week 16.

g For the evaluation of symptoms and signs, each symptom and each sign will be graded as: 0 = Absent; 1 = Mild; 2 = Moderate; 3 = Severe.

h Signs and symptoms other than vaginal dryness to be evaluated are the following:
Symptoms: dyspareunia; pruritus or itching; burning; dysuria.
Signs: pallor; friability; thinning or flattening of folds; presence of petechiae; dry mucosa.

i Global efficacy will be assessed as: highly efficacious; efficacious; non-efficacious; detrimental.

j Global acceptability will be assessed as: excellent; good; acceptable; unacceptable.

k The investigator will administer the first study medication or will supervise the subject doing the administration.

l Medication kits will be allocated via IWRS.

* Endometrial biopsy is mandatory in subjects with a uterus at the Screening and Week 12 visits, and whenever the subject presents with vaginal bleeding at other visits.

5. ANALYSIS POPULATIONS

A total of three populations will be used for all summaries and analyses. Subjects who have satisfied the population criteria will be classified in the designated population and will only be included in analyses for which they have available data.

5.1 Intent-to-Treat (ITT) Population

The ITT population is defined as all randomized subjects. The ITT population will be used to present all efficacy analyses, including the four co-primary efficacy endpoints. Subjects will be summarized according to the treatment to which they were randomized.

5.2 Per Protocol (PP) Population

The PP population is defined as all randomized subjects who complete the study to Week 12 with no major protocol deviations impacting the co-primary efficacy analyses. The major deviations impacting the analysis are defined in more detail in **Section 6.2.2**. The PP population will be used for confirmation of efficacy analyses of interest. Subjects will be summarized according to the treatment to which they were randomized.

5.3 Safety Population

The Safety population includes all randomized subjects who administered at least one dose of study treatment. The Safety population will be used to present all safety analyses. Subjects will be summarized according to the actual treatment received. In the event that a subject is administered incorrect treatment in error, the subject will be categorized according to their randomized treatment assignment if at least one dose of the correct treatment was administered during the subject's time on treatment.

6. STATISTICAL METHODOLOGY

6.1 Statistical and Analytical Issues

6.1.1 Statistical Methods

All statistical methods will be based on the International Conference on Harmonization (ICH) E9 document "Statistical Principles for Clinical Trials".

All data will be summarized by treatment group. In addition, where appropriate, a total overall group column will be included to summarize all subjects. Where appropriate, data will be summarized by treatment group and visit.

In summary and analysis tables of continuous variables, the minimum and maximum statistics will be presented to the same number of decimal places as the original data. The mean, median, quartiles, Least Squares mean (LS mean), and 95% Confidence Interval (CI) will be presented to one more decimal place than the original data. The standard deviation (SD) and standard error (SE) will be presented to two more decimal places than the original data. Standard descriptive statistics (number of subjects with data available [n], mean, SD, median, quartiles, minimum and maximum) will be presented in summary tables. The LS mean, SE and 95% CI will be presented in the statistical analysis outputs as appropriate.

In summary tables of categorical variables, the number of non-missing observations and percentages will be presented. The denominator for each percentage will be the number of subjects within the population and treatment group of interest with non-missing responses, unless otherwise specified.

All hypothesis testing will be 1-sided unless otherwise specified and carried out at the 2.5% significance level, designed to evaluate the superiority of estriol vaginal gel versus placebo vaginal gel.

P-values will be rounded to 3 decimal places. P-values less than 0.001 will be reported in the tables as '<0.001'. P-values greater than 0.999 will be reported as '>0.999'.

Should any of the statistical methods proposed prove unsuitable during the final analysis, more appropriate methods will be used and any changes will be documented in the Clinical Study Report (CSR), including the rationale for use.

All data collected on the electronic case report form (eCRF) as well as the data from central laboratories will be presented in data listings. The subject data listings will be sorted by unique subject number, randomized treatment group and visit or visit date, as appropriate. Further sorting variables may be used if required.

The Baseline assessment is defined as the last assessment prior to first administration of study medication. This will be the last assessment from either the Screening or Baseline visits.

All statistical analysis will be performed using SAS® Version 9.3 or higher.

The statistical analysis will be conducted when all data for all subjects have been entered, cleaned and locked, at which point the study will be formally unblinded.

6.1.2 Handling of Dropouts and Missing Data

For the primary efficacy analysis of all four co-primary efficacy endpoints, and all secondary efficacy analyses of change from Baseline to Week 12 in symptoms and signs of vulvovaginal atrophy and the global symptom score, missing data will be imputed using the last observation carried forward (LOCF) methodology. Only post-Baseline values will be carried forward.

As sensitivity analyses, the four co-primary efficacy variables will also be analyzed on an as observed case basis with no imputation of missing data, in order to assess the robustness of the results.

All other analyses will be conducted on an as observed case basis, that is, no further imputation of missing data will be carried out unless expressly stated otherwise.

6.1.3 Pooling of Investigative Sites

Sites will be pooled according to country for inclusion as an effect in the statistical models. It is recognized that the pooling of sites in this manner preserves the inclusion of all randomized subjects in the analysis. However, the result of pooling is a loss of the "individuality" of each site being pooled as well as the potential for the pooled site to be associated with an inflated sample size compared to sites that do not require pooling. The effect of this cannot be determined a priori but it is an acceptable approach for the retention of subjects in the analysis.

6.1.4 Determination of Sample Size

A total of 280 female subjects will be enrolled in the study.

The estimates of treatment differences and variability were gained from the results of Study ITFE-2026-C2.

A sample size of 57 in each group will have 90% power to detect a treatment difference for change in severity of vaginal dryness, assuming five categories with proportions in each treatment group as shown in **Table 2** below, using a Wilcoxon rank sum test with a 0.025 one-sided significance level.

Table 2: Assumed Proportions for Wilcoxon Rank Sum Test of Vaginal Dryness

Change in Vaginal Dryness Severity	Proportion in Estriol Vaginal Gel Group	Proportion in Placebo Group
-3	0.189	0.118
-2	0.423	0.196
-1	0.270	0.353
0	0.118	0.333

Change in Vaginal Dryness Severity	Proportion in Estriol Vaginal Gel Group	Proportion in Placebo Group
1	0	0

A sample size of 48 in each group will have 90% power to detect a treatment difference of -0.89 for change in vaginal pH, assuming a common standard deviation of 1.33 using a two-group t-test with a 0.025 one-sided significance level.

A sample size of 24 in each group will have 90% power to detect a treatment difference of 21 for change in proportion of superficial cells, assuming a standard deviation of 26.3 for 0.005% estriol vaginal gel and 14.8 for placebo vaginal gel, using a two-group Satterthwaite t-test with a 0.025 one-sided significance level.

A sample size of 9 in each group will have 90% power to detect a treatment difference of -53 for change in proportion of parabasal cells, assuming a standard deviation of 37.9 for 0.005% estriol vaginal gel and 19.6 for placebo vaginal gel, using a two-group Satterthwaite t-test with a 0.025 one sided significance level.

The largest sample size calculated for a single co-primary endpoint has been considered the sample size required for the whole study, thus, 57 subjects per group are estimated. Accounting for an 18% drop-out rate, and rounding numbers, approximately 70 subjects/group have been finally estimated.

6.2 Subject Characteristics

6.2.1 Subject Disposition

The subject disposition table will summarize the following and will be presented for all subjects by treatment group and overall:

- The number of subjects enrolled
- The number of subjects screen failed
- The number of subjects re-screened
- The number of subjects randomized
- The number and percentage of randomized subjects discontinued from the study
- The number and percentage of randomized subjects who completed the study
- The number and percentage of randomized subjects in the ITT population
- The number and percentage of randomized subjects in the PP population
- The number and percentage of randomized subjects in the Safety population

The number and percentage of randomized subjects who discontinue from the study and the primary reason for discontinuation will also be summarized by treatment group and overall.

The number and percentage of enrolled subjects meeting each inclusion/exclusion criteria will be summarized overall.

6.2.2 Protocol Deviations

A protocol deviation (PD) is defined as a divergence or departure from the requirements and procedures described in the study protocol and protocol amendments. The study PD Plan describes the procedures for reporting PDs from all sites into a global PD Log and reconciliation of this PD Log with the clinical database. All PDs in the global PD Log will be classified by the Medical Monitor into one of the categories shown in **Table 3** below.

Table 3: Classification of Protocol Deviations

1	Informed Consent Form Non-Compliance
2	Inclusion/Exclusion Non-Compliance
3	Investigational Product Non-Compliance
4	Prohibited Concomitant Treatment
5	Protocol Procedure Non-Compliance – Safety
6	Protocol Procedure Non-Compliance – Efficacy
7	Protocol Procedure Non-Compliance - Safety AND Efficacy
8	Protocol Procedure Performed Out of Visit Window
9	Randomization Procedure not Followed
10	Other

Appendix 3 of the PD Plan defines how each PD that may occur during the study should be classified.

The number and percentage of subjects with any PD and the number and percentage of subjects with each category of PD will be summarized by treatment group and overall for the ITT population.

Protocol Deviations from the global PD Log that may impact the co-primary efficacy analyses are those classified as "Inclusion/Exclusion Non-Compliance", "Investigational Product Non-Compliance", "Prohibited Concomitant Treatment", "Protocol Procedure Non-Compliance – Efficacy", "Protocol Procedure Non-Compliance – Safety AND Efficacy", "Protocol Procedure Performed Out of Visit Window", "Randomization Procedure not Followed" or "Other". All PDs within these classifications will be reviewed by the Medical Monitor and the sponsor, and those deviations impacting the co-primary efficacy results will be identified in the global PD Log and reviewed by the study statistician. Subjects with such events will be excluded from the PP population.

The following criteria will be derived programmatically using the data from the clinical database and relevant external vendors. Subjects meeting any of these criteria will also be excluded from the PP population.

- Did not complete the study to Week 12.
- Vaginal dryness assessment not available at Baseline.
- Vaginal dryness assessment not available at Week 12.
- pH assessment not available at Baseline.
- pH assessment not available at Week 12.
- Proportion of superficial cells from vaginal smear not available at Baseline.
- Proportion of superficial cells from vaginal smear not available at Week 12.
- Proportion of parabasal cells from vaginal smear not available at Baseline.
- Proportion of parabasal cells from vaginal smear not available at Week 12.
- Received study treatment that was not the allocated treatment for that subject.

Subjects to be excluded from the PP population will be identified prior to database lock and before the treatment codes are unblinded. In the case of administration of incorrect medication kits, it will not be possible to identify whether any subject received study treatment other than their allocated treatment until unblinding is performed. Any subject known to have received incorrect kits will be reviewed post unblinding and excluded from the PP population if necessary.

The number and percentage of subjects meeting any criteria for exclusion from the PP population and the number and percentage of subjects meeting each criterion will be summarized by treatment group and overall for the ITT population.

6.2.3 Background and Demographic Characteristics

The following demographic characteristics will be summarized by treatment group and overall for the ITT population:

- Age (years) recorded directly on the eCRF
- Race (American Indian or Alaska Native; Asian; Native Hawaiian or Other Pacific Islander; White; Black or African American; Other)

Demographic data will be summarized using summary statistics for continuous variables (number of subjects [n], mean, standard deviation [SD], median, quartiles, minimum and maximum) or by way of group frequencies and percentages for categorical variables, as appropriate.

6.2.4 Treatment Exposure and Compliance

Study treatment exposure and compliance will be summarized by treatment group and overall for the Safety population.

Study Treatment Data

Subjects are instructed to bring their used and unused study treatment and unused applicators to the Week 3, Week 8 and Week 12 visits. All used applicators will be discarded by the subjects in their normal household waste. Compliance will be evaluated at the site, by reviewing the study treatment and the returned supplies. The following data related to administration of study treatment will be collected:

[A] The date of first administration of study treatment will be captured on the 'Study Drug Administration – Baseline' page of the eCRF.

[B] Ongoing study treatment data will be captured at the Week 3, Week 8 and Week 12 visits on the 'Study Drug Administration and Compliance – Cumulative' page of the eCRF. This log page includes, for each kit number: the date dispensed, the date returned, whether the tube of gel was returned, the number of applicators returned, the number of applicators used and the number of applicators lost/missed/damaged.

[C] When the subject completes or discontinues from the study, the date of last administration of study treatment will be captured on the 'End of Study' page of the eCRF.

Exposure

Extent of exposure to study treatment is defined as:

$$\textit{Extent of Exposure (days)} = \textit{Date of Last Administration of Study Treatment}^{[C]} - \textit{Date of First Administration of Study Treatment}^{[A]} + 1$$

The date of last administration of study treatment will be taken from the 'End of Study' page of the eCRF wherever available. If this is missing or not known, the date of last administration of study treatment will be assumed to be the latest date that a kit is returned on the 'Study Drug Administration and Compliance – Cumulative' page[B] of the eCRF.

If no kits were dispensed according to the 'Study Drug Administration and Compliance – Cumulative' page[B], but the study treatment was administered at Baseline according to the 'Study Drug Administration – Baseline' page[A], the date of last administration of study treatment will be assumed to be equal to the date of first administration of study treatment, i.e. exposure will be calculated as equal to 1 day.

Extent of exposure (days) will be summarized using summary statistics for continuous variables (number of subjects [n], mean, standard deviation [SD], median, quartiles, minimum and maximum).

Compliance

The first administration of study treatment will be conducted at the Baseline visit, either by the investigator or by the subject in the presence of the investigator. Study treatment will then be administered once daily, at bedtime, until the Week 3 visit, and twice weekly after the Week 3

visit until the Week 12 visit. Therefore, the expected number of total applications for each subject will be calculated in 3 stages, as follows:

(1) The expected number of applications from Baseline to Week 3 will be calculated as:

Expected Applications from Baseline to Week 3 = Date of Week 3 visit – Date of Baseline visit

This calculation assumes no application of study treatment on the day of the Week 3 visit, based on the bedtime dosing regimen.

If a subject discontinues prior to the Week 3 visit, the date of last administration of study treatment from the 'End of Study' page of the eCRF will be used instead, and the expected number of applications will be calculated as:

Expected Applications from Baseline to Week 3 = Date of Last Administration of Study Treatment[C] *– Date of Baseline Visit*

(2) The expected number of applications from Week 3 to Week 12 will be calculated using the algorithm below.

(i) The duration, in weeks, between the Week 3 visit and the Week 12 visit will be calculated as:

Weeks from Week 3 to Week 12 = (Date of Week 12 Visit – Date of Week 3 Visit) / 7

(ii) If (i) is an integer, i.e. a number of whole weeks, the expected number of applications from Week 3 to Week 12 will be calculated as:

Expected Applications from Week 3 to Week 12 = (Weeks from Week 3 to Week 12) × 2

(iii) If (i) is not an integer and the decimal part of (i) is < 0.5, the expected number of applications from Week 3 to Week 12 will be calculated as:

Expected Applications from Week 3 to Week 12 = [(Weeks from Week 3 to Week 12) × 2] + 1

(iv) If (i) is not an integer and the decimal part of (i) is ≥ 0.5, the expected number of applications from Week 3 to Week 12 will be calculated as:

Expected Applications from Week 3 to Week 12 = [(Weeks from Week 3 to Week 12) × 2] + 2

(3) The expected number of total applications from Baseline to Week 12 will be calculated as:

Expected Applications from Baseline to Week 12 = Expected Applications from Baseline to Week 3 + Expected Applications from Week 3 to Week 12

The actual number of total applications for each subject will be calculated as the total of all applicators used from all kits according to the 'Study Drug Administration and Compliance – Cumulative' page[B] of the eCRF.

The percentage compliance for each subject will then be calculated as:

$$\frac{\text{Actual Number of Total Applications}}{\text{Expected Applications from Baseline to Week 12}} \times 100$$

Treatment compliance (%) will be summarized using summary statistics for continuous variables (number of subjects [n], mean, standard deviation [SD], median, quartiles, minimum and maximum). The number and percentage of subjects with compliance by category (< 80%, 80% - 120%, >120%) will also be presented. The denominator of percentage will be the number of subjects who received study treatment and have a calculated compliance value available.

6.2.5 Prior and Concomitant Medications and Therapies

Prior and Concomitant Medications

The minimum requirement is that the drug name and the dates of administration are to be recorded in the eCRF. This will include all prescription drugs, herbal products, and over-the-counter (OTC) medications from 28 days prior to the start of study treatment and up to 30 days (± 5 days) following the last dose of investigational product; the reason for the drug administration must be recorded in the eCRF. Any changes in concomitant medications also will be recorded in the subject's eCRF.

All prior and concomitant medications will be classified using the World Health Organization Drug Dictionary Enhanced (WHO DDE). The WHO DDE Version June 2016 will be used, with one update during the study as per the Data Management Plan. The Anatomical Therapeutic Chemical (ATC) Classification Level 2 and Preferred Term will be used to list and summarize the data.

Prior Medications are defined as the medications (recorded on the Prior and Concomitant Medications page of the eCRF) that started and stopped before the date of first administration of study treatment captured on the 'Study Drug Administration – Baseline' page of the eCRF. Only medications where the stop date is prior to the date of first administration of study treatment exclusive will be considered *prior medications*. If the stop date is unknown or incomplete and the medications cannot definitely be considered as stopped prior to Baseline then the medications will be considered as *concomitant medications*. Medications for subjects who were not treated with study treatment will be considered as *prior medications*.

Concomitant Medications are defined as the medications (recorded on the Prior and Concomitant Medications page of the eCRF) that either started before the date of first administration of study treatment and continued into the study, or medications that started during the study.

The number and percentage of subjects reporting the use of any prior medications, and the number of reported prior medications by ATC level 2, preferred term, treatment and overall will be summarized for the Safety population. This table will be repeated for concomitant medications.

Other Concomitant Diagnostic or Therapeutic Procedures

All other concomitant diagnostic or therapeutic procedure terms (recorded on the Other Concomitant Diagnostic or Therapeutic Procedures page of the eCRF) will be classified using the Medical Dictionary for Regulatory Activities (MedDRA) coding dictionary. MedDRA Version 19.0 will be used, with one update during the study as per the Data Management Plan.

Other concomitant diagnostic and therapeutic procedures will be summarized by system organ class (SOC), preferred term, treatment and overall for the Safety population. The number and percentage of subjects having each coded term, and the number of events under each term will be presented. If a subject reports multiple diagnostic or therapeutic events with the same preferred term (i.e. with distinct start and stop dates), these shall be summarized once within the count of subjects, but each event will be counted within the count of events.

6.2.6 Medical History

All medical history terms will be classified using the MedDRA coding dictionary. MedDRA Version19.0 will be used, with one update during the study as per the Data Management Plan.

Medical history will be summarized by SOC, preferred term, treatment and overall for the Safety population. The number and percentage of subjects having each medical history term, and the number of events under each term will be presented. If a subject reports multiple medical history events with the same preferred term (i.e. with distinct start and stop dates), these shall be summarized once within the count of subjects, but each event will be counted within the count of events.

6.3 Efficacy Analysis

6.3.1 Primary Efficacy Analysis

The four co-primary efficacy variables in this study are:

- Change from Baseline to Week 12 in the severity of vaginal dryness
- Change from Baseline to Week 12 in vaginal pH
- Change from Baseline to Week 12 in the proportion of superficial cells of the vaginal epithelium

- Change from Baseline to Week 12 in the proportion of parabasal cells of the vaginal epithelium

The planned analyses of each of these four endpoints and the gatekeeping procedure to be employed for obtaining conclusions are described in **Sections 6.3.1.1 to 6.3.1.5** below.

The primary analyses will be performed using the ITT population using LOCF methodology, and confirmatory analyses will be performed using the PP population (using observed data only, since having Week 12 data available is a requirement for inclusion in the PP population). It should be noted that, since the statistical tests are one-sided, only one limit of each confidence interval presented (the limit in the direction where Estriol vaginal gel is favorable for each specific endpoint) is of particular interest for analysis. Each statistical test will examine the null hypothesis that Estriol vaginal gel is no different to Placebo against the alternative hypothesis that Estriol vaginal gel is superior to Placebo.

No adjustment for multiplicity will be made, since statistical significance in all 4 co-primary endpoints is required to declare efficacy. The minimal effective dose will be considered to be the lowest dose at which a significant difference is seen (in favor of Estriol) when compared with Placebo.

6.3.1.1 Change from Baseline to Week 12 in Severity of Vaginal Dryness

Vaginal dryness is one of the five symptoms of vulvovaginal atrophy, and is self-assessed by the subject at the Screening, Week 3 and Week 12 / Early Termination visits.

Each symptom will be scored by the subject using a numeric scale from 0 to 3, as follows:

0	Absent:	The symptom is not present
1	Mild:	The symptom is of mild intensity
2	Moderate:	The symptom is of moderate intensity
3	Severe:	The symptom is severe in intensity

The number and percentage of subjects with each severity of vaginal dryness at each visit will be summarized by treatment in a general summary table of symptoms of vulvovaginal atrophy.

Given that the Inclusion Criteria for the study require the subject to have moderate (score of 2) or severe (score of 3) vaginal dryness at Baseline (Screening), the change from Baseline to Week 12 may take the possible values ranging from -3 (Severe at Screening; Absent at Week 12) to 1 (Moderate at Screening; Severe at Week 12).

The number and percentage of subjects with each ordinal change from Baseline to Week 12 value (-3, -2, -1, 0 and 1) will be summarized by treatment group.

The change from Baseline to Week 12 in vaginal dryness will be analyzed using a Wilcoxon Rank Sum test, controlled for country, to obtain the one-sided p-values for each treatment

comparison (0.005% Estriol versus placebo, 0.002% Estriol versus placebo and 0.0008% Estriol versus placebo) by means of the FREQ procedure in SAS with the SCORES = RANK option. The Hodges-Lehmann estimate of median difference between treatment groups will also be presented for each of the three treatment comparisons above, along with 95% CIs, by means of the NPAR1WAY procedure in SAS. Hodges-Lehmann estimates will be presented without the adjustment for country, since SAS does not currently produce Hodges-Lehmann estimates when covariates are included. Missing values at Week 12 will be estimated using LOCF methodology for the ITT population analysis only. Only post-Baseline values will be carried forward.

6.3.1.2 Change from Baseline to Week 12 in Severity of Vaginal pH

A vaginal pH determination will be performed at the Screening, Week 3 and Week 12 / Early Termination visits.

The vaginal pH will be summarized using descriptive statistics for continuous variables by treatment and visit. The change from Baseline (Screening) value will also be summarized for all post-Baseline visits.

The mean change in vaginal pH from Baseline to the Week 12 / Early Termination visit will be analyzed using a parametric analysis of covariance (ANCOVA) model. The model will include treatment and country as fixed effects and the Baseline (Screening) pH will be used as a covariate. The adjusted (least-square) mean (LS Mean) change in vaginal pH from Baseline to Week 12 will be presented for each treatment, along with the corresponding SE. The difference in LS Means between treatment groups (0.005% Estriol versus Placebo, 0.002% Estriol versus Placebo and 0.0008% Estriol versus Placebo) will also be presented with 95% CI and one-sided p-value. The p-values for the model parameters of country and Baseline pH will also be presented. Model assumptions will be examined and, if required, a Wilcoxon Rank Sum test as described for vaginal dryness (**Section 6.3.1.1**) will be used. Missing values at Week 12 will be estimated using LOCF methodology for the ITT population analysis only. Only post-Baseline values will be carried forward.

6.3.1.3 Change from Baseline to Week 12 in Proportion of Superficial Cells of the Vaginal Epithelium

Vaginal smears will be performed at the Screening, Week 3 and Week 12 / Early Termination visits. Two vaginal smears will be collected per subject at each visit, with one cytological sample sent to the central laboratory [redacted] for analysis, and one sample retained at site for backup purposes. The samples sent to [redacted] will be evaluated by two different cytopathologists who will be blinded to subject, treatment assignment and to each other's evaluations. The number and percentage of superficial cells will be provided for each pathologist's cytologic evaluation, and the percentages will be converted to proportions for statistical presentation.

The proportion of superficial cells of a given subject and visit to be used for analysis will be derived as the mean result across the two pathologists. The evaluations of both pathologists must be verified by the Clinical Research Associate (CRA) in order to be considered for final analysis.

The proportion of superficial cells will be summarized using descriptive statistics for continuous variables by treatment and visit in a general summary table of vaginal cytology parameters. The change from Baseline (Screening) value will also be summarized for all post-Baseline visits.

The mean change in proportion of superficial cells from Baseline to the Week 12 / Early Termination visit will be analyzed using a parametric analysis of covariance (ANCOVA) model. The model will include treatment and country as fixed effects and the Baseline (Screening) proportion of superficial cells will be used as a covariate. The adjusted (least-square) mean (LS Mean) change in proportion of superficial cells from Baseline to Week 12 will be presented for each treatment, along with the corresponding SE. The difference in LS Means between treatment groups (0.005% Estriol versus Placebo, 0.002% Estriol versus Placebo and 0.0008% Estriol versus Placebo) will also be presented with 95% CI and one-sided p-value. The p-values for the model parameters of country and Baseline proportion of superficial cells will also be presented. Model assumptions will be examined and, if required, a Wilcoxon Rank Sum test as described for vaginal dryness (**Section 6.3.1.1**) will be used. Missing values at Week 12 will be estimated using LOCF methodology for the ITT population analysis only. Only post-Baseline values will be carried forward.

6.3.1.4 Change from Baseline to Week 12 in Proportion of Parabasal Cells of the Vaginal Epithelium

The proportion of parabasal cells is based on the evaluation of vaginal smears by two pathologists at [redacted] and the change from Baseline to Week 12 will be summarized and analyzed exactly as described for superficial cells in **Section 6.3.1.3**.

6.3.1.5 Gatekeeping Procedure

For the co-primary efficacy endpoints, a gatekeeping procedure will be employed. The 0.005% Estriol dose will be compared to placebo, and if this is significant at a one-sided alpha level of 0.025 for all four co-primary endpoints, then the 0.002% Estriol dose versus placebo can be considered. If this is also significant at a one-sided alpha level of 0.025 for all four co-primary endpoints, then the 0.0008% Estriol dose versus placebo can be considered. The approach to be taken is outlined in **Table 4** below.

The one-sided p-values for all four co-primary endpoints will be presented for all three treatment comparisons in an overall gatekeeping table, so that it may be clearly seen whether or not all four co-primary endpoints are statistically significant and therefore be determined for which treatment comparisons efficacy can be formally declared. The overall gatekeeping table will be presented using the ITT population with LOCF methodology for the co-primary efficacy analyses, and also

using observed data only for the PP population and the ITT population in order to confirm the co-primary efficacy results.

Table 4: Gatekeeping Procedure for Efficacy Analyses

Test	Endpoints	All endpoints significant	Result	Next step
0.005% Estriol versus placebo	Vaginal dryness Vaginal pH Superficial cells Parabasal cells	Yes	If treatment difference exists between 0.005% Estriol and placebo, move to next dose	Continue to test 0.002% Estriol versus placebo
	Vaginal dryness Vaginal pH Superficial cells Parabasal cells	No	Not enough evidence to conclude a difference exists between 0.005% Estriol and placebo	Stop testing
0.002% Estriol versus placebo	Vaginal dryness Vaginal pH Superficial cells Parabasal cells	Yes	If treatment difference exists between 0.002% Estriol and placebo, move to next dose	Continue to test 0.0008% Estriol versus placebo
	Vaginal dryness Vaginal pH Superficial cells Parabasal cells	No	Not enough evidence to conclude a difference exists between 0.002% Estriol and placebo	Stop testing
0.0008% Estriol versus placebo	Vaginal dryness Vaginal pH Superficial cells Parabasal cells	Yes	Treatment difference exists between 0.0008% Estriol and placebo	Stop testing
	Vaginal dryness Vaginal pH Superficial cells Parabasal cells	No	Not enough evidence to conclude a difference exists between 0.0008% Estriol and placebo	Stop testing

6.3.1.6 Sensitivity Analysis

Observed Case Analysis

The robustness of the LOCF approach for missing data will be addressed by completing a sensitivity analysis. This will be an observed case analysis of the endpoints described in **Sections 6.3.1.1** to **6.3.1.5**, with no imputation methods for missing data, in order to assess the robustness of the primary efficacy results.

Mixed Model for Repeated Measures Analysis

As additional sensitivity analyses, the endpoints described in **Sections 6.3.1.1** to **6.3.1.5** will each be analyzed using a mixed effects model for repeated measures (MMRM) on the ITT population. The models will include treatment, visit, treatment-by-visit interaction, country and Baseline value as fixed effects, and will use an unstructured (general) residual covariance matrix for repeated records within subjects. For each endpoint, the LS mean change in the endpoint from Baseline will be presented for each treatment by visit and overall, along with the corresponding SE. The difference in LS means (0.005% Estriol versus Placebo, 0.002% Estriol versus Placebo and 0.0008% Estriol versus Placebo) will also be presented with 95% CI and one-sided p-value for each visit and overall. The p-values will also be presented for the model parameters of visit, treatment-by-visit interaction, country and Baseline value.

Only ITT population subjects with both a Baseline and at least one post-Baseline measurement for a given endpoint will be included in the MMRM for that endpoint.

Due to the categorical nature of the vaginal dryness endpoint, the change from Baseline (Screening) vaginal dryness values and the Baseline (Screening) vaginal dryness values will be assigned ranks according to their order within each treatment group, and the MMRM will be conducted on the ranked values.

6.3.2 Secondary Efficacy Analysis

The secondary efficacy analyses will be performed using the ITT population. Secondary efficacy endpoints include the change from Baseline to Week 12 in the severity of individual symptoms (excluding vaginal dryness, i.e. dyspareunia, pruritus or itching, burning, and dysuria) and signs of vulvovaginal atrophy (pallor, friability, thinning or flattening of folds, presence of petechiae, and dry mucosa) and global symptom scores, as well as the change from Baseline to Week 3 in the severity of individual symptoms (as above, include vaginal dryness) and signs and global symptom scores, vaginal pH and proportions of superficial and parabasal cells, respectively. The planned analyses of each of these endpoints are described in **Sections 6.3.2.1 to 6.3.2.26** below.

6.3.2.1 Change from Baseline to Week 12 in Severity of Dyspareunia

Dyspareunia is one of the five symptoms of vulvovaginal atrophy. It is self-assessed by the subject at the Baseline, Week 3 and Week 12 / Early Termination visits, and is scored using a numeric scale from 0 to 3 as described in **Section 6.3.1.1**. Dyspareunia is only applicable in subjects who have experienced sexual activity with penetration since the previous study visit.

The number and percentage of subjects with each severity of dyspareunia at each visit will be summarized by treatment in a general summary table of symptoms of vulvovaginal atrophy.

The change from Baseline to Week 12 in severity of dyspareunia may take the possible values ranging from -3 (Severe at Baseline, Absent at Week 12) to 3 (Absent at Baseline, Severe at Week 12). The change from Baseline in severity of dyspareunia will only be calculated when the

subject has indicated that they have experienced sexual activity with penetration at both the Baseline and the post-Baseline visit. This endpoint will be summarized and analyzed as described for vaginal dryness in **Section 6.3.1.1**. Missing values at Week 12 will be estimated using LOCF methodology. Only post-Baseline values will be carried forward.

6.3.2.2 Change from Baseline to Week 12 in Severity of Pruritus or Itching

Pruritus or itching is one of the five symptoms of vulvovaginal atrophy. It is collected as described for dyspareunia in **Section 6.3.2.1** (but is applicable at all visits irrespective of sexual activity with penetration), and will be summarized and analyzed as described for vaginal dryness in **Section 6.3.1.1**. Missing values at Week 12 will be estimated using LOCF methodology. Only post-Baseline values will be carried forward.

6.3.2.3 Change from Baseline to Week 12 in Severity of Burning

Burning is one of the five symptoms of vulvovaginal atrophy. It is collected as described for dyspareunia in **Section 6.3.2.1** (but is applicable at all visits irrespective of sexual activity with penetration), and will be summarized and analyzed as described for vaginal dryness in **Section 6.3.1.1**. Missing values at Week 12 will be estimated using LOCF methodology. Only post-Baseline values will be carried forward.

6.3.2.4 Change from Baseline to Week 12 in Severity of Dysuria

Dysuria is one of the five symptoms of vulvovaginal atrophy. It is collected as described for dyspareunia in **Section 6.3.2.1** (but is applicable at all visits irrespective of sexual activity with penetration), and will be summarized and analyzed as described for vaginal dryness in **Section 6.3.1.1**. Missing values at Week 12 will be estimated using LOCF methodology. Only post-Baseline values will be carried forward.

6.3.2.5 Change from Baseline to Week 12 in Global Symptom Score 1

Global Symptom Score 1 is defined as the sum of the numerical scores of all individual symptoms (vaginal dryness; dyspareunia; pruritus or itching; burning; dysuria) for each subject at each time point: Baseline, Week 3 and Week 12 / Early Termination. The Baseline Global Symptom Score 1 will use the vaginal dryness assessment from the Screening visit and the assessments from the Baseline visit for all other symptoms.

The Global Symptom Score 1 may take the possible values ranging from 2 (at least moderate vaginal dryness as the only symptom) to 15 (all five symptoms severe in intensity) at Baseline/Screening, and possible values ranging from 0 (all symptoms absent) to 15 (all five

symptoms severe in intensity) at the Week 3 and Week 12 / Early Termination visits. The Global Symptom Score 1 will only be evaluated when all five individual symptom scores are available at a given time point.

The Global Symptom Score 1 will be summarized using descriptive statistics for continuous variables by treatment and visit. The change from Baseline value will also be summarized for all post-Baseline visits.

The mean change in Global Symptom Score 1 from Baseline to the Week 12 / Early Termination visit will be analyzed using the same parametric ANCOVA model as described for vaginal pH in **Section 6.3.1.2**. Missing values at Week 12 will be estimated using LOCF methodology. Only post-Baseline values will be carried forward.

6.3.2.6 Change from Baseline to Week 12 in Global Symptom Score 2

It is recognized that the number of missing assessments for Global Symptom Score 1 may be considerable due to its requirement for all five individual symptoms to have a response available and the dyspareunia symptom being applicable only in subjects who experienced sexual activity with penetration since the previous study visit. For this reason, the Global Symptom Score 2 is defined as an additional parameter.

Global Symptom Score 2 is defined as the sum of the numerical scores of the four individual symptoms excluding dyspareunia (vaginal dryness; pruritus or itching; burning; dysuria) for each subject at each time point: Baseline, Week 3 and Week 12 / Early Termination. The Baseline Global Symptom Score 2 will use the vaginal dryness assessment from the Screening visit and the assessments from the Baseline visit for all other symptoms.

The Global Symptom Score 2 may take the possible values ranging from 2 (at least moderate vaginal dryness as the only symptom) to 12 (all four symptoms severe in intensity) at Baseline/Screening, and possible values ranging from 0 (all symptoms absent) to 12 (all 4 symptoms severe in intensity) at the Week 3 and Week 12 / Early Termination visits. The Global Symptom Score 2 will only be evaluated when all four individual symptom scores are available at a given time point.

The Global Symptom Score 2 will be summarized using descriptive statistics for continuous variables by treatment and visit. The change from Baseline value will also be summarized for all post-Baseline visits.

The mean change in Global Symptom Score 2 from Baseline to the Week 12 / Early Termination visit will be analyzed using the same parametric ANCOVA model as described for vaginal pH in **Section 6.3.1.2**. Missing values at Week 12 will be estimated using LOCF methodology. Only post-Baseline values will be carried forward.

6.3.2.7 Change from Baseline to Week 12 in Severity of Pallor

Pallor is one of the five signs of vulvovaginal atrophy, and is recorded according to the visual inspection of the vagina as performed by the investigator at the Baseline, Week 3 and Week 12 / Early Termination visits.

Each sign will be scored by the investigator on a numeric Liekert scale, as follows:

0	Absent:	The sign is not present
1	Mild:	The sign is present and is considered a mild alteration
2	Moderate:	The sign is present and is considered a moderate alteration
3	Severe:	The sign is present and is considered a severe alteration

The number and percentage of subjects with each severity of pallor at each visit will be summarized by treatment in a general summary table of signs of vulvovaginal atrophy.

The change from Baseline to Week 12 in severity of pallor may take the possible values ranging from -3 (Severe at Baseline, Absent at Week 12) to 3 (Absent at Baseline, Severe at Week 12). This endpoint will be summarized and analyzed as described for vaginal dryness in **Section 6.3.1.1**. Missing values at Week 12 will be estimated using LOCF methodology. Only post-Baseline values will be carried forward.

6.3.2.8 Change from Baseline to Week 12 in Severity of Friability

Friability is one of the five signs of vulvovaginal atrophy. It is collected and summarized as described for pallor in **Section 6.3.2.7**, and will be analyzed as described for vaginal dryness in **Section 6.3.1.1**. Missing values at Week 12 will be estimated using LOCF methodology. Only post-Baseline values will be carried forward.

6.3.2.9 Change from Baseline to Week 12 in Severity of Thinning or Flattening of Folds

Thinning or flattening of folds is one of the five signs of vulvovaginal atrophy. It is collected and summarized as described for pallor in **Section 6.3.2.7**, and will be analyzed as described for vaginal dryness in **Section 6.3.1.1**. Missing values at Week 12 will be estimated using LOCF methodology. Only post-Baseline values will be carried forward.

6.3.2.10 Change from Baseline to Week 12 in Severity of Presence of Petechiae

Presence of petechiae is one of the five signs of vulvovaginal atrophy. It is collected and summarized as described for pallor in **Section 6.3.2.7**, and will be analyzed as described for

vaginal dryness in **Section 6.3.1.1**. Missing values at Week 12 will be estimated using LOCF methodology. Only post-Baseline values will be carried forward.

6.3.2.11 Change from Baseline to Week 12 in Severity of Dry Mucosa

Dry mucosa is one of the five signs of vulvovaginal atrophy. It is collected and summarized as described for pallor in **Section 6.3.2.7**, and will be analyzed as described for vaginal dryness in **Section 6.3.1.1**. Missing values at Week 12 will be estimated using LOCF methodology. Only post-Baseline values will be carried forward.

6.3.2.12 Change from Baseline to Week 3 in Severity of Vaginal Dryness

The change from Baseline to Week 3 in vaginal dryness will be analyzed as described in **Section 6.3.1.1**. Missing values will not be imputed.

6.3.2.13 Change from Baseline to Week 3 in Severity of Dyspareunia

The change from Baseline to Week 3 in dyspareunia will be analyzed as described in **Section 6.3.1.1**. Missing values will not be imputed.

6.3.2.14 Change from Baseline to Week 3 in Severity of Pruritus or Itching

The change from Baseline to Week 3 in pruritus or itching will be analyzed as described in **Section 6.3.1.1**. Missing values will not be imputed.

6.3.2.15 Change from Baseline to Week 3 in Severity of Burning

The change from Baseline to Week 3 in burning will be analyzed as described in **Section 6.3.1.1**. Missing values will not be imputed.

6.3.2.16 Change from Baseline to Week 3 in Severity of Dysuria

The change from Baseline to Week 3 in dysuria will be analyzed as described in **Section 6.3.1.1**. Missing values will not be imputed.

6.3.2.17 Change from Baseline to Week 3 in Global Symptom Score 1

The mean change in Global Symptom Score 1, as defined in **Section 6.3.2.5**, from Baseline to Week 3, will be analyzed as described in **Section 6.3.1.2**. Missing values will not be imputed.

6.3.2.18 Change from Baseline to Week 3 in Global Symptom Score 2

The mean change in Global Symptom Score 2, as defined in **Section 6.3.2.6**, from Baseline to Week 3 will be analyzed as described in **Section 6.3.1.2**. Missing values will not be imputed.

6.3.2.19 Change from Baseline to Week 3 in Severity of Pallor

The change from Baseline to Week 3 in pallor will be analyzed as described in **Section 6.3.1.1**. Missing values will not be imputed.

6.3.2.20 Change from Baseline to Week 3 in Severity of Friability

The change from Baseline to Week 3 in friability will be analyzed as described in **Section 6.3.1.1**. Missing values will not be imputed.

6.3.2.21 Change from Baseline to Week 3 in Severity of Thinning or Flattening of Folds

The change from Baseline to Week 3 in thinning or flattening of folds will be analyzed as described in **Section 6.3.1.1**. Missing values will not be imputed.

6.3.2.22 Change from Baseline to Week 3 in Severity of Presence of Petechiae

The change from Baseline to Week 3 in presence of petechiae will be analyzed as described in **Section 6.3.1.1**. Missing values will not be imputed.

6.3.2.23 Change from Baseline to Week 12 in Severity of Dry Mucosa

The change from Baseline to Week 3 in dry mucosa will be analyzed as described in **Section 6.3.1.1**. Missing values will not be imputed.

6.3.2.24 Change from Baseline to Week 3 in Severity of Vaginal pH

The mean change in vaginal pH from Baseline to Week 3 will be analyzed as described in **Section 6.3.1.2**. Missing values will not be imputed.

6.3.2.25 Change from Baseline to Week 3 in Proportion of Superficial Cells of the Vaginal Epithelium

The mean change in proportion of superficial cells from Baseline to Week 3 will be analyzed as described in **Section 6.3.1.3**. Missing values will not be imputed.

6.3.2.26 Change from Baseline to Week 3 in Proportion of Parabasal Cells of the Vaginal Epithelium

The mean change in proportion of parabasal cells from Baseline to Week 3 will be analyzed as described in **Section 6.3.1.3**. Missing values will not be imputed.

6.3.3 Exploratory Analysis

6.3.3.1 Evaluation of Global Efficacy and Acceptability

Global efficacy is defined as the subject's perception of the effect of the study treatment on her vaginal-related symptomatology. At the Week 12 / Early Termination visit, the subject will be asked to rate the global efficacy of the treatment as highly efficacious, efficacious, non-efficacious, or detrimental.

Acceptability is defined as the subject's perception of the ease in the method of administration and the cleanliness of the product, respectively. At the Week 12 / Early Termination visit, the subject will be asked to rate the acceptability of the treatment as excellent, good, acceptable or unacceptable.

The number and percentage of subjects with each rating of global efficacy (highly efficacious, efficacious, non-efficacious and detrimental) will be summarized by treatment group and overall for the ITT population.

The number and percentage of subjects with each rating of acceptability of method of administration and cleanliness of the product (excellent, good, acceptable and unacceptable) respectively will be summarized by treatment group and overall for the ITT population.

6.3.3.2 Change from Baseline to Week 12 in Maturation Value

The maturation value for a given pathologists's analysis of a given vaginal smear slide is defined as:

$$[0.2 \times \%\text{ parabasal cells}] + [0.6 \times \%\text{ intermediate cells}] + [1 \times \%\text{ superficial cells}]$$

The maturation value for each of the two pathologists' cytologic evaluation of each vaginal smear sample will be provided by [redacted] The maturation value for a given subject and visit to be used for analysis will be derived as the mean result across the two pathologists. If only one pathologist's result is available, then that result will be used to represent the overall result for that visit.

The mean change in maturation value from Baseline to Week 12 will be summarized and analyzed exactly as described for superficial cells in **Section 6.3.1.3.**

6.3.3.3 Change from Baseline to Week 3 in Maturation Value

The mean change in maturation value from Baseline to Week 3 will be analyzed as described in **Section 6.3.1.3.** Missing values will not be imputed.

6.4 Safety Analysis

Safety analysis will be performed using the Safety population. All outputs will be summarized by actual treatment received. Missing data will not be imputed.

6.4.1 Adverse Events

The assessment of safety will be based on the incidence of AEs occurring during the 12 weeks on study treatment. Adverse events will be recorded and monitored at all study visits. All AEs will be classified using the MedDRA coding dictionary. MedDRA Version 19.0 will be used, with one update during the study as per the Data Management Plan.

A treatment-emergent AE (TEAE) is defined in the study protocol as any event not present before exposure to study treatment, or any event already present that worsens in either intensity or frequency after exposure to study treatment. Events will be classified as *treatment-emergent* in the following situations:

- If the onset date of the AE is later than the date of first dose of study treatment.
- If the onset date of the AE is either unknown or is the same as the date of first dose of study treatment, and the AE is considered on the eCRF to have "onset after first dose of study drug".

- If the onset date of the AE is either unknown or is the same as the date of first dose of study treatment, and the AE is not clarified to have either "onset after first dose of study drug" or "onset before first dose of study drug" on the eCRF.

- If the onset date of the AE is partially known and it cannot be determined whether the AE started before or after the start of study treatment.

Treatment-emergent adverse events are therefore a subset of all adverse events.

The following presentations will be produced by treatment group and overall:

- A general summary table reporting the number and percentage of subjects and the number of events will be presented. The following categories will be included:

    - Subjects with no TEAEs
    - At least one TEAE
    - TEAEs leading to discontinuation of study treatment (defined as having the action taken with study treatment recorded as "drug withdrawn")
    - Treatment-related TEAEs (defined as having the relationship to study treatment recorded as "possibly", "probably", "definitely" or missing)
    - Serious TEAEs
    - Serious treatment-related TEAEs
    - TEAEs leading to death (defined as having the outcome recorded as "fatal")

- TEAEs will be summarized by SOC and preferred term. The number and percentage of subjects having each event, and the number of events will be presented. If a subject records multiple AEs with the same preferred term (i.e. with distinct start and stop dates), these shall be summarized once within the count of subject, yet each event will be counted within the number of reports of each TEAE.

- TEAEs will be summarized by SOC, preferred term and severity. The number and percentage of subjects having each severity for each event and the number of events will be presented. Any missing severity will be left as missing and categorized as such, although missing severity should be queried for completion by Data Management. If a subject has events with different severity for the same preferred term, each event will be summarized under its corresponding severity. This means that a subject may be counted across multiple severities for the same preferred term, which may result in the total number of subjects across all severities in the table appearing higher than the total number of subjects in the Safety population.

- TEAEs leading to discontinuation of study treatment will be summarized by SOC and preferred term. The number and percentage of subjects having each event, and the number of events will be presented.

- Treatment-related TEAEs will be summarized by SOC and preferred term. The number and percentage of subjects having each event, and the number of events will be presented. Any missing relationships will be left as missing and categorized as such, although missing relationships to study treatment will be queried for completion by Data Management.

- Serious TEAEs will be summarized by SOC and preferred term. The number and percentage of subjects having each event, and the number of events will be presented. Any missing serious criteria will be queried and must be completed.

Five general AE listings will be produced:

- All AEs
- Treatment-related AEs
- Serious AEs
- AEs leading to discontinuation of study treatment
- AEs leading to death

6.4.2 Physical Examination

A full physical examination is performed at the Screening and Week 12 / Early Termination visits, and will include assessments of the following body systems: General Appearance; HEENT; Cardiovascular; Respiratory; Gastrointestinal; Dermatological; Neurological; Musculoskeletal; Lymphatic; and Other.

The number and percentage of subjects with normal, abnormal not clinically significant and abnormal clinically significant results respectively will be summarized for each body system at each visit by treatment group and overall.

6.4.3 Vital Signs

Vital signs measurements are taken at the Screening and Week 12 / Early Termination visits. Vital signs collected will include systolic and diastolic blood pressures (millimeters of mercury (mmHg)) and heart rate (beats per minute) recorded after the subject has been seated for 5 minutes, height (cm) [at Screening only] and weight (kg).

Body Mass Index (BMI) will be derived at each visit as:

$$\frac{\text{Weight (kg)}}{[\text{Height (cm)} / 100]^2}$$

The absolute values of systolic blood pressure, diastolic blood pressure, heart rate, height, weight and BMI will be summarized at each visit by treatment group and overall, using summary statistics for continuous variables. Change from Baseline (Screening) values will also be summarized for systolic blood pressure, diastolic blood pressure, heart rate and weight at the Week 12 / Early Termination visit.

6.4.4 Electrocardiogram

Two 12-lead ECG examinations will be performed approximately 5 minutes apart from one another at each of the Screening and Week 12 / Early Termination visits. The ECG vendor, [redacted], will provide the following interval measurements for each ECG: Ventricular Rate (beats per minute); PR Interval (msec); RR Interval (msec); QRS Interval (msec); QT Interval (msec); QT Interval corrected according to Bazett Formula (QTcB, msec); and QT Interval corrected according to Fridericia Formula (QTcF, msec). The ECG vendor will also provide an overall interpretation of each ECG (normal, abnormal not clinically significant or abnormal clinically significant).

For each of the interval measurements, the mean of the two results at each visit will be evaluated for each subject to represent the overall result at that visit. If only one result is available for a given parameter at a particular visit, then that result will be used to represent the overall result for that parameter at that visit.

For the overall interpretation, the worst of the two results at each visit will represent the overall interpretation at that visit (where abnormal clinically significant represents the worst result, followed by abnormal not clinically significant and normal respectively). If only one overall interpretation is available at a particular visit, then that result will be used to represent the overall interpretation at that visit.

The absolute values of each interval measurement (ventricular rate, PR, RR, QRS, QT, QTcB and QTcF) will be summarized at each visit by treatment group and overall, using summary statistics for continuous variables. Change from Baseline (Screening) values will also be summarized for the Week 12 / Early Termination visit, along with the 95% CI for the mean change from Baseline.

The number and percentage of subjects with normal, abnormal not clinically significant and abnormal clinically significant overall interpretations respectively will be summarized at each visit by treatment group and overall.

6.4.5 Laboratory Parameters

Clinical laboratory tests will be performed under fasting conditions at the Screening and Week 12 / Early Termination visits. The following parameters will be assessed:

- Hematology:
  - Basophils
  - Eosinophils
  - Lymphocytes
  - Monocytes
  - Neutrophils
  - Basophils/Leukocytes
  - Eosinophils/Leukocytes
  - Lymphocytes/Leukocytes
  - Monocytes/Leukocytes
  - Neutrophils/Leukocytes
  - Hematocrit
  - Hemoglobin
  - Platelets
  - Erythrocytes
  - Leukocytes
  - Factor V Leiden (either at Screening or at Week 12 / Early Termination if not previously obtained at Screening)

- Blood Chemistry:
  - Albumin
  - Alanine Aminotransferase
  - Aspartate Aminotransferase
  - Blood Urea Nitrogen
  - Creatine Kinase
  - Creatinine
  - Glucose
  - Bilirubin
  - Protein
  - Urate
  - Triglycerides
  - Cholesterol
  - Cholesterol: High-Density Lipoprotein
  - Cholesterol: Low-Density Lipoprotein

- Urinalysis:
  - pH
  - Specific Gravity
  - Clarity
  - Colour
  - Protein
  - Glucose
  - Occult Blood

- Leukocytes – assessed routinely by dipstick; sediment method is used to confirm a positive dipstick result. If both dipstick and sediment results are present at the same visit, the sediment result will be used any applicable statistical summaries.
- Bilirubin
- Ketones
- Urobilinogen
- Nitrites
- Erythrocytes
- Bacteria
- Casts
- Crystals
- Epithelial Cells
- Yeast Cells

- Coagulation:
  - Activated Partial Thromboplastin Time
  - Prothrombin Time
  - Protein C
  - Protein S
  - Thrombin/Antithrombin III
  - Fibrinogen

- Endorinology (Hormones):
  - Estradiol
  - Estriol
  - Estrone
  - Follicle Stimulating Hormone (FSH)
  - Luteinizing Hormone (LH)

The samples will be analyzed by the central laboratory, [redacted] Reference ranges are available for all applicable parameters in Appendix I of the [redacted] Laboratory Manual. All parameters will be summarized in conventional units as provided by [redacted]

The absolute values of continuous laboratory parameters will be summarized at each visit by treatment group and overall, using summary statistics for continuous variables. Change from Baseline (Screening) values will also be summarized at the Week 12 / Early Termination visit. Categorical laboratory parameters will be listed only.

Tables presenting the shift in result (low, normal, high) from Baseline (Screening) to Week 12 / Early Termination will also be presented for each laboratory parameter by treatment group. For the shift tables, the denominator of percentage will be the number of subjects in the Safety population in each treatment group.

6.4.6 Transvaginal Ultrasound

A transvaginal ultrasound is performed at the Screening and Week 12 / Early Termination visits for subjects with an intact uterus, and will include assessments of endometrial thickness and pelvic pathology (ovaries, miomas, etc.).

The number and percentage of subjects with normal, abnormal not clinically significant and abnormal clinically significant results, respectively, will be summarized for each assessment at each visit by treatment group and overall.

A table presenting the shift in result (normal, abnormal not clinically significant, abnormal clinically significant) from Baseline (Screening) to Week 12 / Early Termination visit will also be presented for each assessment by treatment group. For the shift table, the denominator of percentage will be the number of subjects in the Safety population in each treatment group.

6.4.7 Gynecological Examination

A gynecological examination is performed at the Screening and Week 12 / Early Termination visits, and will include the following assessments: Visual Inspection of External Genitalia; Speculum Examination of Vagina; Speculum Examination of Cervix (if present); Bimanual Palpation of Uterus and Ovaries (if present); Breast Palpation; and Papanicolau Exam (in subjects with a cervix, at Screening only).

The number and percentage of subjects with normal, abnormal not clinically significant and abnormal clinically significant results respectively will be summarized for each assessment at each visit by treatment group and overall.

The papanicolau examination is conducted at the Screening visit only and the associated normal / abnormal not clinically significant / abnormal clinically significant result is recorded on the eCRF. These results will be summarized as described above. The smear samples are sent to the central laboratory, [redacted], for analysis, and the detailed results from the cytopathology reports will be listed only.

6.4.8 Endometrial Biopsy

An endometrial biopsy is performed in subjects with a uterus as a mandatory procedure at the Screening and Week 12 / Early Termination visits, and whenever the subject presents with vaginal bleeding at any other visit.

The slides from the endometrial biopsies will be sent to the central laboratory, [redacted] for evaluation. Three independent pathologists, who will be blinded to subject, treatment assignment and to each other's evaluations, will assess the endometrial biopsies obtained at the Week 12 / Early Termination visit plus 10% of the Screening slides and all slides diagnosed with hyperplasia or cancer. The slides from the Screening visit and from any visit where the subject presents with vaginal bleeding will be assessed by only the primary pathologist.

The primary pathologist will be from [redacted], and the other two pathologists will be from different institutions but will travel to [redacted] to evaluate the endometrial biopsies.

Each pathologist will provide their evaluation according to the following categories, which are ordered from best to worst diagnosis, i.e. atypical hyperplasia (carcinoma) is the most severe diagnosis, followed by complex hyperplasia, simple hyperplasia and benign endometrium accordingly.

Table 5: Histologic Characteristics of the Endometrium

0	No tissue
1	Tissue insufficient for diagnosis
2	Atrofic
3	Inactive
4a	Weakly proliferative
4b	Active proliferative
4c	Disordered proliferative
5a	Secretory – Cyclic type
5b	Secretory – Progestational type (including stromal decidualization)
6	Menstrual type
7	Simple hyperplasia without atypia
8	Simple hyperplasia with atypia
9	Complex hyperplasia without atypia
10	Complex hyperplasia with atypia
11	Carcinoma (specify type)

For any subject, the final diagnosis will be accepted as the concurrence of two of the three pathologists. If there is no agreement among the three pathologists, the most severe pathologic diagnosis will be used as the final diagnosis. The evaluations of all three pathologists must be verified by the CRA in order to be considered for final analysis.

If hyperplasia is diagnosed by the single safety reader for a subject who has bled while on study treatment, this diagnosis will be maintained for the efficacy/safety evaluation and the slides become part of the slide set given to the two other pathologists for reading.

The bullet points below specify how many pathologists will review the biopsy samples at each visit and how the final diagnosis of a sample will be programmatically derived at each visit.

- Screening Visit
  - Biopsy sample is mandatory.
  - Samples reviewed by Pathologist 1 only.
    - Final diagnosis = diagnosis of Pathologist 1
  - 10% of samples will also be reviewed by Pathologist 2 and Pathologist 3 **for QC purposes only.** The diagnoses of Pathologist 2 and Pathologist 3 will **not** be used in the determination of final diagnosis.

- All slides diagnosed with hyperplasia or cancer (categories 7, 8, 9, 10 or 11 from Table 5 above) by Pathologist 1 will also be reviewed by Pathologist 2 and Pathologist 3 **for QC purposes only**. The diagnoses of Pathologist 2 and Pathologist 3 will **not** be used in the determination of final diagnosis.

- Baseline, Week 3 and Week 8 Visits
  - Biopsy sample is only taken if the subject presents with vaginal bleeding at the visit, otherwise the eCRF is marked as Not Applicable.
  - Samples reviewed by Pathologist 1 only.
    - Final diagnosis = diagnosis of Pathologist 1
  - All slides diagnosed with hyperplasia or cancer (categories 7, 8, 9, 10 or 11 from Table 5 above) by Pathologist 1 will also be reviewed by Pathologist 2 and Pathologist 3 **for QC purposes only**. The diagnoses of Pathologist 2 and Pathologist 3 will **not** be used in the determination of final diagnosis.

- Week 12 / Early Termination Visit
  - Biopsy sample is mandatory.
  - Samples reviewed by Pathologist 1, Pathologist 2 and Pathologist 3.
    - Final diagnosis derived as follows:
      - If the subject presents with vaginal bleeding at the visit (as indicated by an answer of 'Yes' to the direct question on the eCRF) and Pathologist 1 diagnoses hyperplasia or cancer (categories 7, 8, 9, 10 or 11 from Table 5 above):
        - Final diagnosis = diagnosis of Pathologist 1, **regardless** of the diagnoses of Pathologist 2 and Pathologist 3, even if Pathologist 2 and Pathologist 3 concur on a different diagnosis, either more or less severe than that of Pathologist 1.
      - Otherwise:
        - If two of the three pathologists agree (i.e. diagnose the same category from Table 5 above), their agreed diagnosis is the final diagnosis.
        - If all three pathologists diagnose a different category from Table 5 above), the most severe diagnosis is taken as final diagnosis.

For summary purposes, the Baseline histologic result will be the final diagnosis from the Baseline visit if endometrial samples were taken due to the subject presenting with vaginal bleeding at that visit, or the final diagnosis from the Screening visit otherwise.

The number and percentage of subjects with each histologic characteristic will be summarized (according to the final diagnosis, as defined above) at each visit by treatment group and overall.

A table presenting the shift in result from Baseline to Week 12 / Early Termination (according to the final diagnosis) will also be presented by treatment group. For the shift table, the denominator of percentage will be the number of subjects in the Safety population in each treatment group.

6.5 Analysis of Other Assessments

No other assessments are planned to be analyzed in the protocol for this study.

6.6 Interim Analysis

There is no planned interim analysis for this study.

6.7 Data Monitoring Committee

There is no planned Data Monitoring Committee review for this study.

6.8 Changes to Methods Planned in the Protocol

- Section 6.2.3.3 of the protocol Version 5.0 describes the Global Symptom Score as the sum of the numerical scores of all five symptoms of vulvovaginal atrophy at a given time point, and states that this score will be evaluated only when all five individual symptoms scores are available. This Global Symptom Score is identified in this Statistical Analysis Plan (SAP) as Global Symptom Score 1, and is analyzed as described in **Section 6.3.2.5**, and as planned in the protocol. This score is labeled as Global Symptom Score 1 in order to distinguish it from the alternative score described in the following bullet.

- The Global Symptom Score 1 requires all five individual symptoms to have a response available. Given that the dyspareunia symptom is only applicable in subjects who experienced sexual activity with penetration since the previous study visit, it is acknowledged that there may be a considerable number of subjects / time points where the Global Symptom Score 1 is not available. In order to allow a Global Symptom Score analysis that utilizes as many subjects as possible, an additional score has been defined that was not planned in the protocol Version 5.0. The Global Symptom Score 2 is defined as the sum of the four individual symptoms excluding dyspareunia, requires all 4 of these symptoms to be available, and is analyzed as described in **Section 6.3.2.6**.

- Although not required by the protocol, additional exploratory ANCOVA analyses are planned for the mean maturation values from the vaginal smear samples. The change from Baseline to Week 12 and the change from Baseline to Week 3 will be analyzed using the same ANCOVA approach described in the for the percentage of superficial cells and the percentage of parabasal cells. The analyses of these additional parameters have been included because they are considered to be relevant.

- Additional sensitivity MMRM analyses are planned for the four co-primary efficacy endpoints. These extra analyses have been added for the purposes of robust exploratory comparison against the LOCF approach used for the primary analyses.

7. TABLES, LISTINGS, AND FIGURES

The general layout of the tables, listings and figures will be as follows:

Orientation	Landscape
Paper Size	A4
Margins	Top: 3.2 cm Bottom: 2.5 cm Left: 2.5 cm Right: 2.5 cm
Font	Courier New 8 point
Headers (alignment)	Sponsor name and Protocol number (Left) Page X of Y (Right) Table/Listing/Figure Number and Title (Centre)
Footers (alignment)	SAS program path and file name (Left) Date output generated (Right)

The font size may be reduced as necessary to allow additional columns to be presented, but not at the expense of clarity. Also, the orientation may be changed to portrait, if appropriate.

The date format for all presentations will be 'DDMMMYYYY'.

All statistical analysis will be performed using SAS® Version 9.3 or higher for Windows.

The treatment labels for all outputs will be as shown in **Table 6.**

Table 6: Treatment Labels for Statistical Outputs

0.005% Estriol
0.002% Estriol
0.008% Estriol
Placebo
Total *(Tables only)*

7.1 Preparation of Tables

The intended layouts for unique summary tables are presented in Appendix I. However, it may be necessary to change the table layouts, as appropriate, upon review of the data available.

7.2 Table of Contents for Tables

7.2.1 Demographic and Background Data

Table 14.1.1.1	Summary of Subject Disposition, by Treatment and Overall	All Subjects
Table 14.1.1.2	Summary of Inclusion/Exclusion Criteria Not Met	All Subjects
Table 14.1.2.1	Reasons for Exclusion from Per-Protocol Population	ITT Population
Table 14.1.2.2	Protocol Deviations from the PD Tracker	ITT Population
Table 14.1.3	Summary of Subject Demographics, by Treatment and Overall	ITT Population
Table 14.1.4	Summary of Treatment Exposure and Compliance, by Treatment and Overall	Safety Population
Table 14.1.5	Summary of Medical History by System Organ Class, Preferred Term, Treatment and Overall	Safety Population
Table 14.1.6.1	Summary of Prior Medications by ATC, Preferred Term, Treatment and Overall	Safety Population
Table 14.1.6.2	Summary of Concomitant Medications by ATC, Preferred Term, Treatment and Overall	Safety Population
Table 14.1.7	Summary of Other Concomitant Diagnostic or Therapeutic Procedures by System Organ Class, Preferred Term, Treatment and Overall	Safety Population

7.2.2 Efficacy Data

Table 14.2.1.1.1	Co-Primary Efficacy Analysis of Change from Baseline to Week 12 in Severity of Vaginal Dryness (LOCF)	ITT Population
Table 14.2.1.1.2	Analysis of Change from Baseline to Week 12 in Severity of Vaginal Dryness	PP Population
Table 14.2.1.1.3	Sensitivity Analysis of Change from Baseline to Week 12 in Severity of Vaginal Dryness (Observed Case)	ITT Population
Table 14.2.1.1.4	Sensitivity Analysis of Change from Baseline in Severity of Vaginal Dryness (MMRM)	ITT Population
Table 14.2.1.2.1	Co-Primary Efficacy Analysis of Change from Baseline to Week 12 in Vaginal pH (LOCF)	ITT Population

Table 14.2.1.2.2	Analysis of Change from Baseline to Week 12 in Vaginal pH	PP Population
Table 14.2.1.2.3	Sensitivity Analysis of Change from Baseline to Week 12 in Vaginal pH (Observed Case)	ITT Population
Table 14.2.1.2.4	Sensitivity Analysis of Change from Baseline in Vaginal pH (MMRM)	ITT Population
Table 14.2.1.3.1	Co-Primary Efficacy Analysis of Change from Baseline to Week 12 in Proportion of Superficial Cells of the Vaginal Epithelium (LOCF)	ITT Population
Table 14.2.1.3.2	Analysis of Change from Baseline to Week 12 in Proportion of Superficial Cells of the Vaginal Epithelium	PP Population
Table 14.2.1.3.3	Sensitivity Analysis of Change from Baseline to Week 12 in Proportion of Superficial Cells of the Vaginal Epithelium (Observed Case)	ITT Population
Table 14.2.1.3.4	Sensitivity Analysis of Change from Baseline in Proportion of Superficial Cells of the Vaginal Epithelium (MMRM)	ITT Population
Table 14.2.1.4.1	Co-Primary Efficacy Analysis of Change from Baseline to Week 12 in Proportion of Parabasal Cells of the Vaginal Epithelium (LOCF)	ITT Population
Table 14.2.1.4.2	Analysis of Change from Baseline to Week 12 in Proportion of Parabasal Cells of the Vaginal Epithelium	PP Population
Table 14.2.1.4.3	Sensitivity Analysis of Change from Baseline to Week 12 in Proportion of Parabasal Cells of the Vaginal Epithelium (Observed Case)	ITT Population
Table 14.2.1.4.4	Sensitivity Analysis of Change from Baseline in Proportion of Parabasal Cells of the Vaginal Epithelium (MMRM)	ITT Population
Table 14.2.1.5.1	Gatekeeping Table for Co-Primary Efficacy Analysis of Change from Baseline to Week 12 (LOCF)	ITT Population
Table 14.2.1.5.2	Gatekeeping Table for Analysis of Change from Baseline to Week 12	PP Population
Table 14.2.1.5.3	Gatekeeping Table for Sensitivity Analysis of Change from Baseline to Week 12 (Observed Case)	ITT Population
Table 14.2.1.5.4	Gatekeeping Table for Sensitivity Analysis of Change from Baseline to Week 12 (MMRM)	ITT Population
Table 14.2.2.1	Analysis of Change from Baseline to Week 3 in Severity of Vaginal Dryness	ITT Population
Table 14.2.2.2.1	Analysis of Change from Baseline to Week 12 in Severity of Dyspareunia (LOCF)	ITT Population
Table 14.2.2.2.2	Analysis of Change from Baseline to Week 3 in Severity of Dyspareunia	ITT Population
Table 14.2.2.3.1	Analysis of Change from Baseline to Week 12 in Severity of Pruritus or Itching (LOCF)	ITT Population
Table 14.2.2.3.2	Analysis of Change from Baseline to Week 3 in Severity of Pruritus or Itching	ITT Population

Table 14.2.2.4.1	Analysis of Change from Baseline to Week 12 in Severity of Burning (LOCF)	ITT Population
Table 14.2.2.4.2	Analysis of Change from Baseline to Week 3 in Severity of Burning	ITT Population
Table 14.2.2.5.1	Analysis of Change from Baseline to Week 12 in Severity of Dysuria (LOCF)	ITT Population
Table 14.2.2.5.2	Analysis of Change from Baseline to Week 3 in Severity of Dysuria	ITT Population
Table 14.2.2.6.1	Analysis of Change from Baseline to Week 12 in Global Symptom Score 1 (LOCF)	ITT Population
Table 14.2.2.6.2	Analysis of Change from Baseline to Week 3 in Global Symptom Score 1	ITT Population
Table 14.2.2.7.1	Analysis of Change from Baseline to Week 12 in Global Symptom Score 2 (LOCF)	ITT Population
Table 14.2.2.7.2	Analysis of Change from Baseline to Week 3 in Global Symptom Score 2	ITT Population
Table 14.2.2.8.1	Summary of Symptoms of Vulvovaginal Atrophy by Treatment and Visit	ITT Population
Table 14.2.2.8.2	Summary of Global Symptom Score 1 and 2 by Treatment and Visit	ITT Population
Table 14.2.3.1.1	Analysis of Change from Baseline to Week 12 in Severity of Pallor (LOCF)	ITT Population
Table 14.2.3.1.2	Analysis of Change from Baseline to Week 3 in Severity of Pallor	ITT Population
Table 14.2.3.2.1	Analysis of Change from Baseline to Week 12 in Severity of Friability (LOCF)	ITT Population
Table 14.2.3.2.2	Analysis of Change from Baseline to Week 3 in Severity of Friability	ITT Population
Table 14.2.3.3.1	Analysis of Change from Baseline to Week 12 in Severity of Thinning or Flattening of Folds (LOCF)	ITT Population
Table 14.2.3.3.2	Analysis of Change from Baseline to Week 3 in Severity of Thinning or Flattening of Folds	ITT Population
Table 14.2.3.4.1	Analysis of Change from Baseline to Week 12 in Severity of Presence of Petechiae (LOCF)	ITT Population
Table 14.2.3.4.2	Analysis of Change from Baseline to Week 3 in Severity of Presence of Petechiae	ITT Population
Table 14.2.3.5.1	Analysis of Change from Baseline to Week 12 in Severity of Dry Mucosa (LOCF)	ITT Population
Table 14.2.3.5.2	Analysis of Change from Baseline to Week 3 in Severity of Dry Mucosa	ITT Population
Table 14.2.3.6	Summary of Signs of Vulvovaginal Atrophy by Treatment and Visit	ITT Population
Table 14.2.4.1	Analysis of Change from Baseline to Week 3 in Vaginal pH	ITT Population

Table 14.2.4.2	Summary of Vaginal pH by Treatment and Visit	ITT Population
Table 14.2.5.1	Analysis of Change from Baseline to Week 3 in Proportion of Superficial Cells of the Vaginal Epithelium	ITT Population
Table 14.2.5.2	Analysis of Change from Baseline to Week 3 in Proportion of Parabasal Cells of the Vaginal Epithelium	ITT Population
Table 14.2.5.3.1	Analysis of Change from Baseline to Week 12 in Maturation Value (LOCF)	ITT Population
Table 14.2.5.3.2	Analysis of Change from Baseline to Week 12 in Maturation Value	PP Population
Table 14.2.5.3.3	Analysis of Change from Baseline to Week 12 in Maturation Value (Observed Case)	ITT Population
Table 14.2.5.3.4	Analysis of Change from Baseline to Week 3 in Maturation Value	ITT Population
Table 14.2.5.4	Summary of Vaginal Cytology Parameters by Treatment and Visit	ITT Population
Table 14.2.6	Summary of Evaluation of Global Efficacy and Acceptability by Treatment	ITT Population

7.2.3 Safety Data

Table 14.3.1.1	Summary of Treatment-Emergent Adverse Events (TEAEs) by Treatment and Overall	Safety Population
Table 14.3.1.2	Summary of Treatment-Emergent Adverse Events (TEAEs) by System Organ Class, Preferred Term, Treatment and Overall	Safety Population
Table 14.3.1.3	Summary of Treatment-Emergent Adverse Events (TEAEs) by Severity, System Organ Class, Preferred Term, Treatment and Overall	Safety Population
Table 14.3.1.4	Summary of Treatment-Emergent Adverse Events (TEAEs) Leading to Discontinuation of Study Treatment by System Organ Class, Preferred Term, Treatment and Overall	Safety Population
Table 14.3.1.5	Summary of Treatment-Related Treatment-Emergent Adverse Events (TEAEs) by System Organ Class, Preferred Term, Treatment and Overall	Safety Population
Table 14.3.1.6	Summary of Serious Treatment-Emergent Adverse Events (TEAEs) by System Organ Class, Preferred Term, Treatment and Overall	Safety Population
Table 14.3.2.1	Listing of Serious Adverse Events	All Subjects
Table 14.3.2.2	Adverse Events Leading to Discontinuation of Study Treatment	All Subjects
Table 14.3.2.3	Listing of Adverse Events Leading to Death	All Subjects

Table 14.3.4.1.1.1	Summary of Hematology Parameters by Treatment and Visit	Safety Population
Table 14.3.4.1.1.2	Summary of Factor V Leiden Results by Treatment	Safety Population
Table 14.3.4.1.2	Summary of Blood Chemistry Parameters by Treatment and Visit	Safety Population
Table 14.3.4.1.3	Summary of Urinalysis Parameters by Treatment and Visit	Safety Population
Table 14.3.4.1.4	Summary of Coagulation Parameters by Treatment and Visit	Safety Population
Table 14.3.4.1.5	Summary of Endocrinology (Hormones) Parameters by Treatment and Visit	Safety Population
Table 14.3.4.2.1	Summary of Shifts from Baseline to Week 12 / Early Termination for Hematology Parameters	Safety Population
Table 14.3.4.2.2	Summary of Shifts from Baseline to Week 12 / Early Termination for Blood Chemistry Parameters	Safety Population
Table 14.3.4.2.3	Summary of Shifts from Baseline to Week 12 / Early Termination for Urinalysis Parameters	Safety Population
Table 14.3.4.2.4	Summary of Shifts from Baseline to Week 12 / Early Termination for Coagulation Parameters	Safety Population
Table 14.3.4.2.5	Summary of Shifts from Baseline to Week 12 / Early Termination for Endocrinology (Hormones) Parameters	Safety Population
Table 14.3.5.1	Summary of Transvaginal Ultrasound by Treatment and Visit	Safety Population
Table 14.3.5.2	Summary of Shifts from Baseline to Week 12 / Early Termination for Transvaginal Ultrasound Assessments	Safety Population
Table 14.3.6.1	Summary of Endometrial Biopsy Histologic Characteristics by Treatment and Visit	Safety Population
Table 14.3.6.2	Summary of Shifts from Baseline to Week 12 / Early Termination for Endometrial Biopsy Histologic Characteristics	Safety Population
Table 14.3.7.1	Summary of 12-Lead Electrocardiogram Parameters by Treatment and Visit	Safety Population
Table 14.3.7.2	Summary of 12-Lead Electrocardiogram Overall Interpretation by Treatment and Visit	Safety Population
Table 14.3.8	Summary of Vital Signs by Treatment and Visit	Safety Population
Table 14.3.9	Summary of Physical Examination by Treatment and Visit	Safety Population
Table 14.3.10	Summary of Gynecological Examination by Treatment and Visit	Safety Population

7.3 Preparation of Data Listings

The intended layouts for listings are presented in Appendix II. However, it may be necessary to change the listing layouts, as appropriate, upon review of the data available.

Data will be presented within the data listings according to the following order:

- Unique subject number
- Treatment group
- Visits: Screening, Baseline, Week 3, Week 8, Week 12 / Early Termination

The sorting order may be changed as appropriate for the data in each listing.

All subjects with relevant data will be included in data listings.

7.4 Table of Contents for Data Listings

7.4.1 Subject Data Listings

Listing 16.2.1.1	Completed and Discontinued Subjects
Listing 16.2.1.2	Informed Consent
Listing 16.2.2.1	Subjects Excluded from Per-Protocol Population
Listing 16.2.2.2	Subjects with Protocol Deviations from the PD Tracker
Listing 16.2.3	Subjects Included in Analysis Populations
Listing 16.2.4	Demographic Data
Listing 16.2.5.1	Study Treatment Administration and Compliance Data
Listing 16.2.5.2	Study Treatment Exposure and Compliance: Derived Parameters

7.4.2 Efficacy Data

Listing 16.2.6.1	Co-Primary Efficacy Parameters: Vaginal Dryness, Vaginal pH, Proportions of Superficial and Parabasal Cells
Listing 16.2.6.2	Symptoms of Vulvovaginal Atrophy
Listing 16.2.6.3	Signs of Vulvovaginal Atrophy
Listing 16.2.6.4	Vaginal Cytology
Listing 16.2.6.5	Evaluation of Global Efficacy and Acceptability

7.4.3 Adverse Events

Listing 16.2.7.1	Adverse Events
Listing 16.2.7.2	Treatment-Related Adverse Events

7.4.4 Laboratory Parameters

Listing 16.2.8.1	Laboratory Parameters: Hematology
Listing 16.2.8.2	Laboratory Parameters: Blood Chemistry
Listing 16.2.8.3	Laboratory Parameters: Urinalysis
Listing 16.2.8.4	Laboratory Parameters: Coagulation
Listing 16.2.8.5	Laboratory Parameters: Endocrinology (Hormones)

7.4.5 Other Safety Data

Listing 16.2.9	Transvaginal Ultrasound
Listing 16.2.10	Endometrial Biopsy
Listing 16.2.11	12-Lead Electrocardiogram

7.4.6 Archive Subject Data Listings

Listing 16.4.1	Inclusion/Exclusion Criteria Not Met
Listing 16.4.2	Medical History
Listing 16.4.3.1	Prior Medications
Listing 16.4.3.2	Concomitant Medications
Listing 16.4.3.3	Other Concomitant Diagnostic or Therapeutic Procedures
Listing 16.4.4	Physical Examination
Listing 16.4.5	Vital Signs
Listing 16.4.6.1	Gynecological Examination
Listing 16.4.6.2	Papanicolau Smear Results from Central Laboratory

8. REFERENCES

The 2012 Hormone Therapy Position Statement of the North American Menopause Society (2012). Menopause, 19:257-271.

Cano A, Estevez J, Usandiziga R, et al (2012). The therapeutic effect of a new ultra low concentration estriol gel formulation (0.005% estriol vaginal gel) on symptoms and signs of postmenopausal vaginal atrophy: results from a pivotal phase III study. Menopause, 19:1130-1139.

9. APPENDICES

Appendix I: Table Shells for Unique Summary Tables
Appendix II: Layout for Data Listings

APPENDIX I

Table Shells for Unique Summary Tables

APPENDIX II

Layout for Data Listings