Statistical Analysis Plan for Study M16-283

A Phase 3b Study to Evaluate the Long-Term Safety of Elagolix in Combination with Estradiol/Norethindrone Acetate for the Management of Heavy Menstrual Bleeding Associated with Uterine Fibroids in Premenopausal Women

Date: 20 January 2021

Version 1.0

Table of Contents

List of Tables

List of Figures

List of Appendices

1.0 Introduction

This Statistical Analysis Plan (SAP) describes the statistical analyses for Study M16-283" A Phase 3b Study to Evaluate the Long-Term Safety of Elagolix in Combination with Estradiol/Norethindrone Acetate for the Management of Heavy Menstrual Bleeding Associated with Uterine Fibroids in Premenopausal Women."

Study M16-283 examines the long-term safety of elagolix 300 mg administered twice daily (BID) in combination with estradiol 1 mg/0.5 mg norethindrone acetate (E2/NETA 1 mg/0.5 mg) in premenopausal women with heavy menstrual bleeding associated with uterine fibroids.

The analyses of pharmacokinetic endpoints, pharmacodynamic biomarker endpoints and exploratory research endpoints will not be covered in this SAP.

The SAP will not be updated in case of administrative changes or amendments to the protocol unless the changes impact the analysis.

Unless noted otherwise, all analyses will be performed using SAS Version 9.4 (SAS Institute Inc., Cary, NC 27513) or later under the UNIX operating system.

2.0 Study Design, Objectives and Procedures

2.1 Objectives

The objectives of this study are to:

- Assess the safety of elagolix 300 mg administered twice daily (BID) in combination with E2/NETA (estradiol 1.0 mg/norethindrone acetate 0.5 mg QD) compared to placebo at 12 months in premenopausal women with heavy menstrual bleeding associated with uterine fibroids,
- Characterize the impact of elagolix 300 mg BID with E2/NETA on bone mineral density (BMD) in women with HMB associated with uterine fibroids after up to 48 months of treatment.

2.2 Study Design Overview

This Phase 3b, randomized, 12-month double-blind placebo-controlled, 36 month open-label, multi-center study is designed to evaluate the safety of elagolix with add-back therapy in women with HMB associated with uterine fibroids.

Approximately 500 subjects will be randomized in a 2:1 ratio to one of the following 2 treatment groups:

1. Elagolix 300 mg BID plus E2/NETA (estradiol 1.0 mg/norethindrone acetate 0.5 mg QD) (n = 333) for 48 months
2. Placebo (n = 167) for 12 months

Subjects randomized to elagolix 300 mg BID plus E2/NETA in the 12-month Treatment Period will receive the same treatment during the open-label Treatment Period. Subject randomized to placebo in the 12-month Treatment Period will switch to elagolix 300 mg BID plus E2/NETA during the open-label Treatment Period after completing the 12-month Treatment Period.

The total duration for a subject's participation in this study is approximately 77 months, consisting of the following 4 study periods:

1. Washout Period – up to 10 months prior to Screening (only applicable if subject is taking exclusionary medication at the time of consent; duration of washout depends on the type of exclusionary medication being taken).
2. Screening Period – approximately 2.5 to 5 months prior to first dose of study drug.
3. Treatment Period – up to 48-month treatment duration
4. Post-Treatment Follow-Up Period – up to 12 months duration following the last dose of study drug. Subjects are expected to enter Post-Treatment Follow-Up after completing Treatment Period Month 48, or at any time a subject prematurely discontinues during the Treatment Period.

The Study Schematic is shown below in Figure 1.

Figure 1. Study Schematic

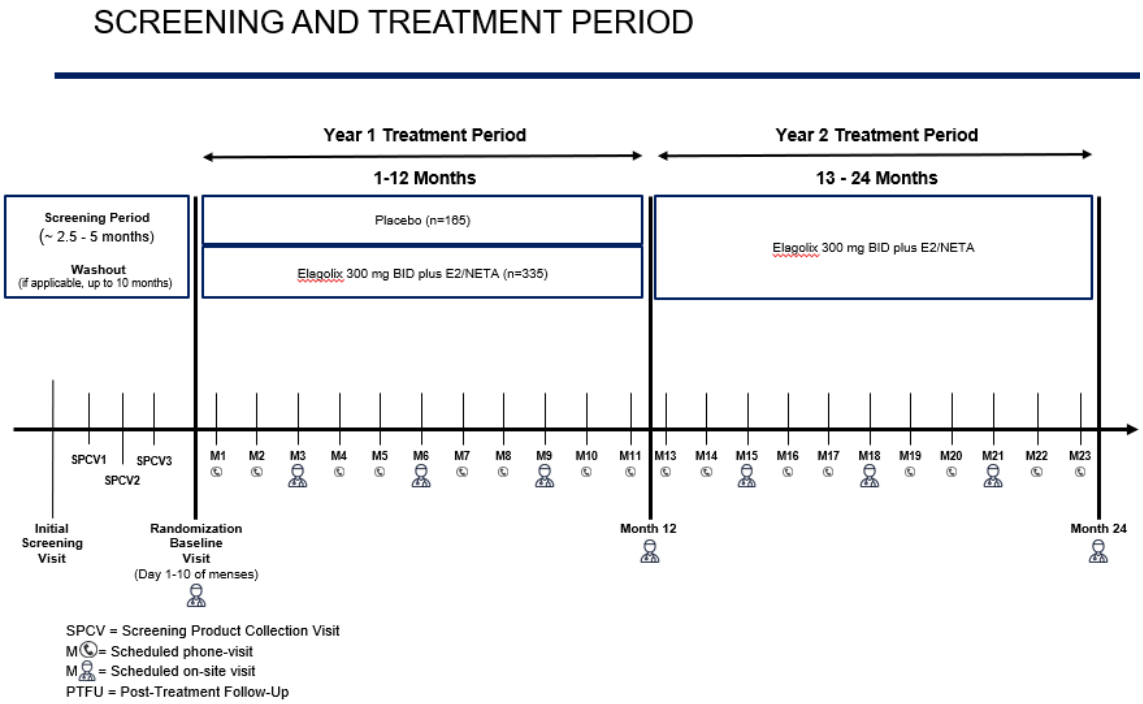
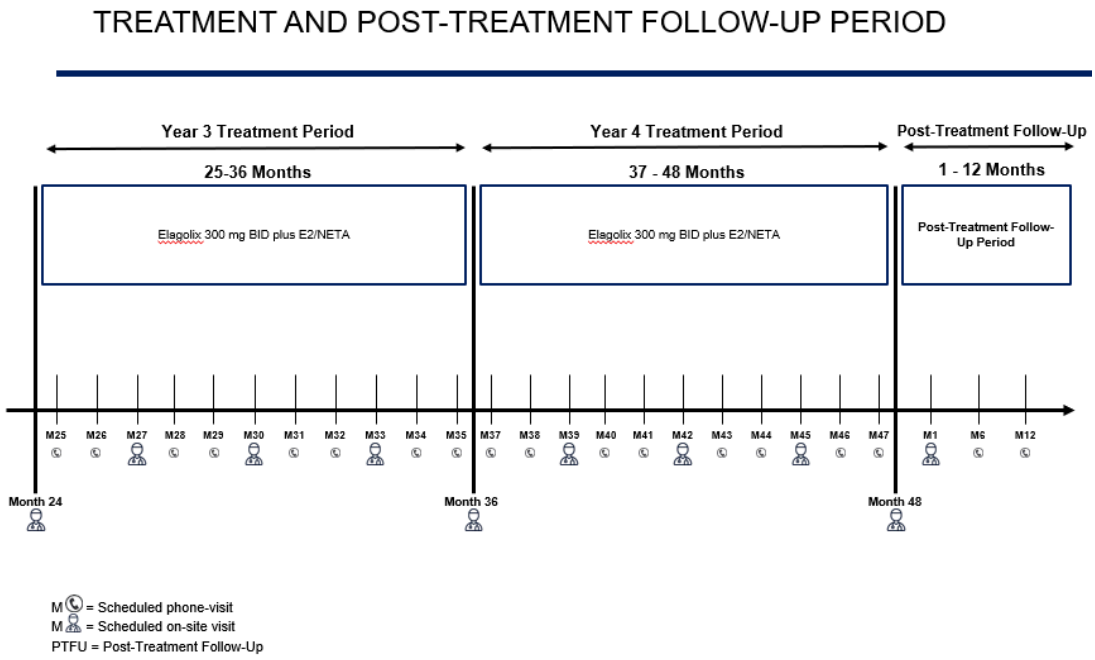


Figure 1. Study Schematic (Continued)



2.3 Treatment Assignment and Blinding

Approximately 500 subjects will be randomily assigned in a 2:1 ratio to one of the following 2 treatment groups:

1. Elagolix 300 mg BID plus E2/NETA (estradiol 1.0 mg/norethindrone acetate 0.5 mg QD) (n = 333)
2. Placebo (n = 167)

Unless otherwise specified, data will be summarized by group as listed above.

2.4 Sample Size Determination

Approximately 500 subjects will be randomily assigned (2:1) to elagolix 300 mg BID plus E2/NETA (N = 333) and placebo (N = 167).

Based on clinical review, a sample size of 500 was selected to gather long term safety exposure data in approximately 50 subjects completing 48 months of study drug dosing, with approximately 30 subjects receiving 48 months of active drug elagolix 300 mg BID plus E2/NETA 1/0.5 mg QD.

2.5 End of 12-Month Placebo-Controlled Treatment Period Analysis

An end-of-placebo-controlled treatment period analysis will be performed after the last subject completes the 12-Month placebo-controlled Treatment Period. These analyses will include data collected during the full Treatment Period (placebo-controlled and open-label Treatment Periods) and the Follow-Up Period up to the data cut-off for the interim lock. The database will be versioned for an interim lock and any discrepant data will be clarified before the versioning. Analyses will be performed by AbbVie.

Since this end-of-placebo-controlled treatment period analysis is the only and final analysis for the comparison between treatment and placebo group at 12 months of safety and efficacy endpoints, no additional adjustment of alpha-level is necessary.

2.6 Treatment Month 24, Month 36 and Month 48 Analyses

At the end of Treatment Month 24, Month 36 and Month 48, analysis of the safety and efficacy safety variables may be performed after the last subject completes the Treatment Month 24, Month 36 and Month 48 visit, respectively. For subjects who prematurely discontinued during the treatment period, their data collected during the Follow-up Period will be included in the analysis, as appropriate. The database will be versioned for an interim lock and any discrepant data will be clarified before the versioning. Analyses will be performed by AbbVie.

3.0 Endpoints

3.1 Primary Endpoint

As the objective of this study is to evaluate the long-term safety and tolerability of elagolix 300 mg BID with E2/NETA, no primary efficacy endpoint is defined.

3.2 Secondary Endpoint

As the objective of this study is to evaluate the long-term safety and tolerability of elagolix 300 mg BID with E2/NETA, no secondary efficacy endpoint is defined.

3.3 Other Efficacy Endpoints

Other efficacy endpoints include the following:

- PGIC for Menstrual Bleeding Symptom at Month 3, Month 6, Month 9 and Month 12;
- Change from baseline for the UFS-QoL at Month 6 and Month 12.

3.4 Safety Endpoints

The safety endpoints will be based on the following evaluations:

- Dual energy X-ray absorptiometry (DXA) scan
- Adverse events monitoring
- Clinical laboratory tests
- Vital signs
- Endometrial biopsy
- Pelvic ultrasound (transabdominal ultrasound [TAU]/transvaginal ultrasound [TVU])
- Columbia Suicide Severity Rating Scale (C-SSRS)
- Pregnancy

4.0 Analysis Populations

The Full Analysis Set (FAS) is composed of all randomized subjects who have received at least 1 dose of study drug in this study. The data from the FAS will be presented by the treatment group assigned at the time of randomization, even if a subject does not receive the correct treatment or does not follow the protocol until completion. The FAS will be used for all efficacy and baseline analyses unless otherwise specified in the SAP. Subjects who do not continue into the open-label Treatment Period will be excluded from the analyses in the open-label Treatment Period.

The following populations will be used for safety analyses:

- The Safety Analysis Set includes all randomized subjects who received at least 1 dose of study drug. Subjects will be analyzed according to the treatment they actually receive, defined as the randomized treatment if the patient took at least one dose of that randomized treatment during the 12-month placebo-controlled double-blind Treatment Period, or the first treatment received if the randomized treatment was never received during the 12-month placebo-controlled double-blind Treatment Period. Subjects who do not continue into the open-label Treatment Period will be excluded from the analyses in the open-label Treatment Period.
- The All Elagolix plus E2/NETA Analysis Set is defined as subjects who received at least 1 dose of elagolix plus E2/NETA in the study (either in the placebo-controlled or open-label Treatment Periods). Subjects will be summarized based on the actual treatment received at the time of randomization (elagolix plus E2/NETA or placebo).

5.0 Subject Disposition

A summary of subject accountability will be provided where the number of subjects in each of the following categories will be summarized for all randomized subjects by treatment group and overall:

- Subjects randomized in the study.

- Subjects who took at least one dose of study drug.
- Subjects who completed the placebo-controlled Treatment Period.
- Subjects who prematurely discontinued during the placebo-controlled Treatment Period.
- Subjects who entered into the open-label Treatment Period.
- Subjects who prematurely discontinued during the open-label Treatment Period.
- Subjects who completed protocol-specified treatment.
- Subjects who prematurely discontinued study drug (all reasons and primary reason).
- Subjects who entered into the Follow-up Period.
- Subjects who completed the Follow-up Period.

The number and percentage of subjects who discontinued study drug will be summarized by treatment group and overall as follows based on the Full Analysis Set:

- By any reason for discontinuation
- By primary reason for discontinuation

In addition, the number and percentage of subjects who discontinued study will be summarized for the FAS by treatment group and overall as follows:

- By any reason for discontinuation
- By primary reason for discontinuation

6.0 Study Drug Duration

Duration (days) of treatment will be summarized for the Safety Analysis Set using the number of subjects treated, mean, standard deviation, median, minimum and maximum.

Study drug duration will be summarized as follows:

Placebo-controlled Treatment Period (Safety Analysis Set):

Duration of treatment during the placebo-controlled Treatment Period is defined as follow:

- For subjects who did not continue into the open-label Treatment Period:
  - Last dose date during the placebo-controlled Treatment Period ‑ First dose date during the placebo-controlled Treatment Period + 1
- For subjects who continued into the open-label Treatment Period:
  Minimum of
  - Last dose date during the placebo-controlled Treatment Period ‑ First dose date during the placebo-controlled Treatment Period + 1
  - First dose date during the open-label Treatment Period ‑ First dose date during the placebo-controlled Treatment Period.

In addition, the number and percentage of subjects in each treatment duration interval (1 – < 28 days, 28 – < 56 days, 56 – < 84 days, 84 – < 112 days, 112 – < 140 days, 140 – < 168 days, 168 – < 196 days, 196 – < 224 days, 224 – < 252 days, 252 – < 280 days, 280 – < 308 days, 308 – < 336 days, ≥ 336 days) will be summarized.

Open-label Treatment Period (Safety Analysis Set):

Duration of treatment during the open-label Treatment Period is defined as last dose date during the open-label Treatment Period minus first dose date during the open-label Treatment Period plus 1 day.

In addition, the number and percentage of subjects in each treatment duration interval (1 – < 84 days, 84 – < 168 days, 168 – < 252 days, 252 – < 336 days, 336 – < 420 days, 420 – < 504 days, 504 – < 588 days, 588 – < 672 days, 672 – < 756 days, 756 – < 840 days, 840 – < 924 days, 924 – < 1008 days, ≥ 1008 days) will be summarized.

All Elagolix plus E2/NETA Analysis Set:

For subjects who received at least 1 dose of elagolix plus E2/NETA (either in the placebo-controlled or open-label Treatment Period), the duration of treatment is defined as the last dose date of receiving elagolix plus E2/NETA minus the first dose date of receiving elagolix plus E2/NETA plus 1 day.

In addition, the number and percentage of subjects in each treatment duration interval (1 – < 84 days, 84 – < 168 days, 168 – < 252 days, 252 – < 336 days, 336 – < 420 days, 420 – < 504 days, 504 – < 588 days, 588 – < 672 days, 672 – < 756 days, 756 – < 840 days, 840 – < 924 days, 924 – < 1008 days, 1008 – < 1092 days, 1092 – < 1176 days, 1176 – < 1260 days, 1260 – < 1344 days, ≥ 1344 days) will be summarized.

7.0 Demographics, Baseline Characteristics, Medical History, and Prior/Concomitant Medications

Demographics, baseline characteristics, medical history, prior medications, concomitant medications, and post-treatment medications will be summarized by treatment group and overall using the Full Analysis Set. Categorical variables will be summarized with the number and percentage of subjects; percentages will be calculated based on the number of non-missing observations. Continuous variables will be summarized with descriptive statistics (number of non-missing observations, mean and standard deviation, median, minimum and maximum).

7.1 Demographics and Baseline Characteristics

Continuous demographic variables include age, weight, height, and body mass index (BMI). Categorical demographic variables include sex, ethnicity, race, age (18 – 19, 20 – 24, 25 – 29, 30 – 34, 35 – 39, 40 – 44, 45 – 49, ≥ 50 years), BMI (≤ 18.5, > 18.5 – < 25, ≥ 25 – < 30, ≥ 30 – < 35, ≥ 35 – < 40, ≥ 40 kg/m^2), tobacco use (current, former, never, unknown), and alcohol use (current, former, never, unknown).

7.2 Medical History

Medical history and gynecological medical/surgical history data will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). The actual version of the MedDRA coding dictionary will be noted in the statistical tables and clinical study report. The number and percentage of subjects in each medical history category (by MedDRA system organ class and preferred term) will be summarized by treatment group and overall for the FAS. The system organ class (SOC) will be presented in alphabetical order, and the preferred terms will be presented in alphabetical order within each SOC. Subjects reporting more than one condition/diagnosis will be counted only once in each row (SOC or preferred term).

7.3 Prior, Concomitant, and Post-Treatment Medications

Prior, concomitant, and post-treatment medications will be summarized by generic name for the FAS.

Prior medications are those medications with a start date prior to the first study drug administration date. Prior medications will be summarized separately for each treatment group and overall using ATC Classification and preferred terms from the World Health Organization (WHO) Drug Dictionary.

Concomitant medications are those medications, other than study drugs, taken during the treatment period with an end date after the first dose of study drug or ongoing at the end of study, and a start date prior to the last dose of study drug. A medication will be considered a concomitant medication if any of the following three cases occur (1) the start date is missing and the end date is either after or on the first study drug dose date; (2) the start date is not missing and the end date is missing; (3) both the start date and the end date are missing.

Concomitant medications will be summarized using ATC Classification and preferred terms from the World Health Organization (WHO) Drug Dictionary with number and percentage for each treatment group and overall.

Post-treatment medications are those medications with an end date after the last dose of study drug or ongoing at the end of study. Post-treatment medications will be summarized by ATC Classification and WHO preferred term with number and percentage for each treatment group and overall.

A subject who reports two or more uses of the same medication will be counted only once within each WHO preferred term. A subject with medications with more than one generic name will be counted only once in the overall total. Prior medications, concomitant medications, and other medications during the Post-Treatment Follow-up Period will be summarized separately.

7.4 Physician Surgery Questionnaire (PSQ)

For each of the PSQ questions, the number and percentage of subjects in each response category will be summarized at Baseline by treatment group and overall. No statistical tests will be performed.

7.5 Reason for Study Participation Questionnaire

Descriptive statistics will be presented for reasons for study participation for each treatment group. No statistical tests will be performed.

8.0 Efficacy Analyses

8.1 General Considerations

All efficacy analyses will be conducted in the FAS Population.

Unless otherwise specified, during the 12-month placebo-controlled Treatment Period, elagolix plus E2/NETA group will be compared against placebo. Statistical tests will be conducted at an alpha level of 0.05 (two-sided). A test will be deemed significant if the P value rounded to three decimal places is less than or equal to 0.05 unless otherwise specified.

Efficacy endpoints collected during each of the placebo-controlled and open-label Treatment Periods will also be summarized as observed cases (by treatment group); no comparisons or statistical tests will be performed.

The End of 12-Month Placebo-Controlled Treatment Period analyses will be performed after all ongoing subjects have completed the 12-month placebo-controlled Treatment Period and the database has been locked. This will be the only and final analysis for the efficacy endpoints collected in the 12-Month Placebo-Controlled Treatment Period.

Unless otherwise specified, categorical data will be summarized by frequencies and percentages; continuous data will be summarized by the mean, standard deviation, median, minimum, and maximum. For the analyses of change from baseline, the within-group changes will be summarized with the mean, standard deviation or standard error, and 95% confidence intervals (CIs); between-group differences will be summarized with the mean, standard error, P value (as appropriate), and the 95% CIs.

For efficacy variables, the baseline during the 12-month placebo-controlled Treatment period (Months 1 – 12) will be the last non-missing value obtained prior to or on Study Day 1.

8.2 Handling of Missing Data

The primary objective of the study is to evaluate the long-term safety and tolerability of Elagolix plus E2/NETA. The efficacy analyses will be based on the observed data. The missing data will not be replaced or imputed. Potential impact due to COVID-19 will be recorded in the database. Listings will be provided to summarize AE and death related to COVID-19 infection, COVID-19 impacted visits, COVID-19 related study drug interruptions, COVID-19 status, and COVID-19 supplemental signs/symptoms.

8.3 Primary Efficacy Endpoint and Analyses

As the objective of this study is to evaluate the long-term safety and tolerability of elagolix 300 mg BID with E2/NETA, no primary efficacy endpoint is defined.

8.4 Secondary Efficacy Endpoint and Analyses

As the objective of this study is to evaluate the long-term safety and tolerability of elagolix 300 mg BID with E2/NETA, no secondary efficacy endpoint is defined.

8.5 Other Efficacy Endpoints and Analyses

Other efficacy assessments include the following:

- PGIC for Menstrual Bleeding Symptom;
- Change from baseline for the UFS-QoL.

8.5.1 Patient Global Impression of Change (PGIC)

For PGIC on menstrual bleeding (PGIC – MB), the number and percentage of subjects in each response category will be summarized at Month 3, Month 6, Month 9, Month 12 and Final Visit during the placebo-controlled double-blind 12-month Treatment Period by treatment group.

The response categories of "Very Much Improved" and "Much Improved" will be combined together. The remaining five categories of the PGIC scale will be combined and labeled "Otherwise." The number and percentage of subjects with response of (a) "Very Much Improved" or "Much Improved," and (b) "Otherwise" will be summarized at Month 3, Month 6, Month 9, Month 12 and Final Visit during the placebo-controlled Treatment Period by treatment group. Comparisons will be made using a Miettinen–Nurminen (M-N) test [1] for Month 3, Month 6, Month 9, Month 12, and Final Visit assessments in placebo-controlled Treatment Period. The p-value of the Miettinen-Nurminen test will be reported.

8.5.2 Uterine Fibroid Symptoms Quality of Life (UFS-QoL)

Improvement in quality of life will be assessed on the UFS-QoL Questionnaire (4-week recall). Each of the UFS-QoL subscales (symptom severity, concern, activities, energy/mood, control, self-conscious, and sexual function) and the HRQL total at

Month 6, Month 12, Month 18, Month 24, Month 30, Month 36, Month 42, Month 48, and Final Visit during the Treatment Period will be calculated and summarized by treatment group. The change from Baseline to Month 6, Month 12, and Final Visit during the placebo-controlled Treatment Period will be analyzed using ANCOVA with treatment as the main effect and corresponding Baseline UFS-QoL as a covariate to compare elagolix dose group to placebo.

9.0 Safety Analyses

9.1 General Considerations

Safety summaries will be provided separately using the Safety Analysis Set during the placebo-controlled Treatment Period, Safety Analysis Set during the open-label Treatment Period, and All Elagolix plus E2/NETA Analysis Set as defined in Section 4.0. Unless otherwise specified, no statistical comparisons or tests will be performed on Safety Analysis Set during the open-label Treatment Period or the All Elagolix plus E2/NETA Analysis Set.

Unless otherwise specified, the baseline for the 12-month placebo-controlled Treatment Period will be the last non-missing value obtained prior to or on Study Day 1. The baseline during the open-label Treatment Period will be the last non-missing value obtained prior to or on the initiation day of elagolix plus E2/NETA. The baseline for All Elagolix plus E2/NETA Analysis Set will be the last non-missing value obtained prior to or on the initiation day of elagolix plus E2/NETA.

All safety analyses will be based on observed data. Unless otherwise specified, missing data will not be imputed. The COVID-19 pandemic has interfered with the conduct of this study. Potential impacts due to COVID-19 are recorded in the database. Listings will be provided to summarize AE and death related to COVID-19 infection, COVID-19 impacted visits, COVID-19 related study drug interruptions, COVID-19 status, and COVID-19 supplemental signs/symptoms.

9.2 Adverse Events

Adverse events (AEs) will be summarized and presented using primary MedDRA System Organ Classes (SOCs) and preferred terms (PTs) according to the version of the MedDRA coding dictionary used for the study at the time of database lock. The actual version of the MedDRA coding dictionary used will be noted in the AE tables and in the clinical study report. Specific adverse events will be counted once for each subject for calculating percentages, unless stated otherwise. In addition, if the same adverse event occurs multiple times within a subject, the highest severity and level of relationship to investigational product will be reported.

9.2.1 Treatment-Emergent Adverse Events

Treatment-emergent adverse events (TEAEs) for each safety population are defined as follows:

TEAEs during the 12-month placebo-controlled Treatment Period on Safety Analysis Set during the placebo-controlled Treatment Period

TEAEs during the 12-month placebo-controlled Treatment Period are defined as any AEs with the onset that is on or after the first dose of study drug during the 12-month placebo-controlled Treatment Period and no more than 30 days after the last dose of study drug for subjects who premature discontinue during the 12-month double-blind Treatment Period or until the first dose of study drug in the open-label Treatment Period for subjects who enter the open-label Treatment Period. AEs starting more than 30 days following discontinuation of study drug will not be included in summaries of treatment-emergent AEs but will be summarized separately.

TEAEs during the open-label Treatment Period on the Safety Analysis Set during the open-label Treatment Period

TEAEs during the open-label Treatment Period are defined as any AEs with the onset that is on or after the first dose of study drug during the open-label Treatment Period and no

more than 30 days after the last dose of study drug in the study. AEs starting more than 30 days following discontinuation of study drug will not be included in summaries of treatment-emergent AEs but will be summarized separately.

TEAEs for All Elagolix plus E2/NETA Analysis Set

TEAEs for All Elagolix plus E2/NETA Analysis Set are defined as any AEs with the onset that is on or after the first dose of elagolix plus E2/NETA (either in the placebo-controlled Treatment Period or open-label Treatment Period) and no more than 30 days after last dose of elagolix plus E2/NETA. AEs starting more than 30 days following discontinuation of study drug will not be included in summaries of treatment-emergent AEs but will be summarized separately.

The number and percentage of subjects experiencing treatment-emergent AEs will be summarized overall, as well as by primary MedDRA SOC and PT by treatment group. The SOCs will be presented in alphabetical order, and the PTs will be presented in alphabetical order within each SOC.

9.2.2 Adverse Event Overview

An overview of treatment-emergent AEs will be presented consisting of the number and percentage of subjects experiencing at least one event for each of the following AE categories:

- Any treatment-emergent AE
- Any treatment-emergent AE related to study drug according to the investigator
- Any severe treatment-emergent AE
- Any serious treatment-emergent AE
- Any treatment-emergent AE leading to discontinuation of study drug
- Any treatment-emergent AE leading to death
- Any treatment-emergent AE of special interest as specified in Appendix A
- All deaths

9.2.3 Treatment-Emergent Adverse Events by SOC and PT

Treatment-emergent adverse events will be summarized by SOC and PT; by maximum relationship to study drug as assessed by the investigator (e.g., reasonable possibility or no reasonable possibility) and SOC and PT; by maximum severity and SOC and PT; and by subject number and SOC and PT. Specific adverse events will be counted once for each subject for calculating percentages, unless stated otherwise. In addition, if the same adverse event occurs multiple times within a subject, the highest severity and level of relationship to investigational product will be reported.

In addition, treatment-emergent adverse events will be summarized by PT and sorted by decreasing frequency for the elagolix plus E2/NETA group.

9.2.4 Treatment-Emergent Adverse Events per Patient-Years of Exposure

An overview of AEs for All Elagolix plus E2/NETA Analysis Set will be presented per 100 patient-years. In addition, TEAEs for All Elagolix plus E2/NETA Analysis Set will be presented per 100 patient-years by SOC and PT. AEs per 100 patient-years of exposure are defined as the number of AEs divided by the total exposure in 100 patient-years. Note that one event per preferred term per day per subject will be counted in the calculation of the number of AEs (i.e., a preferred term will not be counted twice on the same day for the same subject).

$$\text{TEAEs per 100 patient-years of exposure} = 100 \times \frac{\text{Number of TEAEs}}{\text{Total Patient Years}}$$

where total patient years are defined as the sum of the study drug exposure (defined as date of last dose of elagolix plus E2/NETA – date of first dose of elagolix plus E2/NETA + 30 days) of all subjects normalized by 365.25, and rounded to one decimal place.

9.2.5 Serious Adverse Events (Including Deaths) and Adverse Events Leading to Study Drug Discontinuation

The number and percentage of subjects experiencing the following categories will be tabulated according to the primary MedDRA SOC and PT. In addition, listing of subjects experiencing the following categories will be generated.

- Treatment-emergent SAEs
- Treatment-emergent AEs leading to discontinuation of study drug
- Treatment-emergent AEs leading to death

9.2.6 Adverse Events of Special Interest

Adverse events of special interest will be summarized by SOC and PT and will be based on standardized or company MedDRA queries (SMQs or CMQs), or based on adjudication results.

Detailed information about the search criteria are provided in Appendix A.

9.2.7 Post-Treatment Adverse Events

Post-Treatment Adverse Events are defined as AEs starting more than 30 days following discontinuation of study drug in the Treatment Period. The Post-Treatment AEs will be summarized as follows:

- Overview of Post-Treatment AEs
- Post-Treatment AEs by SOC and PT
- Post-Treatment SAEs by SOC and PT

In addition, listing of subjects with Post-Treatment AEs by SOC and PT and Post-Treatment SAEs by SOC and PT will be provided.

9.3 Analysis of Laboratory Data

Data collected from central and local laboratories, including additional laboratory testing due to an SAE, will be used in all analyses, except for Baseline where SAE-related laboratory assessments on or before the first dose of study drug will be excluded. The clinical laboratory tests defined in the protocol operations manual (i.e., hematology and clinical chemistry) will be summarized.

Each laboratory variable will be summarized for all time points (starting with Baseline) with the number of non-missing observations, mean and standard deviation, median, minimum and maximum. For selected laboratory variables including lipid variables, liver variables (alkaline phosphatase, ALT, AST, bilirubin), and hemotological variables (hemoglobin, hematocrit, platelet count, and red blood cell (RBC) count), mean change from baseline to each applicable post-baseline visit will be summarized, with the number of observations, baseline mean, and visit mean. The change from baseline mean, standard error, and 95% confidence interval will be presented for the mean change from baseline within each treatment group and difference between elagolix group and placebo. For lipid variables, in addition to low-density lipoprotein cholesterol (LDL-C), high-density lipoprotein cholesterol (HDL-C), total cholesterol, triglycerides (TG) and apolipoprotein A and B, the following ratios will be included: the ratio of total cholesterol to HDL-C, the ratio of LDL-C to HDL-C, the ratio of TG to HDL-C, and the ratio of non-HDL-C (calculated as total cholesterol minus HDL-C) to HDL-C.

Changes in selected laboratory parameters will be tabulated using shift tables by NCI CTC criteria and categorized as low, normal, or high based on the normal ranges of the laboratory used for each sample. A shift table from baseline either to the worse value (based on NCI CTC criteria) during treatment or to minimum and maximum value (based on normal range), will be created. A similar shift table will be provided to summarize shifts from baseline to the final post-baseline value.

Laboratory abnormalities meeting CTC criteria grade 3 or 4 will be summarized.

The number and percentage of subjects meeting the following sponsor-defined potentially clinically significant lipid values will be summarized at baseline and each relevant visit by treatment group in the appropriate Treatment Period:

- Total cholesterol: ≤ 300, > 300 – ≤ 400, > 400 – ≤ 500, and > 500 mg/dL
- HDL-C: < 40 and ≥ 40 mg/dL
- LDL-C: < 130, ≥ 130 – < 160, ≥ 160 – < 190, and ≥ 190 mg/dL
- TG: ≤ 150, > 150 – ≤ 300, > 300 – ≤ 500, > 500 – ≤ 1000, and > 1000 mg/dL
- LDL-C/HDL-C ratio: ≤ 3, and > 3
- Total cholesterol/HDL-C ratio: ≤ 4.5, and > 4.5

In addition, the number and percentage of subjects who have potentially clinically significant (PCS) lipid values meeting the following criteria any time during the appropriate Treatment Period will be summarized by treatment group. Listings will be provided to summarize subject-level laboratory data for subjects meeting PCS criteria.

- Total cholesterol > 200 mg/dL
- LDL-C > 130 mg/dL
- LDL-C > 160 mg/dL
- HDL-C < 40 mg/dL
- TG > 150 mg/dL
- TG/HDL-C ratio > 3.5
- LDL-C/HDL-C ratio > 4

The following plots will be provided by treatment group for HDL-C, LDL-C, and triglycerides during the appropriate Treatment Period:

- Final post-baseline lab values vs. Baseline lab values
- Final post-baseline lab values vs. Baseline lab values for subjects with NCI CTC Grade 3 or 4
- Worst post-baseline lab values vs. Baseline lab values

- Worst post-baseline lab values vs. Baseline lab values for subjects with NCI CTC Grade 3 or 4
- Mean percent change from baseline in lipid parameters over time

In addition, plots of mean lab values over time during the Treatment Period will be provided for hemoglobin, total cholesterol, HDL-C, LDL-C, triglycerides, LDL-C/HDL-C ratio, ALT, AST, bilirubin, apolipoprotein A and apolipoprotein B.

The number and percentage of subjects in each treatment group with maximum on treatment laboratory values meeting the following criteria compared to the upper limit of normal (ULN) during the appropriate Treatment Period will be summarized to assess potential hepatotoxicity.

- ALT $\geq 3 \times$ ULN and total bilirubin $\geq 2 \times$ ULN
- AST $\geq 3 \times$ ULN and total bilirubin $\geq 2 \times$ ULN
- ALT and AST $\geq 3 \times$ ULN and total bilirubin $\geq 2 \times$ ULN
- ALT $\geq 3 \times$ ULN and total bilirubin $< 2 \times$ ULN
- AST $\geq 3 \times$ ULN and total bilirubin $< 2 \times$ ULN
- ALT and AST $\geq 3 \times$ ULN and total bilirubin $< 2 \times$ ULN
- ALT $\geq 3 \times$ ULN and total bilirubin $\geq 1.5 \times$ ULN
- AST $\geq 3 \times$ ULN and total bilirubin $\geq 1.5 \times$ ULN
- ALT and AST $\geq 3 \times$ ULN and total bilirubin $\geq 1.5 \times$ ULN
- ALT $\geq 3 \times$ ULN, $\geq 5 \times$ ULN, $\geq 10 \times$ ULN, $\geq 20 \times$ ULN
- AST $\geq 3 \times$ ULN, $\geq 5 \times$ ULN, $\geq 10 \times$ ULN, $\geq 20 \times$ ULN
- Total bilirubin $\geq 1.5 \times$ ULN, $\geq 2.0 \times$ ULN

The maximum ratio relative to the ULN is used to determine if subjects met the criteria listed above. The ALT, AST, and total bilirubin values do not need to be concurrent in order to meet the defined criteria. For ALT, AST, and total bilirubin, a subject is counted if the post-baseline laboratory value during the Treatment Period meets the above criteria

regardless of Baseline laboratory value (i.e., the post-baseline laboratory value does not need to be worse than Baseline laboratory value).

A listing of all ALT, AST, total bilirubin, and alkaline phosphatase values will be provided for each subject who meets any of the criteria defined above.

The following plots will be generated: peak AST vs. peak bilirubin; peak ALT vs. peak bilirubin.

9.4 Analysis of Vital Signs and Weights

Vital sign measurements of pulse rate, sitting systolic blood pressure, and sitting diastolic blood pressure will be summarized.

Each vital sign variable and weight will be summarized for all time points (starting with Baseline) with the number of non-missing observations, mean and standard deviation, median, minimum and maximum. Mean change from baseline to each applicable post-baseline visit will be summarized for each vital sign variable, with the number of observations, baseline mean, and visit mean. The change from baseline mean, standard error, and 95% confidence interval will be presented for the mean change from baseline within each treatment group and difference between elagolix group and placebo.

The number and percentage of subjects who have potentially clinically significant (PCS) vital sign and weight values meeting the following criteria will be summarized by treatment group. All increase/decrease is calculated from Baseline to a post-baseline visit in the Treatment Period.

- Diastolic blood pressure
  - ≤ 50 mmHg and ≥ 15 mmHg decrease
  - > 90 mmHg and ≥ 15 mmHg increase
  - ≥ 100 mmHg
- Systolic blood pressure
  - ≤ 90 mmHg and ≥ 20 mmHg decrease

- ≥ 140 mmHg and ≥ 20 mmHg increase
- ≥ 160 mmHg

- Pulse rate
  - ≤ 45 bpm and ≥ 15 bpm decrease
  - > 100 bpm and ≥ 15 bpm increase
  - ≥ 120 bpm
- Weight
  - ≥ 5% decrease
  - ≥ 7% increase.

Listings will be provided to summarize subject-level vital sign and weight data for subjects meeting PCS criteria.

The number and percentage of subjects who had a sustained PCS vital sign and weight values will be summarized by treatment group and a listing of these subjects will be provided. A sustained PCS value is defined as 3 consecutive PCS values.

9.5 Other Safety Analyses

9.5.1 Columbia Suicide Severity Rating Scale (C-SSRS)

The C-SSRS data will be summarized as observed by treatment group for the Safety Analysis Set.

The C-SSRS – Baseline/Screening measured at Day 1 will be considered as Baseline C-SSRS. Baseline C-SSRS will be summarized. Other analysis of C-SSRS will only include subjects who have at least one post-baseline C-SSRS measurement, regardless of whether they had a baseline C-SSRS measurement. The number of subjects with suicidal ideation, suicidal behavior, and self-injurious behavior without suicidal intent at each scheduled assessment during the appropriate Treatment Period will be summarized. In addition, this table will be repeated for providing a summary of lifetime outcomes and past year outcomes at screening and Day 1. No statistical test will be performed.

The number of subjects with suicide-related treatment-emergent events based on the C-SSRS during the appropriate Treatment Period will be summarized. No statistical tests will be performed.

A listing of subjects with suicidal ideation, suicidal behavior, or self-injurious behavior without suicidal intent based on the C-SSRS during treatment period will be provided.

9.5.2 Bone Mineral Density

All analyses and summaries of bone mineral density (BMD) will be performed for each anatomic location, i.e., femoral neck, lumbar spine, and total hip. For subjects who had a right-side scan performed (rare instances), their data for right femoral neck and right femoral total hip will be included in the analysis with the data for the left femoral neck and left total hip, respectively (available for the majority of subjects). If more than one scan is reported for an anatomic location within an analysis window, the worse (the lower value) of the multiple measurements will be used for analysis for each anatomic location. Unless otherwise specified, the analysis of BMD at each post-baseline visit will exclude subjects who switch machine manufacturer type (Lunar or Hologic).

9.5.2.1 Bone Mineral Density during the Treatment Period

BMD

A summary of continuous BMD at each scheduled assessment will be provided during the appropriate Treatment Period. This summary will include the sample size, mean, SD, median, minimum, and maximum. This analysis will be repeated including subjects who switch machine manufacturer type (Lunar or Hologic), for subjects with scans only on Lunar machines, and for subjects with scans only on Hologic machines.

The analysis of percent change in BMD from baseline to each relevant visit during the Treatment Period will be based on an ANCOVA model with treatment as the main effect and baseline value of corresponding parameter as a covariate. Elagolix plus E2/NETA group will be compared to placebo at Months 6 and 12 during the placebo-controlled

Treatment Period. This analysis will be repeated including subjects who switch machine manufacturer type (Lunar or Hologic), for subjects with scans only on Lunar machines, and for subjects with scans only on Hologic machines.

The number and percentage of subjects with percent change from Baseline to each scheduled assessment during the Treatment Period in mutually exclusive categories of BMD decrease: ≤ 0%, > 0% – ≤ 1.5%, > 1.5% – ≤ 3%, > 3% – ≤ 4%, > 4% – ≤ 5%, > 5% – < 8%, and ≥ 8% will be summarized by treatment group. Elagolix plus E2/NETA group will be compared to placebo at Months 6 and 12 during the placebo-controlled Treatment Period using a chi-squared test.

Cross-tabulation of the DXA scans performed versus Treatment Period completion will be provided after all subjects complete the Treatment Period or premature discontinue from the Treatment Period.

Listings of subjects meeting the following criteria during the Treatment Period will be provided:

- Listing of subjects with BMD decrease ≥ 8%

In addition, the following plot will be provided as needed:

- Percent change from Baseline to each scheduled assessment (Months 6, 12, 18, 24, 30, 36, 42, 48) during the Treatment Period in BMD vs. Baseline BMD values for subjects with greater than 8% BMD decrease at any location during the Treatment Period.
- Cumulative distribution function of percent change from baseline in BMD at each scheduled assessment (Months 6, 12, 18, 24, 30, 36, 42, 48) during the Treatment Period.
- Percent change from Baseline to each scheduled assessment (Months 6, 12, 18, 24, 30, 36, 42, 48) in BMD during the Treatment period.

Z-Score and T-Score

A continuous summary of the Z-score and T-score at each scheduled assessment will be provided by treatment group during the Treatment Period. This summary will include the sample size, mean, SD, median, minimum, and maximum.

A categorical summary of Z-score will be provided at each scheduled assessment during the Treatment Period for the following categories: ≤ -2.0, > -2.0 to ≤ -1.5, > -1.5 to ≤ -1.0, and > -1.0. A categorical summary of T-score will be provided at each scheduled assessment during the Treatment Period for the following categories: ≤ -2.5, > -2.5 to < -1.0, and ≥ -1.0. In addition, a categorical summary of the worst Z-score and T-score during the Treatment Period for the above categories will be provided.

9.5.2.2 Bone Mineral Density during the Follow-up Period

BMD

A summary of continuous BMD at each scheduled assessment will be provided by treatment group during the Follow-up Period. This summary will include the sample size, mean, SD, median, minimum, and maximum. This analysis will be repeated including subjects who switch machine manufacturer type (Lunar or Hologic), for subjects with scans only on Lunar machines, and for subjects with scans only on Hologic machines.

The analysis of percent change in BMD from baseline to each relevant visit during the Follow-up Period will be based on an ANCOVA model with treatment as the main effect and baseline value of corresponding parameter as a covariate. No comparison will be performed during the Follow-up Period. This analysis will be repeated including subjects who switch machine manufacturer type (Lunar or Hologic), for subjects with scans only on Lunar machines, and for subjects with scans only on Hologic machines.

The analysis of percent change in BMD from baseline will be repeated for the following subset of subjects:

- subjects who have a Treatment Month X (e.g., Month 12) scan and a Follow-up Month Y (e.g., Follow-up Month 6) scan and premature discontinue prior to next scheduled assessments.

The number and percentage of subjects with percent change from Baseline to each scheduled assessment during the Follow-up Period in mutually exclusive categories of BMD decrease: ≤ 0%, > 0% – ≤ 1.5%, > 1.5% – ≤ 3%, > 3% – ≤ 4%, > 4% – ≤ 5%, > 5% – < 8%, and ≥ 8% will be summarized by treatment group.

In addition, the number and percentage of subjects with percent change from final on-treatment to each scheduled assessment during the Follow-up Period in mutually exclusive categories of BMD decrease: ≤ 0%, > 0% – ≤ 1.5%, > 1.5% – ≤ 3%, > 3% – ≤ 4%, > 4% – ≤ 5%, > 5% – < 8%, and ≥ 8% will be summarized by treatment group.

For the purpose of assessing post-treatment BMD recovery, the following summaries will be provided. The "recovery" statistic was defined as:

Recovery at Follow-up Month X = 100 × ((% change from baseline to final) – (% change from baseline to Follow-up Month X))/(% change from baseline to final)

The "recovery" statistic is the proportion of BMD decrease at the final treatment scan recovered at Follow-up Month X. For example, a subject who has a change of –2% at the end of treatment and has a change of –1% at Follow-up Month 6 has a recovery value of 50% at Follow-up Month 6. This statistic will only be defined for subjects who experienced a decrease from baseline at final treatment.

A continuous summary of recovery at each relevant visit during the Follow-up Period will be provided including mean, standard deviation, median, and a within-group 95% confidence interval for the following subset of subjects:

- All subjects who experienced a decrease from baseline at final treatment;
- Subjects who premature discontinue after Month 6 and on or prior to Month 12;

- Subjects who premature discontinue after Month 12 and on or prior to Month 18;
- Subjects who premature discontinue after Month 18 and on or prior to Month 24;
- Subjects who premature discontinue after Month 24 and on or prior to Month 30;
- Subjects who premature discontinue after Month 30 and on or prior to Month 36;
- Subjects who premature discontinue after Month 36 and on or prior to Month 42;
- Subjects who premature discontinue after Month 42 or completed the Treatment Period.

Additionally, the number and percentage of subjects in each of the following categories will be provided for the subset of subjects listed above: < 0%, 0 – 25%, > 25 – 50%, > 50 – 75%, > 75 – 100%, > 100%.

Bone loss will be defined as > 3% at the spine, > 4% at the total hip, or > 5% at the femoral neck. The following listings will be provided:

- Listing of subjects with bone loss from baseline to final on-treatment visit and continued bone loss from final on-treatment to Follow-up visits will be provided. Subjects subsequently with an increase from Follow-up Month 6 to a subsequent Follow-up assessment will be excluded from this listing.

In addition, the following plots will be provided:

- Cumulative distribution function of recovery at Follow-up Month 6 and Follow-up Month 12.
- Percent change from baseline to treatment and follow-up months of interest.

9.5.3 Transvaginal Ultrasound

9.5.3.1 Ovarian Cysts

The presence of ovarian cysts will be noted at Baseline and each relevant post-baseline visit during the placebo-controlled and open-label Treatment Periods. The ovarian findings included complex ovarian cyst > 3.5 cm, simple ovarian cyst > 5 cm or endometriomas > 3 cm. For the Safety Analysis Set, the number and percentage of subjects with the pre-defined complex ovarian cysts and simple ovarian cysts will be summarized by treatment group at each relevant visit during the Treatment Period.

For each TAU/TVU assessment, information for ovarian findings from the cysts assessment may be available for more than one cyst at more than one ovary location (left and/or right). The subject with multiple cyst findings was counted once in the numerator and denominator when reporting the number and percentage of subjects in the relevant category of ovarian findings. A listing included all results across multiple cysts findings from multiple assessments (where available) from subjects who had ovarian findings.

9.5.3.2 Endometrial Thickness

For the Safety Analysis Set, the endometrial thickness at each relevant visit during the Treatment Period will be summarized with descriptive statistics. Additionally, analysis of change in endometrial thickness from Baseline to each relevant visit during the Treatment Period will be performed. The mean change from Baseline each relevant visit during the 12-month placebo-controlled Treatment Period will be compared between the elagolix treatment group and placebo using one-way ANOVA with treatment as the main effect; no comparison will be conducted during the open-label Treatment period.

In addition, the number and percentage of subjects with endometrial thickness of < 8 mm, ≥ 8 mm and ≤ 12 mm, > 12 mm and ≤ 18 mm, and > 18 mm will be summarized by treatment group at each relevant visit during the Treatment Period. Elagolix group will be compared to placebo using a Chi-squared test during the 12-month placebo-controlled

Treatment Period. No comparisons will be conducted during the open-label Treatment Period.

9.5.4 Endometrial Biopsy

For the Safety Analysis Set, the number and percentage of subjects in each category of endometrial biopsy results will be summarized by treatment group at each relevant visit during the placebo-controlled and open-label Treatment Periods. If multiple assessments existed in a specific time window, all assessments will be displayed.

9.5.5 Pregnancy Results

For the Safety Analysis Set, pregnancies and outcomes will be summarized by treatment group for the Treatment Period (placebo-controlled and open-label) and Follow-up Period. Subjects with conception date by sponsor more than 30 days after the last dose of study drug will be included in the summaries for the Follow-up Period.

Listings will be prepared of all pregnancy test results for any subject who ever had a positive pregnancy test at any time point during the study.

Positive pregnancy test results (serum, urine, and both) and a listing of all positive pregnancy test results will be provided.

Pregnancy outcome will be summarized by treatment group. Additionally, a listing of pregnancy information will be provided. Pregnancy outcome includes:

- Number of pregnancies
- Live births
  - Pre-term (extremely pre-term: < 28 weeks; very pre-term: 28 - < 32 weeks; moderate to late preterm: 32 - < 37 weeks)
  - Term (37 - < 42 weeks)
  - Post-term (≥ 42 weeks)
- Spontaneous abortion (< 6, 6 - 13, > 13 weeks gestation)
- Termination of pregnancy

- Ectopic pregnancy
- Still birth
- Congenital anomaly
- Lost to follow-up
- Subjects refused to provide information
- Other
- Ongoing (pregnancy ongoing at end of study)
- Maternal exposure to Elagolix/Study drug (0, > 0 ‑ 4, > 4 ‑ 6, > 6 ‑ 12, > 12 ‑ 23, > 23 weeks)
- Embryo/Fetal exposure to Elagolix/Study drug assessed by the sponsor (0, > 0 ‑ 4, > 4 ‑ 6, > 6 ‑ 12, > 12 weeks)

Annualized pregnancy rate during treatment will be reported by treatment group. The annualized pregnancy rate will be calculated as

Annualized pregnancy rate = (number of pregnancies/days of exposure to treatment) × 365

In addition, listings of fetal/newborn, infant post-delivery assessments, and infant during 6 – 12 month follow-up assessment will be provided.

The above information will be summarized primarily based on data collected on the pregnancy case report form. Information not directly obtained from the case report form will use the following definitions.

Gestation day

Gestation Day 0 was calculated by subtracting the gestational age provided according to the first trimester ultrasound date from the first trimester ultrasound date. If no gestational age was provided, it will be imputed by subtracting 14 days from the midpoint between the subject's last negative pregnancy test date and the date of the positive pregnancy test.

Gestational age

Gestational age at delivery was calculated from gestation Day 0 to the date of delivery. Gestational age at abortion was calculated from gestation Day 0 to the date of abortion.

Conception date

Conception date by sponsor was calculated from the gestational age provided from the first trimester ultrasound, by adding 14 days to gestation Day 0. If no gestational age was provided, the conception date by sponsor was imputed as the midpoint between the subject's last negative pregnancy test date and the date of the positive pregnancy test.

Maternal exposure

Maternal exposure to study drug was calculated as the last dose date of study drug – the first dose date of study drug + 1.

Embryo/fetal exposure

Embryo/fetal exposure to study drug by sponsor was calculated from the conception date by site/sponsor to the date of maternal last dose of study drug.

10.0 Overall Type-I Error Control

The objective of this study is to evaluate the long-term safety and tolerability of elagolix 300 mg BID with E2/NETA. No primary or secondary efficacy endpoint is defined.

Since the End of 12-Month Placebo-Controlled Treatment Period analysis is the only and final analysis for the comparison between treatment and placebo group at 12 months of safety and efficacy endpoints, no additional adjustment of alpha-level is necessary.

11.0 Version History

Table 1. SAP Version History Summary

Version	Date	Summary
1.0	20 January 2021	Original version

12.0 References

1. Miettinen O, Nurminen M. Comparative analysis of two rates. Stat Med. 1985;4(2):213-26.

Appendix A. Definition of Adverse Events of Special Interest (AESI) Adverse Events of Special Interest

Item of Safety Interest	Method of Surveillance
Hot flashes/night sweats	Non-bone related hypoestrogenic effects CMQ
Bone mineral density loss	Osteoporosis/Osteopenia SMQ DXA results from clinical trials
Anemia	Cases are identified through the Non-Hemolytic and Non-Aplastic Anemias CMQ Haematopoietic erythropenia SMQ
Bone Fractures	Osteoporosis/osteopenia SMQ Bone fracture CMQ
Rash and hypersensitivity reactions	Anaphylactic reaction SMQ Severe cutaneous adverse reactions SMQ Drug induced rash CMQ
Lipid abnormalities	Dyslipidemia SMQ
Uterine bleeding change	Female reproductive bleeds CMQ
Endometrial safety	Uterine and fallopian tube neoplasms, malignant and unspecified SMQ Reproductive Premalignant Disorders SMQ Endometrial biopsy results
Hypoestrogenic AEs (excluding hot flashes, BMD loss, and fractures)	Non-bone related hypoestrogenic effects CMQ
Spontaneous abortion	Termination of pregnancy and risk of abortion SMQ
Teratogenicity	Review of pregnancy outcome All pregnancies will be followed up to 6 to 12 months post-delivery and reviewed at least quarterly and as they occur
Obstetrical complications (maternal and infant)	Pregnancy, labor, and delivery complications and risk factors (excluding abortions and stillbirths) SMQ
Psychiatric events	Depression and suicide/self-injury SMQ
Cardiovascular events	Cardiac arrhythmias SMQ, Cardiomyopathy SMQ, and Ischemic heart disease SMQ
Thromboembolic events	Embolic and thrombotic events SMQ
Alopecia	Alopecia, diffuse alopecia, androgenic alopecia
Hepatic events	Drug related hepatic disorders comprehensive SMQ, Enhanced pharmacovigilance hepatic terms PMQ
Hormonally mediated malignancies	Breast neoplasms, malignant and unspecified SMQ Uterine and fallopian tube neoplasms, malignant and unspecified SMQ Ovarian neoplasms, malignant and unspecified SMQ