STATISTICAL ANALYSIS PLAN

Study Number:	**MVT-601-050**
Study Title:	A Phase 3, Single-Arm, Open-Label Study to Evaluate the Safety and Contraceptive Efficacy of Relugolix Combination Therapy in Women with Uterine Fibroids or Endometriosis Who Are 18 to 50 Years of Age and at Risk for Pregnancy
Study Drug	Relugolix+E2/NETA
Indication	Contraception
Protocol Version: Date	Original: 12AUG2020 Amendment 1: 21DEC2020 Amendment 2: 07JUL2021 Amendment 3: 15SEP2021 Amendment 4: 21DEC2022
SAP Version/Date	Original SAP: 22MAR2024 SAP Amendment 1: 18AUG2025
Author	[redacted]

Sumitomo Pharma America, Inc.
84 Waterford Drive
Marlborough, MA 01752
USA

SIGNATURE PAGE

Author:

[redacted]

Signing Reason: I am the author of this document
Signing Time: 2025/08/18 | 13:42 JST
5A9BFCD607CC4C46B53F46AE41FB6B5C

2025/08/18 | 13:42 JST

Date

[redacted]

Reviewed and approved by:

[redacted]

Signing Reason: I approve this document
Signing Time: 2025/08/18 | 08:57 EDT
44DD1738BFFD481C8A2E5A6FCF993C7A

2025/08/18 | 08:57 EDT

Date

[redacted]

Signing Reason: I approve this document
Signing Time: 2025/08/18 | 15:29 JST
F2B425AE147D438FAF5519F081C9CC5E

2025/08/18 | 15:29 JST

Date

[redacted]

Signing Reason: I approve this document
Signing Time: 2025/08/18 | 07:56 EDT
80032B425FFA4FD9B4056764D5D1F83D

2025/08/18 | 07:56 EDT

Date

[redacted]

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Term	Description
AE	adverse event
ALP	alkaline phosphatase
ALT	alanine aminotransferase
AST	aspartate aminotransferase
ATC	Anatomical Therapeutic Chemical
BMD	bone mineral density
BMI	body mass index
C-SSRS	Columbia Suicide Severity Rating Scale
CI	confidence interval
CSR	clinical study report
CTCAE	Common Terminology Criteria for Adverse Events
CUSUM	cumulative summing
DSM-5	Diagnostic and Statistical Manual of Mental Disorders 5
DSMB	Data Safety Monitoring Board
DXA	dual-energy x-ray absorptiometry
E2	estradiol
ECD	estimated conception date
eCRF	electronic case report form
EDD	estimated date of delivery
eDiary	electronic diary
EOT	end-of-treatment
ET	early termination
FDA	Food and Drug Administration
HbA1c	hemoglobin A1c
HDL	high-density lipoprotein
ICH	International Conference on Harmonisation
IPD	important protocol deviation
ITT	intent-to-treat
IUD	intrauterine device
IWRS	Interactive Web Response System
KM	Kaplan-Meier
LDL	low-density lipoprotein
MedDRA	Medical Dictionary for Regulatory Activities
mITT	modified intent-to-treat
NA	not applicable
NETA	norethindrone acetate
NDA	New Drug Application
PHQ-9	Patient Health Questionnaire-9
PI	Pearl Index

Term	Description
PP	per-protocol
PT	preferred term
PTFU	post-treatment follow-up
QD	once daily
rITT	restricted intent-to-treat
SAP	statistical analysis plan
SAS	statistical analysis system
SD	standard deviation
SI	International System of Units
SOC	system organ class
ULN	upper limit of normal
WHO	World Health Organization

1. INTRODUCTION

This statistical analysis plan (SAP) describes the statistical analyses to be conducted for study protocol MVT-601-050 Amendment 4, dated 21DEC2022, entitled "A Phase 3, Single-Arm, Open-Label Study to Evaluate the Safety and Contraceptive Efficacy of Relugolix Combination Therapy in Women with Uterine Fibroids or Endometriosis Who Are 18 to 50 Years of Age and at Risk for Pregnancy".

This SAP will provide details to further elaborate statistical methods as outlined in the protocol MVT-601-050 and will describe analysis conventions to guide the statistical programming work. The SAP readers should consult the study protocol for more background information on the study.

This SAP is developed in accordance with the International Council on Harmonisation (ICH) guidelines:

- ICH guidelines E3 (Structure and Content of Clinical Study Reports) and
- ICH E9 (Statistical Principles for Clinical Trials).

All decisions regarding statistical analyses of the study, as defined in this SAP, will be made prior to the database lock. The SAP is to be finalized and approved by the sponsor before the lock of study database. Any changes to the final SAP (or SAP amendments, if applicable) after database lock will be clarified in the clinical study report (CSR) with "change from the planned analyses".

An independent Data and Safety Monitoring Board (DSMB) consisting of experts in women's health, clinical study safety monitoring, and statistics was established to evaluate the safety of study participants on an ongoing basis. A separate DSMB SAP outlines the specific safety data analyses that are performed for the DSMB on an ongoing basis during the study. Analyses will be performed using SAS® version 9.4 or higher (SAS Institute, Inc., Cary, NC 27513).

2. STUDY OBJECTIVES AND DESIGN

2.1. Study Objectives and Endpoints

The study objectives and corresponding endpoints are provided in Table 1.

Table 1: Study Objectives and Endpoints

Objectives	Endpoints
Primary Efficacy	
To assess the contraceptive efficacy of relugolix combination therapy in women with uterine fibroids or endometriosis who are 18 to 50 years of age and at risk for pregnancy, as expressed by the At-Risk PI in the rITT population, based on the number of on-treatment pregnancies occurring during cycles without use of any contraceptive method other than relugolix contraceptive therapy and with confirmed vaginal intercourse.	The contraceptive efficacy of relugolix combination therapy will be assessed using the At-Risk PI, defined as the number of on-treatment pregnancies per 100 woman-years of treatment: $PI = \frac{1300 \times (\text{number of on-treatment pregnancies})}{\text{Number of 28-day at-risk cycles of treatment}}$ On-treatment pregnancies are pregnancies with an estimated conception date between the first day of study intervention intake up to and including 7 days after the last intake of study medication. The At-Risk PI will be calculated on the basis of cycles considered at-risk of pregnancy, ie, consecutive 28-day periods without use of any other contraceptive methods and with affirmed occurrence of vaginal intercourse.
Secondary Efficacy	
To assess the following for relugolix combination therapy in women with uterine fibroids or endometriosis who are 18 to 50 years of age and at risk for pregnancy:	
Contraceptive efficacy, as expressed by the modified At-Risk PI, based on the number of on-treatment pregnancies occurring during cycles without use of any contraceptive method other than relugolix contraceptive therapy, regardless of occurrence of vaginal intercourse.	Modified At-Risk PI: PI based on the number of on-treatment pregnancies occurring during cycles without any other contraceptive methods regardless of occurrence of vaginal intercourse.
"Typical use" contraceptive efficacy, as expressed by the Gross PI, based on the number of on-treatment pregnancies occurring during all cycles in the ITT population, regardless of use of other contraceptive methods, confirmed vaginal intercourse, or protocol compliance.	Gross PI: PI based on the number of on-treatment pregnancies occurring during all cycles regardless of the use of other contraceptive methods, confirmed vaginal intercourse, or protocol compliance.

Table 1: Study Objectives and Endpoints (Continued)

Objectives	Endpoints
"Perfect use" contraceptive efficacy, as expressed by the Method Failure PI, based on the number of on-treatment pregnancies in the PP population.	Method Failure PI: PI based on the number of on-treatment pregnancies occurring during cycles that are at risk and without selected important protocol deviations.
Contraceptive efficacy in various populations and analysis sets.	Cumulative 1-year pregnancy rates.
Safety	
To describe the safety of relugolix combination therapy.	Incidence of treatment-emergent adverse events, change in vital signs (including body weight), and clinical laboratory tests.
To evaluate change in bone mineral density during treatment with relugolix combination therapy.	Percent change in bone mineral density from baseline to 6 and 12 months on-treatment at lumbar spine (L1-L4), total hip, and femoral neck.
To evaluate post-treatment change in bone mineral density.	Percent change from baseline and 12 months on-treatment (or Early Termination) to the 6-month and 12-month post-treatment time points at lumbar spine (L1-L4), total hip, and femoral neck.
To estimate discontinuation rate.	Proportion of enrolled patients who do not complete 13 treatment cycles.
To summarize bleeding profile	Number of bleeding/spotting days, bleeding days by bleeding intensity category, and proportion of patients with sustained absence of bleeding and spotting by visit Proportion of patients who discontinue the study drug due to bleeding

Abbreviations: ITT = intent-to-treat (Analysis Set); PI = Pearl Index; PP = per protocol (Analysis Set); rITT = restricted intent-to-treat (Analysis Set).

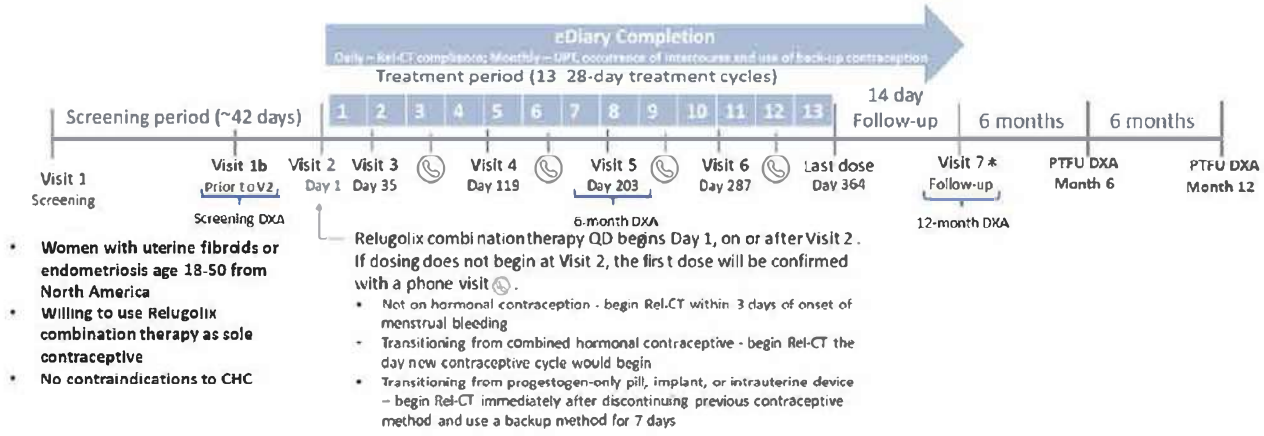
2.2. Study Design

Study MVT-601-050 is a single-arm, open-label, phase 3 study to assess the contraceptive efficacy of relugolix combination therapy (relugolix 40 mg, estradiol [E2] 1 mg, and norethindrone acetate [NETA] 0.5 mg). Approximately 1020 sexually active women with uterine fibroids or endometriosis who are 18 to 50 years of age and will not be using any other form of contraception will receive relugolix combination therapy orally once daily (QD) for up to 1 year (i.e., 13 consecutive 28-day treatment cycles) and will be assessed for the occurrence of pregnancy.

The study period consists of a Screening Period (Visit 1), a 52-week Treatment Period (Visits 2 to 14 days prior to Visit 7), a 14-day post-treatment Safety Follow-Up Period (Visit 7), as well as a 12-month Post-Treatment Follow-Up (PTFU) period. Throughout the Treatment Period, compliance with study intervention administration will be recorded daily in the electronic diary (eDiary). The eDiary is also used for the assessment of the primary and several secondary efficacy endpoints. Assessments of safety will be performed throughout the study. Bone mineral density (BMD) is monitored via dual-energy X-ray absorptiometry (DXA) scan at 6-month intervals during the treatment period and the post-treatment follow-up period. The study schema is presented in Figure 1.

Upon completion of 13 cycles of treatment under this protocol, patients will have the option of entering into MVT-601A-006 for an additional three years of treatment. If patients have opted to enroll in MVT-601A-006, they will still undergo all assessments at the follow-up/end-of-treatment (EOT) visit (Visit 7) 14 days after the last dose under MVT-601-050, in order to ensure that all on-treatment pregnancies are noted, as well as fulfill requirements related to the 12-month visit in MVT-601A-006. Post-treatment follow-up for these patients will be completed under the MVT-601A-006 protocol.

Figure 1: Schematic of Study MVT-601-050



Abbreviations: CHC = combined hormonal contraceptive; DXA = dual energy X-ray absorptiometry; E2 = estradiol; eDiary = electronic diary; NETA = norethindrone acetate; PTFU = post-treatment follow-up; QD = once daily; Rel-CT = relugolix combination therapy (relugolix 40 mg, E2 1 mg, NETA 0.5 mg); UPT = urine pregnancy test.
* Or 14 days after the last dose.
Note: During the treatment period, Visits occur approximately 7 days after the end of the previous cycle. Phone visits occur approximately 6 weeks after the subsequent Visit.

2.3. Randomization

This is an open-label, single-arm study. Randomization is not applicable to this study.

2.4. Sample Size and Power

2.4.1. Sample Size Estimation

The study is designed to measure contraceptive efficacy of relugolix combination therapy in women with uterine fibroids or endometriosis using the Pearl Index (defined as the number of on-treatment pregnancies per 100 woman-years of treatment, where 1 woman-year means 13 consecutive 28-day treatment cycles). Based on the assumptions below, the study will provide > 90% power for the primary efficacy analysis using two-sided type-I error rate of 5%.

To support the primary contraceptive efficacy analysis, the sample size of this trial has been set to attain at least 10,000 menstrual cycles of drug exposure in study patients. The following assumptions are used in the sample size and power calculation.

- 60% patients will complete the study, with each patient contributing 13 treatment cycles.
- 40% patients will discontinue the study, with each discontinuer contributing 5 treatment cycles on average.
- 70% of treatment cycles contributed by all patients will be “at-risk” for pregnancy (i.e., 7000 at-risk cycles).
- Pearl Index is approximately 2.0.

Pearl Index value 2.0 has been chosen based on the pooled estimate of the Pearl Index by Food and Drug Administration (FDA) in the New Drug Application (NDA) review of NuvaRing. The discontinuation rate of 40% has been assumed based on the historical contraceptive trials.

2.4.2. Justification

The sample size and power calculations were based on the simulation work using the above assumptions. The methodology described by (Benda et al. 2004) was used for calculating the 95% confidence interval (CI) for the PI with the assumption of a Poisson model for number of on-treatment pregnancies. In the Poisson model, a Poisson distributed random variable X is assumed with parameter $\theta = 100\lambda T$ where λ is the pregnancy rate per year and T is the total exposure time in 100-woman years. X represents the number of on-treatment pregnancies occurring during a total exposure time of 100T woman-years. The PI is given by $\rho = \theta/T$ which is the expected number of pregnancies within 100-woman years (or 1300 28-day cycles). Based on this model, a two-sided 95% CI was derived for each simulation. Power (probability of success) was estimated as the proportion of the simulated 95% CIs meeting the criterion that the upper bound of the 95% CI is less than 5, consistent with FDA Guidance for Industry (FDA 2019).

Based on the above assumptions and model, approximately 1020 patients will be enrolled to achieve at least 7000 at-risk cycles. Additional patients may be enrolled to ensure a minimum of 200 patients completing the study. With 7000 at-risk cycles, the study will have 95% power for the upper limit of the two-sided 95% CI for the Pearl Index to be < 5 in the overall population, assuming that the true Pearl Index is 2 and that a Poisson model is used to achieve the CI.

3. PLANNED ANALYSES

There will be two formal analyses in this study. This SAP describes both analyses to be performed for the study.

3.1. Interim Analysis

No formal interim efficacy analyses are planned for this study.

3.2. Primary Analysis

The primary analysis of all efficacy and safety data from study MVT-601-050 will occur after all enrolled patients have completed the 52 weeks of study treatment or discontinued the study and completed the 14-day safety follow-up visit. The results of the primary analyses will be used for the clinical study report (CSR) of MVT-601-050.

3.3. Post-Treatment Follow-Up Analysis

The post-treatment follow-up analysis will occur after all patients have completed 12 months post-treatment follow-up for those patients who have not opted to enroll in MVT-601A-006. This analysis will concentrate on two post-treatment DXAs to be performed at 6 months and 12 months after the last dose of study drug. Data collected during the post-treatment follow-up period will be summarized and reported in an addendum to the CSR of MVT-601-050.

4. GENERAL ANALYISIS DEFINITIONS

4.1. Treatment

In this study, patients will be allocated to receive the following open-label oral study treatment once daily for the entire duration of the treatment period (thirteen 28-day treatment cycles, total 52 weeks):

Relugolix+E2/NETA: relugolix combination therapy (relugolix 40 mg, estradiol 1 mg, and norethindrone acetate 0.5 mg).

4.2. Analysis Populations

The analysis populations and their definitions are provided below. Before implementation of protocol amendment 2, eight healthy participants were enrolled under protocol amendment 1 and six of them had at least one dose of study intervention. These healthy participants will not be included in the analysis populations. In addition, their data, including but not limited to, demographics and baseline characteristics, study drug exposure, and adverse events, will be reported separately. Starting from protocol amendment 2, patients in this study had a diagnosis of uterine fibroids or endometriosis as defined in the protocol inclusion criteria. The number and percent of patients meeting the definition of each analysis population will be summarized by diagnosis cohort (uterine fibroids or endometriosis) and total.

Management of patients with duplicate records is detailed in Section 4.8.

4.2.1. Enrolled Population

The enrolled population is defined as all patients who have completed the informed consent process, completed screening procedures, and have been allocated to treatment.

4.2.2. Intent-to-treat Population (ITT)

The ITT population is defined as all enrolled patients who receive at least one dose of study intervention. The ITT population is the same as the safety population (See Section 4.2.6).

4.2.3. Modified Intent-to-treat Population (mITT)

The mITT population is defined as the subset of patients included in the ITT who have at least one treatment cycle without use of other contraceptive methods (regardless of vaginal intercourse), or patients with a treatment cycle (at-risk or not) in which a pregnancy has occurred.

4.2.4. Restricted Intent-to-treat Population (rITT)

The rITT population is defined as the subset of patients included in the ITT population who have at least one at-risk treatment cycle or who have a treatment cycle (at-risk or not) in which a pregnancy has occurred. This population will be used for the primary analysis of the primary endpoint described in Section 7.3.1.

4.2.5. Per-Protocol Population (PP population)

The PP population is defined as the subset of patients in the rITT population with at least one treatment cycle and without meeting prespecified important protocol deviations. Important protocol deviations that will exclude patients from the PP population are specified in Section 5.3.

If the PP population comprises > 95% of the rITT population, the sensitivity analysis of the primary endpoint using the PP population will not be performed. The PP population and its associated subset of important protocol deviations that occur prior to the primary analysis timepoint will be finalized prior to the database lock for the primary analysis.

4.2.6. Safety Population

The safety population is defined as all patients who receive at least one dose of study intervention. The safety population is the same as the ITT population.

Safety analyses will be performed using the safety population unless otherwise specified.

4.3. General Rules for Analysis Populations

The general rules for handling analyses using an analysis population are discussed in Table 2. The definitions of analysis populations are described in Section 4.2.

In general, data for efficacy will be presented for total and the other data such as demographics and safety, etc. will be presented for diagnosis cohort and total.

Table 2: General Considerations for the Usage of Analysis Populations

Type of Endpoint	Analysis Population
Disposition	Enrolled population
Screen failure	All screened participants
Demographics and baseline characteristics Protocol deviations	ITT population
Medical history Prior medications Concomitant medications Study drug exposure and compliance AEs Labs Vital signs and body weight Bone mineral density Bleeding profile	Safety population
Efficacy	At-Risk PI (primary efficacy endpoint): rITT population. Modified At-Risk PI (secondary efficacy endpoint): mITT population. Gross PI (secondary efficacy endpoint): ITT population. Method Failure PI (secondary efficacy endpoint): PP population. Cumulative 1-year pregnancy rate: ITT population, mITT population, rITT population, and PP population.

4.4. Data Presentation Conventions

All statistical analyses will be conducted using SAS® Version 9.4 or higher.

Statistical tests for the primary and secondary efficacy endpoints will be assessed at a two-sided $\alpha = 0.05$ significance level, and all CIs will be reported as two-sided, unless otherwise stated.

Where appropriate, variables will be summarized descriptively by study visit. For the categorical variables, the count and proportions of each possible value will be tabulated by diagnosis cohort and total. For continuous variables, the number of patients with non-missing values, mean, median, standard deviation (SD), minimum, and maximum values will be tabulated.

Unless otherwise specified, the following conventions will be applied to all analyses:

- Unless otherwise specified, mean and median values will be formatted to one more decimal place than the measured value; standard deviation values will be formatted to two more decimal places than the measured value; minimum and maximum values will be presented to the same number of decimal places as the measured value. If the measured value is large (e.g., > 100), fewer decimal places may be displayed.
- Percentages will be rounded to one decimal place.
- P-values will be rounded to four decimal places. P-values less than 0.0001 will be presented as "< 0.0001" and p-values greater than 0.9999 will be presented as "> 0.9999."
- 1 month = 30.4375 days. Month is calculated as (days/30.4375) rounded to 1 decimal place.
- 1 year = 365.25 days. Year is calculated as (days/365.25) rounded to 1 decimal place.
- Age will be calculated using the date of first dose as a reference. If only year of birth is collected, then 1 July of the year of birth will be used to calculate age. If year and month of birth are collected, then 1st of the month of birth will be used to calculate age. If year is missing, no imputation will be performed.
- 1 pound = 0.454 kg.
- 1 inch = 2.54 cm.
- Missing efficacy or safety data will not be imputed, unless otherwise specified.
- For laboratory results above or below sensitivity limits displayed as "<" or ">" a quantification threshold, 0.0000000001 will be subtracted or added, respectively, to the threshold to derive a numeric result for analysis.
- For laboratory analysis datasets, both SI units and conventional units will be derived if applicable.
- For safety analyses, percentages will be calculated based on the number of patients in the analysis population in each diagnosis cohort and total.

- For by-visit observed data analyses, percentages will be calculated based on the number of patients with non-missing data as the denominator, unless otherwise specified.
- For other continuous endpoints, the summary statistics will include mean, SD, median, minimum, and maximum.
- For time-to-event endpoints, the summary statistics will include median time to event-free survival, the 25th and 75th percentiles, and the number of patients at risk at specified time points.
- For categorical endpoints, the summary statistics will include counts and percentages.
- Confidence intervals, when presented, will generally be constructed at the 95% level. For binomial variables, exact methods will be employed, unless otherwise specified.

4.5. Definitions, Computations, and Conversions

4.5.1. Definition of Date of First Dose and Date of Last Dose of Study Drug

The date of the first dose of study drug is defined as the date when a patient receives the first dose of study drug. The date of the last dose of study drug is defined as the date a patient receives the last dose of study drug in study MVT 601-050. If the complete date of last dose of study drug is unknown, the latest date the study drug was known to have been taken will be used based on data collected in the eCRF.

4.5.2. Study Day

Study day will be calculated with respect to the date of the first dose of study drug (Study Day 1). For assessments conducted on or after the date of the first dose of study drug, study day will be calculated as:

(Assessment date – date of first dose of study drug) + 1

For assessments conducted before the date of the first dose of study drug, study day will be calculated as:

(Assessment date – date of first dose of study drug)

For patients who do not receive any amount of study drug, study day will be calculated as above with respect to the date of treatment allocation in Interactive Web Response System (IWRS).

4.5.3. Definition of Treatment Duration

Treatment duration is defined as the duration of time from the date of the first dose of study drug to the date of the last dose of study drug as follows:

(Date of last dose of study drug – Date of first dose of study drug) + 1

For patients without complete date of last dose of study drug, the last date of study drug was known to have been taken will be derived as described in Section 4.5.1 and used to calculate treatment duration.

4.5.4. Definition of Baseline Value and Post-Baseline Value

Unless otherwise specified, baseline values are defined as the last available measurement before the first administration (date and time) of study drug. A post-baseline value is defined as a measurement taken after the first administration of study drug. Change from baseline is defined as (post-baseline value − baseline value). Percent change from baseline is defined as (post-baseline value – baseline value)/ (baseline value) × 100%.

Both date and time of study drug administration and measurement will be considered when calculating baseline value. If the time is not available, the date alone will be used. For patients who receive no study drug, the date of treatment allocation will be used in place of the date of first dose in determining baseline and post-baseline values.

For BMD only, the last on-treatment BMD value is defined as the last available BMD measurement within 90 days of the last dose of study drug. A BMD measurement taken after 90 days of the last dose of study drug will be considered the PTFU BMD measurement. Change from last on-treatment BMD is defined as (PTFU BMD value – last on-treatment BMD value). Percent change from last on-treatment BMD is defined as (PTFU BMD value – last on-treatment BMD value) / (last on-treatment BMD value) × 100%.

4.5.5. Start Date of Treatment Cycle

Start date of treatment cycle is defined as the date of the first dose of study drug. If the date of the first dose of study drug and DAY1 on eDiary did not match, the start date is defined as DAY1 anchor date on eDiary. This start date rule will be applied to all analyses for efficacy endpoints, including but not limited to cumulative one-year pregnancy rate and sensitivity analyses for primary endpoint.

4.5.6. Visit Analysis Windows

Visit analysis windows, which will be used to associate assessments with a scheduled visit, will be used only for summarizing data by study visit.

Data will be handled according to the following rules to define the visit analysis window, unless otherwise specified. All data will be used in determining the most extreme on-treatment value, unless otherwise specified. On-treatment period is defined in Section 4.5.7.

If the results from more than one assessment are within a given visit analysis window, the non-missing result from the assessment closest to the target date will be used. If multiple assessments are done on the same day, the latest assessment will be used for the analysis. If two assessments are equally close to the target day, the later assessment will be used.

For end-of-cycle data on eDiary used for the efficacy assessments, the analysis window follows Table 3 below.

Table 3: End-of-Cycle Data on eDiary Assessment Window

End of Cycle	Start Day	Target Day	End Day
1	21	28-30	37
2	49	56-58	65
3	77	84-86	93
4	105	112-114	121
5	133	140-142	149
6	161	168-170	177
7	189	196-198	205
8	217	224-226	233
9	245	252-254	261
10	273	280-282	289
11	301	308-310	317
12	329	336-338	345
13	357	364-366	373

Evaluable end-of-cycle used for the efficacy analysis is defined the treatment cycles within + or – 7 days of the target days. The day for each end-of-cycle will be based on start date of treatment cycle (See Section 4.5.5).

For all safety assessments except bone mineral density (BMD), visit analysis window follows Table 4 as below.

Table 4: Visit Analysis Windows

Visit	Start Day	Target Day	End Day
Visit 3 (7 days after Cycle 1)	2	35	77
Visit 4 (7 days after Cycle 4)	78	119	161
Visit 5 (7 days after Cycle 7)	162	203	245
Visit 6 (7 days after Cycle 10)	246	287	332
EOT: Visit 7 (14 days after Cycle 13)	333	378	420

BMD is determined using DXA scan and will be assessed at the lumbar spine (L1, L2, L3, and L4), total hip, and femoral neck (same leg within each patient). A screening, 6- and 12-month on-treatment, and 6- and 12-month PTFU DXA scans will be performed. The 12-month on-treatment DXA scan may occur at the EOT visit.

For early termination due to reasons unrelated to bone density loss occurring prior to completing 6 cycles of study medication, an early-termination DXA scan is not required. For early termination occurring after completing 6 cycles of study medication, the patient should complete a DXA at the time of early termination. If the 6-month on-treatment DXA was completed within the last 6 weeks and the patient has completed fewer than 8 cycles of study medication, an early termination DXA does not need to be performed.

BMD assessment window is shown in Table 5 below. Patients who develop BMD decline (compared to pre-treatment baseline) of > 3% or a Z-score ≤ -2.0 at any anatomic site (lumbar spine, total hip, or femoral neck) will undergo repeat DXA scan to confirm this measurement. The repeat scan will only be used for analysis if it occurs within 60 days of the corresponding first DXA scan demonstrating BMD loss of > 3% or a Z-score ≤ -2.0. The 12-month on-treatment DXA will be assumed to be the last on-treatment DXA, unless a patient early terminates or does not complete a 12-month on-treatment DXA, in which case the last on-treatment DXA will be defined as the last DXA performed within 90 days from the last dose date.

Table 5: BMD Assessment Window

Visit	Start Day	Target Day	End Day
6-month on Treatment	93	183	274
12-month on Treatment	275	364	456
6-month post Treatment	Date of the last dose of study drug + 93 days	Date of the last dose of study drug + 183 days	Date of the last dose of study drug + 274 days
12-month post Treatment	Date of the last dose of study drug + 275 days	Date of the last dose of study drug + 365 days	Date of the last dose of study drug + 456 days

Assessments will be excluded from analysis if they occur on or after the date of initiation of another investigational agent or hormonal therapy affecting the hypothalamic-pituitary-gonadal axis or surgical intervention for uterine fibroids or endometriosis. Assessments after date of last dose of study drug + 92 days will be excluded from on-treatment analysis. Repeat DXA scan within 60 days of the corresponding first DXA scan will be included for analysis.

4.5.7. Definition of On-Treatment Period

- For **contraceptive efficacy**, an on-treatment pregnancy is defined as a pregnancy with an estimated date of conception that occurs during the On-Treatment Period, which is defined as the period of time from start date of treatment cycle up to and including 7 days after the last dose of study drug.
- For **treatment-emergent adverse events**, the On-Treatment Period is defined as the period of time from the date of first dose of study drug up to and including 14 days after the date of last dose of study drug.
- For **prior, concomitant or post-treatment medications**, the On-Treatment Period is defined as the period of time from the date of first dose of study drug to the date of last dose of study drug, inclusive.
- For **other safety data** (e.g. laboratory results and vital signs), the On-Treatment Period is defined as the period of time from the date of first dose of study drug up to and including 14 days after the date of last dose of study drug. Unless otherwise specified, this On-Treatment Period will be applied to the analyses that only include on-treatment results (e.g., abnormal results during the on-treatment period), but not applied to the by-visit analyses that include results for EOT/Visit 7 derived as per the visit analysis window algorithm specified in Section 4.5.6.

Dosing interruptions will not change the definition of the on-treatment period. If a patient had a period of time when study drug was not being taken between the date of first dose of study drug and the last dose of study drug, this time will still be considered to be during the On-Treatment Period.

4.6. General Rules for Missing Data

The rules for handling missing data for by-visit endpoints and incomplete dates associated with adverse events and concomitant medications are discussed in this section. The rules for handling missing data for efficacy endpoints are described in Section 7.3.

4.6.1. By-Visit Endpoints

By-visit endpoints will be analyzed using observed data, unless otherwise specified. For observed data analyses, missing data will not be imputed and only the observed records will be included.

4.6.2. Adverse Events and Concomitant Medications

The following imputation rules for the safety analyses will be used to address the issues with partial dates. The imputed dates will be used to determine the treatment-emergent period. For adverse events with a partial date, available date parts (year, month, and day) of the partial date will be compared with the corresponding date components of the start date and end date of the treatment emergent period (see definition in Section 4.1) to determine if the event is treatment emergent. When in doubt, the adverse event will be considered treatment emergent by default.

The following rules will be applied to impute partial dates for adverse events:

- If start date of an adverse event is partially missing, impute as follows:
  - If both Month and Day are missing and Year = Year of treatment start date, set to treatment start date if adverse event end date is not prior to treatment start date, or set to January 1 if adverse event end date is prior to treatment start date;
  - If both Month and Day are missing and Year ≠ Year of treatment start date, set to January 1;
  - If Day is missing and Month and Year = Month and Year of treatment start date, set to treatment start date if adverse event end date is not prior to treatment start date, or set to first of the month if adverse event end date is prior to treatment start date;
  - If Day is missing and Month and Year ≠ Month and Year of treatment start date, set to first of the month;
  - If start date is completely missing, set to treatment start date if adverse event end date is not prior to treatment start date.

- If end date of an adverse event is partially missing, impute as follows:
  - If both Month and Day are missing, set to December 31;
  - If only Day is missing, set to last day of the month;
  - If end date is completely missing, do not impute.

When the start date or end date of a medication is partially missing, the date will be imputed to determine whether the medication is prior or concomitant (or both).

The following rules will be applied to impute partial dates for medications:

- If start date of a medication is partially missing, impute as follows:
  - If both Month and Day are missing, set to January 1;
  - If only Day is missing, set to the first of the month.
- If end date of a medication is partially missing, impute as follows:
  - If both Month and Day are missing, set to December 31;
  - If only Day is missing, set to last day of the month.
- If start date or end date of a medication is completely missing, do not impute.

For purpose of determining medication status as prior or concomitant, the medication will be treated as it began prior to the first dose of study drug if start date is completely missing; the medication will be considered as ongoing if end date is completely missing.

4.7. Handling Multiple Values on the Same Day

4.7.1. Questionnaires

If a questionnaire (ie, Patient Health Questionnaire-9 [PHQ-9], the Columbia Suicide Severity Rating Scale [C-SSRS], and Diagnostic and Statistical Manual of Mental Disorders 5 [DSM-5]) is completed multiple times on the same day, the questionnaire assessment with the worst score will be used for analysis.

4.7.2. eDiary Entries

Participants will complete daily eDiary entries including compliance with study intervention dosing, occurrence of vaginal bleeding and its severity, monthly assessments of occurrence of sexual intercourse, use of back-up contraception, and home urine pregnancy test results. If there are multiple eDiary entries on the same day for a patient, the worst level will be used for analysis. That is, occurrence of vaginal bleeding will be chosen in the order of Yes (patient experienced vaginal bleeding) > No (patient did not experience vaginal bleeding); the amount of bleeding will be chosen in the order of extremely heavy > heavy > moderate > light > spotting.

For end-of-cycle questions of vaginal sexual intercourse and other form of birth control, if there are multiple eDiary entries on the same treatment cycle for a patient, values will be chosen in a conservative way for consideration of at-risk cycles in efficacy analysis. That is, value of vaginal sexual intercourse will be chosen in the order of No > Yes; value of other form of birth control will be chosen in the order of Yes > No.

4.8. Handling Duplicate Patients

Duplicate patients were defined as individuals who enrolled at two or more study sites, either concurrently or sequentially. Potential duplicates were identified, investigated, and confirmed during on-site visits. In every case, the sites involved were unaware of the participant's enrolment at another location. Upon confirmation, the sites were instructed to terminate these participants from the study and document the event as a major protocol deviation. Due to inconsistencies across multiple parameters within the duplicate records, it was not possible to determine which data were accurate. Therefore, data from these patients will be excluded from all analysis populations as defined in the SAP. The disposition, demographics, administration of study drug, pregnancy details (if any), treatment cycles and safety data of duplicate patients will be provided in listings, separately from the listings of the analysis populations, and will be discussed in the clinical study report (CSR) separately.

5. STUDY POPULATION

5.1. Patient Disposition

The number and percentage of patients for each of the following categories will be summarized by diagnosis cohort and total based on all enrolled patients:

- Patients who completed the treatment period
- Patients who did not receive any study drug.
- Patients who discontinued early from the treatment period and reasons for discontinuation
  - Patients who were discontinued from treatment due to bone loss meeting the threshold for withdrawal per protocol are separately quantified under the "Other, Specify" category.
- Patients who completed the post-treatment follow-up period
- Patients who discontinued from the post-treatment follow-up period and reasons for discontinuation
  - Patients who rolled over into the MVT-601A-006 (MOXIE) study are separately quantified under the "Other, Specify" category.

This summary will be also performed by age group (≤ 35 years old / > 35 years old).

A by-patient listing of patient disposition will be provided.

5.2. Screen Failure

For all screened patients, the number and percentage of patients with screen failure will be summarized. In addition, reason for screen failure and criterion not met will be summarized. The number and percentage will be based on the patients with screen failure.

5.3. Protocol Deviations

Protocol deviations will be categorized as important per the protocol deviation plan. Important protocol deviations (IPDs) include, but are not limited to, the following categories:

- Enrolled patient who did not satisfy key eligibility criteria;
- Enrolled patient who met withdrawal criteria during the study but was not withdrawn;
- Enrolled patient who received a prohibited concomitant medication that met criteria for an important protocol deviation;
- Enrolled patient who had low compliance (< 75%) with daily intake of study drug;
- Enrolled patient who missed key study procedures or visits;

IPDs will be summarized by deviation category for all patients in the ITT population. A by-patient listing of IPDs will be provided.

In addition, patient eligibility, including inclusion criteria that are not met and exclusion criteria that are met, will be summarized for all patients in the ITT population.

Prior to the database lock for the primary analysis, the selected subset of the IPDs that are likely to affect the primary efficacy outcome will be finalized according to the protocol deviation plan to define the Per-Protocol population.

This subset of the IPDs includes, but is not limited to, the following IPD categories:

1. Did not satisfy key entry criteria.
2. Drug compliance < 75% during the primary analysis time frame.
3. Patient received prohibited concomitant medications that met criteria for important protocol deviation: restricted to patients who received prohibited analgesics or concomitant medications that may cause significant drug-drug interactions during the primary analysis time frame.

5.4. Demographics and Baseline Characteristics

Demographic and baseline characteristics will be summarized by diagnosis cohort and total for the ITT population. Categorical and binary data will be summarized using frequencies and percentages, by diagnosis cohort and total (Table 6). Summaries of continuous data will display the mean, SD, median, minimum, and maximum. The numbers of missing values will also be summarized.

Table 6: Categories for Demographic and Baseline Characteristics

Variable	Category
Age Group (years)	≤ 35, > 35
Race	Black or African American, White, American Indian or Alaska Native, Asian, Native Hawaiian or Other Pacific Islander, Other
Ethnicity	Hispanic or Latino, Not Hispanic or Latino, or Not reported
BMI (kg/m^2)	< 30, ≥ 30
Alcohol Consumption History	None, Moderate, Heavy
Smoking History	Never smoker, Former smoker, Current smoker
Prior Contraceptives	Yes, No
Baseline Vitamin D Level (ng/mL)	< 12, 12 to ≤ 19, > 19 to ≤ 30, > 30
Method of Contraception Used	Barrier methods (Condom, Spermicide, Withdrawal, Contraceptive Diaphragm), Progestin-only contraceptive pills, Long-acting reversible contraceptives (Hormonal IUD, Non-hormonal IUD, Contraceptive implant), Contraceptive injection, Combined hormonal contraceptives (Combined oral contraceptive, Contraceptive trans-dermal patch, Hormonal contraceptive vaginal ring), Other
Timing of First Dose (Day 1)	In clinic, On the day patient would normally initiate new contraceptive cycle of hormonal contraception, Within 3 days of onset of menses, Other
Renal Function [a]	CrCl: < 60, ≥ 60 to < 90, ≥ 90 mL/min
Endometriosis-associated pain severity[b]	Worst pelvic pain on days patients were having most recent period or Worst pelvic pain on days patients were not having period: Absent, Mild, Moderate, Severe and Very severe
Transvaginal Ultrasound[c]	Number of fibroids visualized: One, Two or more and None Type (s) of Fibroids: Intramural, Subserosal and Submucosal

Abbreviations: BMI = body mass index; CrCl = creatinine clearance; IUD = intrauterine device.
[a] Creatinine clearance by Cockcroft-Gault Equation = 0.85 × {(140 – age(years)) × weight(kg)} / (72 × Serum Creatinine(mg/dL)) for women.
[b] For women with endometriosis
[c] For women with uterine fibroids

This summary will be also performed by age group (≤ 35 years old / > 35 years old).

5.5. Medical History

General medical history and post-treatment medical and gynecological conditions will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) version 26.0 or higher and will be summarized by system organ class (SOC) and preferred term (PT) by diagnosis cohort and total for the safety population. A patient with multiple occurrences of medical history within a PT will be counted only once in that PT.

Gynecologic and obstetric history, including age at first menstrual period, usual menstrual cycle length, usual menstrual period duration, and number of pregnancies (total, stillbirths, live births, spontaneous abortions, therapeutic or elective abortions, ectopic or molar pregnancies) will be summarized.

By-patient medical history data will be listed.

5.6. Prior, Concomitant and Post-treatment Medications

Prior medications and concomitant medications taken during the study treatment period and post-treatment concomitant medications will be summarized for all patients in the safety population by diagnosis cohort and total. Medications are considered prior if exposure occurs before the first dose of study drug. Medications are considered concomitant if exposure occurs during the on-treatment period as defined in Section 4.5.7. Medications will be considered post-treatment if exposure is started or changed in the post-treatment period.

The number and percentage of patients who took at least one dose of a prior or concomitant medication or post-treatment concomitant medication will be summarized by diagnosis cohort and total using the World Health Organization (WHO) Drug Dictionary and summarized according to the Anatomical Therapeutic Chemical (ATC) classification system Level 3, Level 4 and generic medication name. ATC Level 1 term and Level 2 term will be used if both ATC Level 3 term and ATC Level 4 term are not available. A patient who has been administered several medications with the same preferred medication name will be counted only once for that preferred medication name. The ATC terms and the preferred medication names within each ATC term will be sorted alphabetically by ATC Level 3 term, ATC Level 4 term and generic medication name.

For the prohibited medications, certain medication types will be summarized separately, including hormonal medications, medications that could potentially affect concentrations of relugolix, E2, NETA, and medications known to affect BMD.

5.7. Prior and Concomitant Surgeries/Procedures

Prior and concomitant surgeries/procedures data will be collected and listed.

6. STUDY DRUG EXPOSURE AND COMPLIANCE

Patients in the safety population will be summarized for the extent of exposure and compliance to study drug by actual treatment received based on electronic case report form (eCRF).

6.1. Study Drug Exposure

Study drug exposure summaries will include the total dosage taken in milligrams, the total number of tablets taken, and the treatment duration (Section 4.5.3).

6.2. Study Drug Compliance

Study drug compliance will be summarized for the treatment period and will be calculated as follows:

Study drug compliance (%) = (total tablets taken / total tablets expected to be taken) × 100

The total tablets taken will be calculated as:

(Total tablets dispensed – total tablets returned)

The total tablets expected to be taken is calculated as the total number of tablets a patient is expected to take each day times the length of time (in days) that the patient was in the treatment period of the study excluding the days of study drug interruption as collected in the eCRF. During the period prior to the last dispensing visit, for patients who did not return all kits from a set of kits dispensed, the number of kits dispensed and returned will not contribute to overall compliance calculation. For patients who did not return tablets that were dispensed at the last drug-dispensing visit, they will be assumed to have taken study drug daily as planned during the period from the last drug-dispensing date to the date of last dose of study drug. For patients who did not return for any post-baseline visits and did not return dispensed study drug, or whose drug return information is not complete enough to support the compliance calculation, study drug compliance will not be calculated and will be categorized as “not able to calculate” in summaries of study drug compliance.

Summary statistics of study drug compliance (e.g., mean, median, etc.) will be presented, along with a categorical summary.

This summary will be also performed by age group (≤ 35 years old / > 35 years old)

By-patient data of administration of study drug and overdose will be listed.

7. EFFICACY ANALYSIS

7.1. General Considerations

Efficacy analyses will be conducted on the ITT population, mITT population, rITT population, or PP population.

7.1.1. Analysis of Binary Data

In general, binary data will be summarized by frequency counts and percentages.

7.1.2. Analysis of Categorical Data

Qualitative variables will be summarized by frequency counts and percentages. Unless otherwise specified, the calculation of proportions will include the missing category. Therefore, counts of missing observations will be included in the denominator and presented as a separate category.

7.1.3. Analysis of Continuous Data

Continuous variables will be summarized using descriptive statistics (e.g., n, mean, median, SD, minimum, maximum, and first and third quartiles). For the analyses of change from baseline, the mean at baseline will be calculated for all patients with at least one post-baseline value. Additionally, the mean will also be calculated for each visit, including only the patients who are in the analysis who have data for that visit.

7.2. Multiplicity Adjustment

This is a single-arm study and there will be no treatment comparison in the analyses of all efficacy and safety endpoints. No formal interim efficacy analyses are planned for this study. There will be a single primary analysis of efficacy and safety data, and a post-treatment follow-up analysis of safety data (see Sections 3.2 and 3.3). There is a single primary efficacy endpoint planned. No testing hierarchy is planned. Therefore, no multiplicity adjustment will be made for this study.

7.3. Primary Efficacy Endpoint

The primary efficacy endpoint in this study is the At-Risk PI, defined as the number of on-treatment pregnancies per 100 woman-years of treatment (where one year consists of thirteen 28-day treatment cycles).

$$\text{Pearl Index (PI)} = \frac{1300 \times (\text{number of on treatment pregnancies})}{\text{number of 28-day at-risk cycles of treatment}}$$

7.3.1. Primary Analysis for the Primary Efficacy Endpoint

There is no hypothesis testing associated with the primary endpoint. The primary contraceptive efficacy will be assessed by the point estimate of At-Risk PI and its corresponding 95% CI. The At-Risk PI will be calculated using on-treatment pregnancies and cycles considered at-risk of pregnancy (i.e., cycles of consecutive 28-day periods without use of any other contraceptive methods and with affirmed occurrence of vaginal intercourse or cycles in which a pregnancy occurred).

Estimation of the At-Risk PI will be conducted using the rITT population. The At-Risk PI will be presented together with the number of treatment cycles, the number of at-risk cycles, the number of at-risk cycles on-treatment pregnancies and the two-sided 95% CI of the At-Risk PI calculated based on a Poisson distribution (Benda et al. 2004).

7.3.2. Derivation of Primary Efficacy Endpoint

Pregnancy is monitored closely during the study. Pregnancy testing is conducted as follows:

- A serum β-hCG pregnancy test is performed at Visit 1 and Visit 7. Serum testing is also performed during the treatment period if a pregnancy is suspected or to confirm a pregnancy reported by a patient (unless already confirmed by transvaginal ultrasound). A urine pregnancy test is performed at all site visits after screening and as needed.
- A home urine pregnancy test performed by the patient is required prior to the start of each cycle, and the result must be negative for the patient to continue on treatment. Any positive pregnancy test should be confirmed by a serum β-hCG pregnancy test (and transvaginal ultrasound when applicable).
- A pregnancy test must be performed prior to each DXA scan on the day of the DXA.

If a participant has an on-treatment pregnancy, transvaginal ultrasonography will be performed to determine gestational age/estimated date of delivery (EDD).

On-treatment pregnancies are pregnancies with an estimated conception date (ECD) between the first day of study drug intake up to and including 7 days after the last intake of study drug.

Transvaginal ultrasound will be the primary method for determining the ECD. From the ultrasound, the EDD will be ascertained. The ECD will then be calculated as:

$$ECD = EDD - 38 \text{ weeks}$$

If no transvaginal ultrasound is available, the ECD will be determined by principal investigator, whereby the EDD will be assessed through a collection of date of last menses, β-hCG levels, ultrasounds and the clinical picture.

A treatment cycle is expected to include 28 treatment days. Patients who complete the study are expected to have 13 cycles of treatment, approximately 364 days in total. Start date and end date of a cycle will be extrapolated from the date of first dose and tracked via the eDiary.

At the completion of each 28-day treatment cycle, in the eDiary the patient will record whether vaginal intercourse or use of any other contraceptive method occurred during the previous 28 days. At-risk cycles will be defined as cycles in which the following criteria are both met:

1. No other methods of birth control are used by the patient as confirmed in the eDiary;
2. Vaginal sexual intercourse occurred during the cycle as confirmed by the patient in the eDiary.

Cycles will not be considered at-risk if the answers to one or both of the two eDiary questions are missing.

If an on-treatment pregnancy occurs in a cycle, that cycle will be counted as an at-risk cycle, regardless of use of other birth-control methods or occurrence of sexual intercourse during that cycle. Cycles after the date of conception will not be included as at-risk cycles. If a pregnancy test is positive and the conception date can't be determined, cycle(s) up to and including the one in which the positive pregnancy test occurred will be considered for the derivation of the Pearl Index.

For patients without an on-treatment pregnancy, treatment cycles will be counted up to the last at-risk cycle ending on or before the date of last dose of study drug. Cycles ending after the date of last dose of study drug will not be included as at-risk cycles.

The two-sided 95% CI for Pearl Index will be calculated using a Poisson distribution with exact method. The CI limits are calculated by the following equations:

$$CI_{lower} = \frac{1300 \times qchisq(0.025, 2x) / 2}{\text{number of 28-day at-risk cycles of treatment}}$$

$$CI_{upper} = \frac{1300 \times qchisq(0.975,\ 2(x+1)) / 2}{\text{number of 28-day at-risk cycles of treatment}}$$

where x is the number of on-treatment pregnancies and qchisq is the quartile of Chi-square distribution.

On-treatment pregnancies will be summarized. Patients with on-treatment pregnancies will be listed.

7.3.3. Primary Efficacy Endpoint by Group

Analyses of the primary efficacy endpoint will be performed to assess whether contraceptive efficacy is consistent across clinically important groups. Variables in Table 7 will be used for the analyses by group.

Table 7: Variables for Group Analyses of the Primary Efficacy Endpoint

Group Variable	Categories	Analysis population
Age group (years)	≤ 35, > 35	rITT
BMI group (kg/m^2)	< 30, ≥ 30	rITT
Age group (years) and BMI group (kg/m^2)	Age group ≤ 35 and BMI group < 30, Age group ≤ 35 and BMI group ≥ 30, Age group > 35 and BMI group < 30, Age group > 35 and BMI group ≥ 30	rITT
Diagnosis cohort	Uterine Fibroids, Endometriosis	rITT
Race group	Black or African American, Other	rITT
Ethnicity	Hispanic or Latino, Not Hispanic or Latino	rITT
Smoking History	Never Smoker, Former Smoker, Current Smoker	rITT

7.3.4. Sensitivity Analysis for the Primary Efficacy Endpoint

To assess the robustness of the primary endpoint, the following sensitivity analyses of the primary endpoint will be performed (Table 8)

Table 8: Sensitivity Analysis for the Primary Endpoint

Sensitivity Analysis	Purpose	Analysis Strategy
Sensitivity analysis 1	To assess the impact of important protocol deviations that may affect the primary efficacy	Estimation of At-Risk PI will be conducted using the PP population.
Sensitivity analysis 2	To assess the impact of missing data in eDiary	At-risk cycles with < 75% eDiary compliance rate will not be considered at-risk cycles and be excluded from the denominator when calculating the PI, unless on-treatment pregnancy occurred in that cycle. The analysis will be conducted using the rITT population.
Sensitivity analysis 3	To assess the impact of prohibited concomitant medications	Patients who started any prohibited concomitant medication with potential impact on the primary efficacy during treatment period will be censored on the start date of prohibited concomitant medication. The analysis will be conducted using the rITT population.
Sensitivity analysis 4	To assess the impact of missed study drug	Patients who missed 3 or more consecutive days of study drug within 28 days prior to the conception will not be counted as having on-treatment pregnancies. The analysis will be conducted using the rITT population.

7.3.4.1. Sensitivity Analysis 1

To assess the impact of important protocol deviations that may affect the primary efficacy, estimation of At-Risk PI will be conducted using the PP population.

The difference between the primary analysis and this sensitivity analysis is the analysis population (ie, PP population instead of rITT population).

7.3.4.2. Sensitivity Analysis 2

To assess the impact of missing data in eDiary, at-risk cycles with < 75% eDiary compliance rate (ie, < 21 days of eDiary entry in a given cycle) will not be considered at-risk cycles and be excluded from the denominator when calculating the PI, unless on-treatment pregnancy occurred in that cycle. Only cycles with ≥ 75% eDiary compliance rate, without use of any other contraceptive methods and with affirmed occurrence of vaginal intercourse will be included in the estimation of the At-Risk PI in this sensitivity analysis.

The numerator will be the same as the primary analysis of At-Risk PI.

7.3.4.3. Sensitivity Analysis 3

To assess the impact of prohibited concomitant medications, patients who started any prohibited concomitant medication with potential impact on the primary efficacy during treatment period will be censored at the cycle, in which the prohibited concomitant medication was taken. Number of treatment cycles for the censored patients will be counted up to the start date of prohibited concomitant medication rather than the last date of study drug in this sensitivity analysis. A pregnancy with an ECD on or after the start date of prohibited concomitant medication will not be considered an on-treatment pregnancy and will be excluded from the numerator in estimation of At-Risk PI. The analysis will be conducted using the rITT population.

7.3.4.4. Sensitivity Analysis 4

To assess the impact of study drug dosing interruption, patients who missed 3 or more consecutive days of study drug within 28 days prior to the conception date will not be counted as having on-treatment pregnancies in this sensitivity analysis. The analysis will be conducted using the rITT population.

7.4. Secondary Efficacy Endpoints

Secondary efficacy endpoints include three PI-related secondary endpoints and one other secondary endpoint (as described in Section 7.4.2). The three PI-related secondary endpoints include:

- Modified At-Risk PI:

Modified At-Risk PI is based on the number of on-treatment pregnancies occurring during treatment cycles without any other contraceptive methods regardless of occurrence of vaginal intercourse. Evaluable cycles include treatment cycles without use of any other contraceptive methods regardless of occurrence of vaginal intercourse. The analysis will be performed using the mITT population.

- Gross PI:

Gross PI is based on the number of on-treatment pregnancies occurring during all treatment cycles regardless of the use of other contraceptive methods, confirmed vaginal intercourse, or protocol compliance. Evaluable cycles include all treatment cycles. The analysis will be performed using the ITT population.

- Method Failure PI:

Method failure PI is based on the number of on-treatment pregnancies occurring during treatment cycles that are at risk and without selected important protocol deviations that are used to define PP population as described in Section 5.3 where the eDiary documented regular pill intake, excludes cycles with ≥ 7 days forgotten tablets/missed diary entries and/or ≥ 3 consecutive days with forgotten tablets/missed diary entries prior to the estimated date of conception. Evaluable cycles include at-risk cycles without the selected important protocol deviations where the eDiary documented regular pill intake, excludes cycles with ≥ 7 days forgotten tablets/missed diary entries and/or ≥ 3 consecutive days with forgotten tablets/missed diary entries (prior to the estimated date of conception if pregnancy occurred). The analysis will be performed using the PP population.

Comparisons of the primary efficacy endpoint and the PI-related secondary endpoints are provided in Table 9.

Table 9: Summary of the PI-related Efficacy Endpoints

Tier	Endpoint	Analysis Population	Pregnancies (Numerator)	Evaluable Cycles (Denominator)
Primary Endpoint	At-Risk PI	rITT (who have at least one at-risk treatment cycle or who have a treatment cycle (at-risk or not) in which a pregnancy has occurred)	On-treatment pregnancies	At-risk treatment cycles (ie, cycles without use of any other contraceptive methods and with affirmed occurrence of vaginal intercourse), and cycles during which an on-treatment pregnancy occurred.
Secondary Endpoint	Modified At-Risk PI	mITT (who have at least one treatment cycle without use of other contraceptive methods, or patients with a treatment cycle (at-risk or not) in which a pregnancy has occurred)	On-treatment pregnancies occurring during cycles without any other contraceptive methods regardless of occurrence of vaginal intercourse	Treatment cycles without use of any other contraceptive methods regardless of occurrence of vaginal intercourse, and cycles during which an on-treatment pregnancy occurred.
Secondary Endpoint	Gross PI	ITT (who have at least one treatment cycle)	On-treatment pregnancies	Treatment cycles (completed treatment regardless of the use of other contraceptive methods, vaginal intercourse, or protocol compliance)
Secondary Endpoint	Method Failure PI	PP (who is in the rITT population without pre-specified protocol deviations)	On-treatment pregnancies occurring during at-risk cycles without the selected important protocol deviations where the eDiary documented regular pill intake, excludes cycles with ≥ 7 days forgotten tablets/missed diary entries and/or ≥ 3 consecutive days with forgotten tablets/missed diary entries prior to the estimated date of conception	At-risk treatment cycles without the selected important protocol deviations where the eDiary documented regular pill intake, excludes cycles with ≥ 7 days forgotten tablets/missed diary entries and/or ≥ 3 consecutive days with forgotten tablets/missed diary entries (prior to the estimated date of conception if pregnancy occurred)

Abbreviations: ITT = intent-to-treat; mITT = modified intent-to-treat; PI = Pearl Index; PP = per protocol; rITT = restricted intent-to-treat.

Derivation method for primary efficacy endpoint (as described in Section 7.3.2) will be applied to the PI-related secondary endpoints, if applicable. Estimation of PI with the two-sided 95% CI calculated based on a Poisson distribution will be provided for the PI-related secondary efficacy endpoints.

7.4.1. Testing Hierarchy for Key Secondary Endpoints

The testing hierarchy is not applicable to this study.

7.4.2. Other Secondary Efficacy Endpoints

The following describes the analysis methods for the other secondary efficacy endpoint – cumulative 1-year pregnancy rate.

Cumulative 1-year pregnancy rate and the associated 95% CI will be estimated on each of the efficacy analysis populations (ITT, mITT, rITT, and PP) by the Kaplan-Meier (KM) survival analysis. All patients will be followed until they either have an outcome of on-treatment pregnancy or are censored at the date of their last dose of study drug. The unit of time to pregnancy/censoring in the KM analysis will be the cycle, with pregnancies recorded by cycle in which the date of conception occurred. The time to event will be calculated as (conception date/the date of last dose of study drug – date of the start date of treatment cycle + 1)/28. Estimation of conception date is described in Section 7.3.2. All on-treatment pregnancies and all treatment cycles will be counted, regardless of use of other contraceptive methods or confirmed vaginal intercourse during the cycle. The median, quartiles, and probabilities of pregnancy at one year will be estimated by the Kaplan-Meier method. Confidence interval for the Kaplan-Meier estimation is calculated using the exponential Greenwood formula via log-log transformation of the survival function.

7.4.3. Secondary Efficacy Endpoints and Other Secondary Efficacy Endpoints by Group

Analyses of the secondary efficacy endpoints and other secondary efficacy endpoints will also be performed by age group (≤ 35 years old / > 35 years old).

7.5. Exploratory Efficacy Endpoints

Not applicable.

7.6. Change From Protocol Specified Efficacy Analysis

Not applicable.

8. SAFETY ANALYSIS

Unless otherwise specified, safety analyses will be conducted using the safety population according to the treatment received by the patients.

8.1. Adverse Events

Adverse events will be reported from the time of the first dose of study drug through the safety follow-up visit approximately 14 days after the last dose of study drug, or enrollment to MVT-601A-006, whichever occurs first. Serious adverse events reported to the investigator after the safety reporting period should be reported to the sponsor if the investigator assesses the event as related to study drug.

A treatment-emergent adverse event is defined as any adverse event that occurs after administration of the first dose of study drug, and on or before 14 days after the last dose of study drug or before the enrollment to MVT-601A-006 if applicable, whichever is earlier.

The severity of all treatment-emergent adverse events will be evaluated by the investigator based on the National Cancer Institute's Common Terminology Criteria for Adverse Events (CTCAE, Version 5.0, dated 17 Nov 2017) and will be coded to PT and SOC using MedDRA version 26.0 or higher.

Adverse event summaries will be based on treatment-emergent adverse events, unless otherwise specified. All adverse events will be listed. Adverse events occurring prior to administration of any study drug will be listed and flagged in by-patient listings only.

The following tabular summaries (including number and percentage of patients) will be provided:

- Overview of adverse events
- All adverse events
  - By SOC and PT
  - By decreasing frequency of PT
  - By SOC, PT, and maximum severity
  - By SOC, PT, and relationship to study drug
- Grade 3 or above adverse events
  - By SOC and PT
  - By decreasing frequency of PT
  - By SOC, PT, and maximum severity
  - By SOC, PT, and relationship to study drug
- Adverse events leading to study drug withdrawal
  - By SOC and PT
  - By decreasing frequency of PT

  - By time to onset, SOC, and PT
- Adverse events leading to dose interruption
  - By SOC and PT
  - By decreasing frequency of PT
- Adverse events resulting in fatal outcome
  - By decreasing frequency of PT
- Serious adverse events
  - By SOC and PT
  - By decreasing frequency of PT
  - By SOC, PT, and maximum severity
  - By SOC, PT, and relationship to study drug
- Adverse events of clinical interest (alanine aminotransferase [ALT] or aspartate aminotransferase [AST] ≥ 3 × upper limit of normal [ULN] and bone fractures)
  - By SOC, PT, and maximum severity
  - By decreasing frequency of PT

A patient with multiple occurrences of AE within a PT will be counted only once in that PT. System organ class will be sorted alphabetically. Preferred term will be displayed in descending order of frequency based on the total number of patients.

8.1.1. Relationship to Study Drug

Adverse events will be classified as "related" to study drug if the relationship was rated by the investigator as possibly related or probably related.

8.1.2. Severity of Adverse Events

Grade 3 or above adverse events will be summarized by SOC, PT, maximum severity, and relationship to study drug.

8.1.3. Serious Adverse Events

Serious adverse events will be summarized by SOC, PT, maximum severity, and relationship to study drug. All serious adverse events and serious adverse events related to study drug will be listed.

The data handling conventions for and the definition of a serious adverse event are discussed in this section. All deaths during the study, including the post-treatment follow-up period, and deaths that resulted from a process that began during the study, should be included in the analysis.

For more details, deaths occurring during the following time periods or under the following conditions should be considered:

- Deaths occurring during participation in this study;
- Deaths occurring after a patient leaves a study, or otherwise discontinues study drug, whether or not the patient completes the study to the nominal endpoint, if the death:
  - Is the result of a process initiated during the study or other drug exposure, regardless of when it actually occurs; or
  - Occurs within a time period that might reflect drug toxicity for a patient leaving a study or otherwise discontinuing drug.

Death reported prior to database lock date will be included in the analysis. Death reported after database lock date will be saved in the safety database but will not be analyzed.

8.1.4. Adverse Events Leading to Withdrawal of Study Drug

Adverse events leading to withdrawal of study drug are those adverse events collected from the adverse event eCRF page with “drug withdrawn” as the action taken with study drug.

Adverse events with “drug withdrawn” as action taken will be considered as leading to withdrawal of study drug.

Adverse events leading to withdrawal of study drug will be listed.

8.1.5. Adverse Events Leading to Dose Interruption

Adverse events leading to dose interruption are those adverse events collected from the adverse event eCRF page with “drug interrupted” as the action taken with study drug.

Adverse events with “drug interrupted” as action taken will be considered as leading to dose interruption.

8.1.6. Adverse Events Resulting in Fatal Outcomes

Adverse events resulting in a fatal outcome are those adverse events collected from the adverse event eCRF page with “fatal” as the outcome.

The fatal events, if any, will be provided in a by-patient listing.

8.1.7. Adverse Events of Clinical Interest

Adverse events of clinical interest in this study are defined as any increase in ALT or AST $\geq 3 \times$ ULN and bone fractures as pre-specified and collected in eCRF Adverse Events.

Adverse events of clinical interest will be listed.

8.2. Laboratory Data

Laboratory parameters, including chemistry and hematology panels, specified as per protocol, and collected from the central laboratory will be tabulated and presented in by-patient listings. Any local laboratory (if central laboratory results are not available) and hepatitis virus serological test results will be provided in by-patient listings only.

The National Cancer Institutes' CTCAE Grading Scale (Version 5.0, dated 17 Nov 2017) with numeric component will be used to categorize toxicity grade for laboratory parameters. Parameters that have criteria available for both low and high values (e.g., hypercalcaemia for a high value of calcium and hypocalcemia for a low value of calcium) will be summarized for both criteria (low and high). Patients will only be counted once for each criterion. The same patient can be counted for both criteria if the patient has laboratory values meeting each criterion. Shift tables will be provided for each gradable parameter to summarize baseline toxicity grade versus worst post-baseline toxicity grade.

Box plots of laboratory values over time will be plotted for key laboratory parameters. These laboratory parameters include, but are not limited to, hematology (hemoglobin, platelets, leukocytes, neutrophils), creatinine, and hepatic function panel (ALT, AST, alkaline phosphatase [ALP], and total bilirubin).

The change from baseline to each post-baseline study visit will be presented by diagnosis cohort and total for each key laboratory test in both tables and figures.

The number and proportion of patients with liver test elevations will be presented. Liver test elevations are assessed by using the worst post-baseline toxicity grade across all assessments including unscheduled visits for ALT, AST, ALP, and total bilirubin based on the definitions presented in Table 10. Concurrent measurements are defined as measurements taken on the same day.

Table 10: Categories for Liver Test Elevations

Laboratory Test	Category
ALT or AST	ALT or AST > ULN to < 3 × ULN ALT or AST ≥ 3 × ULN to < 5 × ULN ALT or AST ≥ 5 × ULN to < 10 × ULN ALT or AST ≥ 10 × ULN to < 20 × ULN ALT or AST ≥ 20 × ULN
Total bilirubin	Total bilirubin > 2 × ULN
ALT or AST and total bilirubin	ALT or AST ≥ 3 × ULN and total bilirubin > 2 × ULN (concurrent)
ALT or AST, total bilirubin, and ALP	ALT or AST ≥ 3 × ULN and total bilirubin > 2 × ULN and ALP < 2 × ULN (concurrent)

Abbreviations: ALP = alkaline phosphatase; ALT = alanine aminotransferase; AST = aspartate aminotransferase; ULN = upper limit of normal.

8.2.1. Pre-defined Limits of Change in Laboratory Parameters

Summaries based on the pre-defined limits of change for all chemistry and hematology parameters in Appendix 1 will be provided separately using the last on-treatment observations and most extreme on-treatment observations. Summary of the pre-defined limits of change for chemistry and hematology parameters listed in Appendix 1 will be provided using both SI units and conventional units.

8.2.2. Pre-defined Categories for Laboratory Shift Tables

Shift tables based on the pre-defined categories for selected parameters (fasting glucose/glucose, HbA1c, total cholesterol, LDL, HDL, triglycerides) in Appendix 2 will be provided using the most extreme on-treatment observations by both SI units and conventional units.

8.3. Other Safety Data

8.3.1. Vital Signs and Body Weight

Blood pressure (systolic and diastolic), heart rate, body weight, and BMI will be summarized at baseline and each subsequent scheduled assessment by diagnosis cohort and total. Change from baseline will be calculated and presented for each parameter at all scheduled post-baseline assessment time points in both tables and figures. All vital sign and body weight data will also be provided in by-patient listings.

Potentially clinically significant abnormalities in vital signs and body weight are defined in Table 11, and they will be summarized by using post-baseline values that meet the defined criteria. Potentially clinically significant abnormalities will also be flagged in by-patient listings.

Table 11: Categories of Potentially Clinically Significant Abnormalities in Vital Signs and Body Weight

Vital Sign Parameter	Category
Systolic blood pressure	≥ 140 mmHg and > baseline ≥ 180 mmHg and > baseline
	≤ 90 mmHg and < baseline
	Increase of ≥ 20 mmHg from baseline Decrease of ≥ 20 mmHg from baseline
Diastolic blood pressure	≥ 90 mmHg and > baseline ≥ 105 mmHg and > baseline
	≤ 50 mmHg and < baseline
	Increase of ≥ 15 mmHg from baseline Decrease of ≥ 15 mmHg from baseline
Heart rate	≥ 120 bpm and > baseline
	< 45 bpm and < baseline
	Increase of ≥ 15 bpm from baseline Decrease of ≥ 15 bpm from baseline
Body weight	Increase of ≥ 5% from baseline
	Increase of ≥ 10% from baseline
	Decrease of ≥ 5% from baseline
	Decrease of ≥ 10% from baseline

Abbreviations: bpm = beats per minute; mmHg = millimeters of mercury.

8.3.2. Bone Mineral Density

BMD data will be collected and analyzed by the central radiology laboratory. BMD (g/cm^2) and Z-score at lumbar spine (L1-L4), total hip, and femoral neck will be reported. Patients who develop BMD decline (compared to pre-treatment baseline) of > 3% or a Z-score ≤ -2.0 at any anatomic site (lumbar spine, total hip, or femoral neck) will undergo repeat DXA scan to confirm this measurement. For the anatomic location that demonstrated BMD decline of > 3% or a Z-score ≤ -2.0, the two DXA results will be averaged for confirmation of result. The average results will be used in all BMD analyses, unless otherwise specified. If the averaged result meets the criteria (BMD decline of > 3% or a Z-score ≤ -2.0), it is considered confirmed. If the averaged result does not meet the criteria, it is not considered BMD loss. If only one DXA was conducted and meets the criteria, it is considered unconfirmed.

The BMD analyses will be performed in the primary analysis (Section 3.2), as well as in the PTFU analysis (Section 3.3). In the primary analysis, BMD analyses will be performed for all on-treatment BMD data as well as the PTFU BMD data available by the time to do the primary analysis. By-patient data of BMD will be listed.

Corrected BMD data will be used for the analyses. Corrected BMD values are BMD measurements that have been multiplied by a correction factor calculated by the central radiology laboratory using a CUSUM (Cumulative Summing) program to quantify calibration changes in scanners longitudinally over the study.

8.3.2.1. On-Treatment Bone Mineral Density Analysis

The percent change from baseline to the 6- and 12-month -on-treatment, DXA and last on-treatment DXA, will be summarized for each anatomic location using descriptive statistics including 95% CI. Box plots of the percent change from baseline will be also plotted by visit (6- and 12-month on-treatment).

A mixed-effects model with repeated measures (MMRM) will be used to analyze the percent change from baseline in BMD at 6- and 12-month on-treatment for each anatomic location. The model will have age at baseline, visit, baseline BMD value, race (Black or African American vs. Other), and BMI at baseline as fixed effects using an unstructured variance-covariance matrix. If the mixed-effects model fails to converge, a first-order autoregressive variance-covariance matrix will be used. Least square means at each anatomic location will be presented at each visit with associated 95% CIs. in both tables and figures.

The same model described above will be used to analyze the observed BMD value for each anatomic location except that baseline BMD is removed from the model and the BMD at baseline is used for analysis (along with BMD at Month 6 and Month 12 on-treatment). The BMD will be also reported by subgroups, such as age group (≤ 35 years old / > 35 years old), race group (Black or African American / Other), BMI group (< 30 kg/m^2 / ≥ 30 kg/m^2) and ethnicity group (Hispanic or Latino / Not Hispanic or Latino). For each subgroup, BMD measurements at baseline and post-baseline visits, and percent change from baseline in BMD to each post-baseline visit will be summarized by visit and anatomic location using descriptive statistics (mean, SD, median, minimum and maximum) including 95% CI and displayed using forest plots.

The number and percentage of patients with a confirmed BMD decline of > 3% from baseline and the number and percentage of patients with a confirmed Z-score ≤ -2.0 will be presented by anatomic site (lumbar, total hip, and femoral neck) at 6-month, 12-month and last on-treatment DXA.

Categorical representation of percent change from baseline to 6 and 12 months of treatment will be presented by the number and proportion (and 95% CI based on Clopper-Pearson method) of patients who had BMD increase (> 0%), no change (0%), and declines of > 0 to < 2%, ≥ 2% to 3%, > 3% to 5%, > 5% to 8%, and > 8% by anatomic location. The proportion of each category will be plotted in stacked bar chart by anatomic location and visit (6 and 12 months) Similar categorical summary will be performed by baseline Vitamin D (≤ 19 ng/mL, > 19 to ≤ 30 ng/mL, > 30 ng/mL, missing, and overall) at each anatomic location. The analysis of categorical percent change from baseline to the worst on-treatment and last on-treatment DXA will be also presented by anatomic location.

Z-scores and their changes from baseline will be summarized by anatomic location and visit with descriptive statistics including 95% CIs. The number and percentage (and 95% CI based on Clopper-Pearson method) of patients with a Z-score ≤ -2.0 will be presented by anatomic location and visit.

8.3.2.2. Post-Treatment Bone Mineral Density Analysis

The PTFU Month 6 and PTFU Month 12 DXA assessments will be waived for patients who enroll in MVT-601A-006 following completion of 13 cycles of treatment under this protocol. The rest of the patients should remain in MVT-601-050 for PTFU DXA assessments.

Descriptive statistics including 95% CI will be used to summarize percent change in BMD from baseline to PTFU Month 6 and PTFU Month 12 by anatomic location. The mean and its 95% CI will be plotted by anatomic location and visit (PTFU Month 6 and PTFU Month 12). Box plots of the percent change from baseline will be also plotted by visit (PTFU Month 6 and PTFU Month 12).

The number and percentage of patients with a confirmed BMD decline of > 3% from baseline and the number and percentage of patients with a confirmed Z-score ≤ -2.0 will be presented by anatomic site (lumbar, total hip, and femoral neck) at PTFU Month 6, PTFU Month 12 and last PTFU.

Percent change in BMD from baseline to PTFU Month 6, PTFU Month 12 and last PTFU will be presented by the number and proportion (and 95% CI) of patients in each of the following categories: increase (> 0%), no change (0%), declines of > 0 to < 2%, ≥ 2% to ≤ 3%, > 3% to ≤ 5%, > 5% to ≤ 8%, and > 8% by anatomic location. The proportion of each category will be also plotted in stacked bar chart by anatomic location and visit (PTFU Month 6 and PTFU Month 12).

Descriptive statistics including 95% CI will be used to summarize the percent change from last on-treatment DXA to 6- and 12-month PTFU DXA. Similarly, the percent change from last on-treatment DXA to 6- and 12-month PTFU DXA will also be summarized categorically.

Percent change in BMD from last on-treatment DXA to PTFU Month 6, PTFU Month 12 and last PTFU will be presented by the number and proportion (and 95% CI) of patients in each of the following categories: increase (> 0%), no change (0%), declines of > 0 to < 2%, ≥ 2% to ≤ 3%, > 3% to ≤ 5%, > 5% to ≤ 8%, and > 8% by anatomic location.

The BMD will be also reported by subgroups, such as age group (≤ 35 years old / > 35 years old), race group (Black or African American / Other), BMI group (< 30 kg/m^2 / ≥ 30 kg/m^2) and ethnicity group (Hispanic or Latino / Not Hispanic or Latino). For each subgroup, BMD measurements and percent change from baseline in BMD from baseline to PTFU Month 6, PTFU Month 12 and last PTFU will be summarized by anatomic location using descriptive statistics (mean, SD, median, minimum and maximum).

For the subset of patients who experienced a decline in BMD at an anatomic location at the last on-treatment DXA scan (defined as percent change from baseline < 0%), BMD recovery during the PTFU period is defined as follows:

Percent recovery (%) = 100% × ([percent change from baseline to last on-treatment measurement] - [percent change from baseline to PTFU measurement] / [percent change from baseline to last on-treatment measurement]).

Percent recovery will be summarized as a continuous variable using descriptive statistics. In addition, percent recovery will be categorized and summarized by frequency in two ways: Category1: no recovery (≤ 0%), partial recovery (> 0% to 25%, > 25% to 50%, > 50% to 75%, > 75% to < 100%), and full recovery (≥ 100%)

Category 2: no recovery (≤ 0%), partial recovery (> 0% to 50%, > 50% to 100%), and full recovery (≥ 100%).

For subgroup of a decline in BMD (> 3% to ≤ 5%, > 5%), the summary will also be performed in the same manner.

8.3.3. Bleeding

When a patient is menstruating, the amount of bleeding during the past 24 hours is captured daily in the eDiary with five response options: spotting, light, moderate, heavy, and extremely heavy. The bleeding related analysis will be based on the amount of bleeding reported by patients in the eDiary. It will be assumed that there is no bleeding or spotting on days that the patient either uses the eDiary and reports no bleeding or spotting, or when the patient misses reporting menstruating within eDiary cycle.

The analyses of bleeding and spotting will be performed over four 90-day periods: Day 1 to Day 90, Day 91 to Day 180, Day 181 to Day 270, and Day 271 to Day 364, unless otherwise specified. The number of days when bleeding or spotting (including light, moderate, heavy, and extremely heavy) is reported will be summarized with descriptive statistics for each of the four periods. The mean of days for each level of bleeding intensity (or bleeding flow) categories (no bleeding or spotting, light, moderate, heavy, and extremely heavy) will be summarized in table and plotted in stacked bar chart for first month (Day 1 to Day 30) and each of the four 90-day periods.

The bleeding flow categories collected in eDiary will be used to define the bleeding/spotting episode and bleeding pattern categories per 90-day period. A bleeding/spotting episode is defined as a bleeding or spotting event with at least one day with bleeding or spotting reported and bordered by at least 2 consecutive full days with no bleeding or spotting reported. Each episode will be categorized by its duration into prolong (> 7 days) or not prolonged (≤ 7 days). The bleeding patterns will be determined per the definitions in Table 12 (Creinin et al. 2022) based on number of bleeding/spotting episodes over a 90-day period.

The number and percentage of patients belonging to each of the bleeding pattern categories defined in Table 12 will be summarized for each of the four 90-day periods. For patients with frequent bleeding, infrequent bleeding, or normal frequency in each period, the number and percentage of patients with prolonged or not prolonged episodes will be presented. If a bleeding episode spans consecutive 90-day reference periods, it is counted only once, within the period in which it initiates. The number and percentage of patients with bleeding, spotting, and bleeding/spotting days in first month and each of the four 90-day periods will also be summarized. Bleeding includes light, moderate, heavy, and extremely heavy bleeding. Patients with both bleeding and spotting in a period are classified as "bleeding". Only patients with no bleeding or spotting are classified as "no bleeding or spotting"

A patient will not be included in the summary of a period if the patient does not complete the treatment in the period, that is, the patient does not reach the end of period when she discontinues the study drug. No analyses of bleeding pattern after Day 364 will be performed.

Sustained absence of bleeding and spotting is defined as lack of bleeding and spotting for at least 90 days that continues until Day 364 or the date of last dose for early terminated patients who terminated before Day 364. The number and percentage of patients will be summarized over the four periods. In each period, the rate of sustained absence of bleeding and spotting will be calculated as the number of patients with absence of bleeding and spotting beginning on or prior to the end of period divided by the number of patients in safety population. The 95% CI of the rate estimated using Clopper-Pearson exact method will be presented.

Table 12: Definitions of Bleeding Outcomes Criteria and Descriptors

Bleeding Criteria	Descriptors	Definitions [a]
Flow	No bleeding or spotting	No bleeding or spotting during the entire reference period
	Spotting	Light flow with no menstrual product use
	Light	Flow with menstrual product use; subjective assessment as compared to typical flow when not using contraceptive
	Moderate	
	Heavy	
	Extremely heavy	
Pattern	Absence of bleeding/spotting	No bleeding or spotting
	Infrequent bleeding	1 or 2 bleeding/spotting episodes
	Normal frequency	3 or 4 bleeding/spotting episodes
	Frequent bleeding	More than 4 bleeding/spotting episodes

Table 12: Definitions of Bleeding Outcomes Criteria and Descriptors (Continued)

Bleeding Criteria	Descriptors	Definitions [a]
Duration [b]	Prolonged	Bleeding/spotting episode lasting more than 7 days
	Not prolonged	Bleeding/spotting episode lasting no more than 7 days
	Total number of days	Total number of bleeding, spotting, and bleeding/spotting days per reference period

[a] Over a 90-day reference period.
[b] Duration not used when absence of bleeding/spotting occurs.

The number and percentage of patients with bleeding (light, moderate, heavy, and extremely heavy), spotting, and no bleeding/spotting in the first month and each of the four 90-day periods will be summarized. If a patient has both bleeding and spotting, she will be classified as "bleeding". Only patients with no bleeding or spotting will be classified as "no bleeding/spotting".

In addition, the number and percentage of patients discontinuing the study drug due to bleeding in the first month and each of the four 90-day periods will be summarized based on the adverse events of bleeding events which led to withdrawal of study drug. The cumulative number and percentage of patients discontinuing the study drug due to bleeding up to the end of each 90-day period (cumulative up to Day 90, Day 180, Day 270, and Day 364) will also be presented in the summary.

8.3.4. Mammogram

Patients who are ≥ 40 years of age at the time of enrollment or who will turn 40 during the trial will need to undergo a screening mammogram. Patients who are ≥ 40 years of age at the time of the end of treatment visit will need to undergo a mammogram at that time. By-patient data will be listed.

8.3.5. Questionnaires

Patient Health Questionnaire-9 (PHQ-9) will be used for depression screening at each in-person visit after screening (Visits 2-7). The Columbia Suicide Severity Rating Scale (C-SSRS) will be used to screen for suicidal ideation and behavior at the screening visit. Patients who report an answer of 1+ to the question of being bothered in the last two weeks of "Thoughts that you would be better off dead, or of hurting yourself in some way" in PHQ-9 must be further assessed with the C-SSRS. By-patient questionnaire data will be listed.

8.4. Changes From Protocol Specified Safety Analysis

Not applicable.

9. REFERENCES

Benda N, Gerlinger C, van der Meulen E, Endrikat J. Sample size calculation for clinical studies on the efficacy of a new contraceptive method. Biometrical Journal. 2004;46(1):141–50.

Creinin M, Vieira C, Westhoff C., Mansour D. Recommendations for standardization of bleeding data analyses in contraceptive studies. Contraception. 2022;112: 14-22.

FDA. Establishing Effectiveness and Safety for Hormonal Drug Products Intended to Prevent Pregnancy Guidance for Industry. 2019 Jul;9.

10. CHANGE HISTORY

SAP Version	Date	Corresponding protocol version	Changes (including section numbers) and reasons
Original Version: 1.0	22MAR2024	Protocol Amendment 4, Dated 21DEC2022	NA
Amendment 1	18AUG2025	Protocol Amendment 4, Dated 21DEC2022	Added general rules to present data in Section 4.3. Revised the analysis population for screen failure in Section 4.3 (Table 2). Added definition of start date of treatment cycle in Section 4.5.5 Added analysis time window for end-of-cycle data from eDiary in Section 4.5.6 to specify evaluable end-of-cycle data for the efficacy analysis Revised wording to match definition of start of treatment cycle (Section 4.5.5) for contraceptive efficacy in Section 4.5.7 Clarified the management of multiple eDiary entries in the same treatment cycle for a patient in Section 4.7.2. Added Section 4.8 to specify the management of duplicate patients for the analysis Added further details in Sections 5.1 (patient disposition) and 5.2 (screen failure). Added the summary by age group in Section 5.1. Revised typographical error in Section 5.3 to clarify that IPDs will be based on the ITT population, similar to Section 4.3, Table 2. Provided additional details for a few categories of demographic and baseline characteristics in Section 5.4. Added the summary by age group in Section 5.4. Added further detail for the definition of post-treatment medications and level of ATC classification system in Section 5.6. Further clarified drug compliance calculations in Section 6.2. Added the summary by age group in Section 6.2.

SAP Version	Date	Corresponding protocol version	Changes (including section numbers) and reasons
			Clarified cycles at-risk of pregnancy in Section 7.3.1. Clarified which additional statistical parameters will be summarized with the primary efficacy endpoint in Section 7.3.1. Moved "Primary Efficacy Endpoint by Group" to Section 7.3.3 and added summary by age and BMI groups. Revised and clarified the Method Failure PI definition in Section 7.4 Added the definition of the time to event in Section 7.4.2 Added new section "Secondary Efficacy Endpoints and Other Secondary Endpoints by Group" as Section 7.4.3. Added more details for the statistical parameters and statistical methods for BMD in Sections 8.3.2.1 and 8.3.2.2. Updated the Vitamin D categories to be consistent with protocol in Section 8.3.2.1 and Table 6. Revised bleeding categories to be summarized in Section 8.3.3 Added more details for the statistical methods for bleeding in Sections 8.3.3. Removed creatine kinase from Appendix 1 since it was not collected.

APPENDIX 1. LIST OF PRE-DEFINED LIMITS OF CHANGE IN SELECTED CHEMISTRY AND HEMATOLOGY TEST RESULTS

Note: if the SI units are the same as the conventional units, it is left blank. Please use the values (or limits) in the "Conventional units" column.	
Liver Function	
Conventional units	**SI units [a]**
ALT > ULN and < 3 × ULN	
ALT ≥ 3 × ULN and < 5 × ULN	
ALT ≥ 5 × ULN and < 10 × ULN	
ALT ≥ 10 × ULN and < 20 × ULN	
ALT ≥ 20 × ULN	
AST > ULN and < 3 × ULN	
AST ≥ 3 × ULN and < 5 × ULN	
AST ≥ 5 × ULN and < 10 × ULN	
AST ≥ 10 × ULN and < 20 × ULN	
AST ≥ 20 × ULN	
ALT or AST > ULN and < 3 × ULN	
ALT or AST ≥ 3 × ULN and < 5 × ULN	
ALT or AST ≥ 5 × ULN and < 10 × ULN	
ALT or AST ≥ 10 × ULN and < 20 × ULN	
ALT or AST ≥ 20 × ULN	
Total BILI > ULN	
Total BILI > 2 × ULN	
ALT or AST ≥ 3 × ULN and Total BILI > 2 × ULN	

Note: if the SI units are the same as the conventional units, it is left blank. Please use the values (or limits) in the "Conventional units" column.	
ALT or AST ≥ 3 × ULN and Total BILI > 2 × ULN and ALP < 2 × ULN	
GGT > ULN and < 3 × ULN	
GGT ≥ 3 × ULN and < 5 × ULN	
GGT ≥ 5 × ULN and < 10 × ULN	
GGT ≥ 10 × ULN and < 20 × ULN	
GGT ≥ 20 × ULN	
Renal Function	
Conventional Units	**SI units**
CR > 1.5 mg/dL and > BL	CR > 132.60 umol/L and > BL
CR > 50% increase from BL	
GFR < 15 mL/min per 1.73 m^2	
GFR ≥ 15 to < 30 mL/min per 1.73 m^2	
GFR ≥ 30 to < 60 mL/min per 1.73 m^2	
GFR ≥ 60 to < 90 mL/min per 1.73 m^2	
GFR ≥ 90 mL/min per 1.73 m^2	
Metabolic Parameters	
Conventional Units	**SI units**
Fasting Glucose	
< 100 mg/dL at BL	< 5.55 mmol/L at BL
< 100 mg/dL	< 5.55 mmol/L
≥ 100 to < 126 mg/dL	≥ 5.55 to < 6.99 mmol/L
≥ 126 mg/dL	≥ 6.99 mmol/L

Note: if the SI units are the same as the conventional units, it is left blank. Please use the values (or limits) in the "Conventional units" column.	
≥ 100 to < 126 mg/dL at BL	≥ 5.55 to < 6.99 mmol/L at BL
< 100 mg/dL	< 5.55 mmol/L
≥ 100 to < 126 mg/dL	≥ 5.55 to < 6.99 mmol/L
≥ 126 mg/dL	≥ 6.99 mmol/L
≥ 126 mg/dL at BL	≥ 6.99 mmol/L at BL
< 100 mg/dL	< 5.55 mmol/L
≥ 100 to < 126 mg/dL	≥ 5.55 to < 6.99 mmol/L
≥ 126 mg/dL	≥ 6.99 mmol/L
Highest Postbaseline Glucose	
Gluc ≥ 200 mg/dL and > BL	Gluc ≥ 11.1 mmol/L and > BL
Gluc ≥ 200 mg/dL and ≥ 126 mg/dL at BL	Gluc ≥ 11.1 mmol/L and ≥ 6.99 mmol/L at BL
Gluc ≥ 500 mg/dL and > BL	Gluc ≥ 27.75 mmol/L and > BL
Gluc ≥ 500 mg/dL and ≥ 126 mg/dL at BL	Gluc ≥ 27.75 mmol/L and ≥ 6.99 mmol/L at BL
Total CHOL > 200 mg/dL and > BL	Total CHOL > 5.18 mmol/L and > BL
Total CHOL increase > 30 mg/dL from BL	Total CHOL increase > 0.78 mmol/L from BL
Total CHOL Normal < 200 mg/dL and > BL	Total CHOL Normal < 5.18 mmol/L and > BL
Total CHOL Borderline High 200 to < 240 mg/dL and > BL	Total CHOL Borderline High 5.18 to < 6.22 mmol/L and > BL
Total CHOL High ≥ 240 mg/dL and > BL	Total CHOL High ≥ 6.22 mmol/L and > BL
LDL > ULN and > BL	
LDL Normal < 100 mg/dL and > BL	LDL Normal < 2.59 mmol/L and > BL
LDL Normal high 100 - < 130 mg/dL and > BL	LDL Normal high 2.59 - < 3.37 mmol/L and > BL
LDL Borderline High 130 - < 160 mg/dL and > BL	LDL Borderline high 3.37 - < 4.14 mmol/L and > BL
LDL High 160 - < 190 mg/dL and > BL	LDL High 4.14 - < 4.92 mmol/L and > BL
LDL Very high ≥ 190 mg/dL and > BL	LDL Very high ≥ 4.92 mmol/L and > BL

Note: if the SI units are the same as the conventional units, it is left blank. Please use the values (or limits) in the "Conventional units" column.	
HDL < LLN and < BL	
HDL Low < 40 mg/dL and < BL	HDL Low < 1.04 mmol/L and < BL
HDL High ≥ 60 mg/dL and < BL	HDL High ≥ 1.55 mmol/L and < BL
TRIG > ULN and > BL	
TRIG Normal <150 mg/dL and > BL	TRIG Normal < 1.69 mmol/L and > BL
TRIG Borderline high 150 - < 200 mg/dL and > BL	TRIG Borderline high 1.69 - < 2.26 mmol/L and > BL
TRIG High 200 - < 500 mg/dL and > BL	TRIG High 2.26 – < 5.65 mmol/L and > BL
TRIG Very High ≥ 500 mg/dL and > BL	TRIG Very High ≥ 5.65 mmol/L and > BL
Electrolytes and other Chemistry Parameters	
Conventional Units	**SI units**
ALB < LLN and < BL	
ALB > ULN and > BL	
ALP > 2 × ULN and > BL	
ALP > 5 × ULN and > BL	
ALP > 10 × ULN and > BL	
CA < LLN and < BL	
CA > ULN and > BL	
K < LLN and < BL	
K > ULN and > BL	
MG < LLN and < BL	
MG > ULN and > BL	
PHOS < LLN and < BL	
PHOS > ULN and > BL	

Note: if the SI units are the same as the conventional units, it is left blank. Please use the values (or limits) in the "Conventional units" column.	
NA < LLN and < BL	
NA> ULN and > BL	
Hematology Laboratory	
Conventional Units	**SI Units**
HCT < LLN and < BL	
HCT decrease ≥ 10% from BL	
HGB ≤ 10.5 g/dL and < BL	HGB ≤ 105 g/L and < BL
HGB decrease > 1 g/dL from BL	HGB decrease > 10 g/L from BL
MCV < LLN and < BL	
MCV > ULN and > BL	
WBC < LLN and < BL	
WBC > ULN and > BL	
LYM < LLN and < BL	
LYM > ULN and > BL	
MONO < LLN and < BL	
MONO > ULN and > BL	
NEUT < LLN and < BL	
NEUT > ULN and > BL	
BASO < LLN and < BL	
BASO > ULN and > BL	

Note: if the SI units are the same as the conventional units, it is left blank. Please use the values (or limits) in the "Conventional units" column.	
EOS < LLN and < BL	
EOS > ULN and > BL	
EOS > 5% and > BL	
PLT < LLN and < BL	
PLT < 100 × 10^3/uL and < BL	PLT < 100 × 10^9/L and < BL
PLT > ULN and > BL	
HbA1c ≤ 5.6% and > BL	
HbA1c 5.7 - 6.4% and > BL	
HbA1c ≥ 6.5% and > BL	
HbA1c increase > 1.0% from BL	

Abbreviations: ALB = albumin; ALP = alkaline phosphatase; ALT = alanine aminotransferase; AST = aspartate aminotransferase; BASO = basophils; BILI = bilirubin; BL = baseline; CA = calcium; CHOL = cholesterol; CK = creatine kinase; CR = creatinine; E2 = estradiol; EOS = eosinophils; GFR = glomerular filtration rate; GGT = gamma glutamyl transferase; Gluc = glucose; HbA1c = hemoglobin A1c; HCT = hematocrit; HDL = high density lipoprotein cholesterol; HGB = hemoglobin; K = potassium; LDL = low density lipoprotein cholesterol; LLN = lower limit of normal; LYM = lymphocytes; MCV = mean corpuscular volume; MG = magnesium; MONO = monocytes; NA = sodium; NETA = norethindrone acetate; NEUT = neutrophils; PHOS = phosphate; PLT = platelets; TRIG = triglycerides; ULN = upper limit of normal; WBC = white blood cell.

[a] Only list those different from conventional units.

Reference for conversion factors: *AMA Manual of Style*: *A Guide for Authors and Editors*. 11th ed. https://www.amamanualofstyle.com/page/si-conversion-calculator;jsessionid=8564876503a346c5c168bcd673e934c5;jsessionid=8564876503A346C5C168BCD673E934C5

Note: the conversion between conventional units and SI units are used ONLY for TFL display purpose. In the actual data sets conversion between conventional units and SI units, please use the strictest available conversional factors.

APPENDIX 2. PRE-DEFINED CATEGORIES FOR LAB SHIFT TABLES

Note: if the SI units are the same as the conventional units, it is left blank. Please use the values (or categories) in the "Conventional units" column.

Clinical Laboratory Parameter	Category in Conventional units	Category in SI units
Fasting Glucose/Glucose	Category 1 (< 100 mg/dL)	Category 1 (< 5.55 mmol/L)
	Category 2 (100 - < 126 mg/dL)	Category 2 (5.55 - < 6.99 mmol/L)
	Category 3 (126 - 200 mg/dL)	Category 3 (6.99 - 11.10 mmol/L)
	Category 4 (> 200 mg/dL)	Category 4 (> 11.10 mmol/L)
HbA1c	Normal (< 5.7%)	
	Prediabetes (5.7 – 6.4%)	
	Diabetes category 1 (6.5 – 8%)	
	Diabetes category 2 (8.1 – 9.4%)	
	Diabetes category 3 (9.5 – 11%)	
	Diabetes category 4 (> 11%)	
Total Cholesterol (NCEP ATP III Guidelines)	Desirable < 200 mg/ dL;	Desirable < 5.18 mmol/L;
	Borderline high 200 - < 240 mg/dL;	Borderline high 5.18 - < 6.22 mmol/L;
	High ≥ 240 mg/dL;	High ≥ 6.22 mmol/L;
Low Density Lipoprotein Cholesterol (NCEP ATP III Guidelines)	Normal (< 100 mg/dL)	Normal (< 2.59 mmol/L)
	Normal high (100 - < 130 mg/dL)	Normal high (2.59 - < 3.37 mmol/L)
	Borderline High (130 - < 160 mg/dL)	Borderline High (3.37 - < 4.14 mmol/L)
	High (160 - < 190 mg/dL)	High (4.14 - < 4.92 mmol/L)
	Very High (≥ 190 mg/dL)	Very High (≥ 4.92 mmol/L)
High Density Lipoprotein Cholesterol (NCEP ATP III Guidelines)	Category 1 (< 40 mg/dL)	Category 1 (< 1.04 mmol/L)
	Category 2 (40 - < 60 mg/dL)	Category 2 (1.04 - <1.55 mmol/L)
	Category 3 (≥ 60 mg/dL)	Category 3 (≥ 1.55 mmol/L)
Triglycerides (NCEP ATP III Guidelines)	Normal: TRIG < 150 mg/dL;	Normal: TRIG < 1.69 mmol/L;
	Borderline High: TRIG 150 - < 200 mg/dL;	Borderline High: TRIG 1.69 - < 2.26 mmol/L;
	High: TRIG 200 - < 500 mg/dL;	High: TRIG 2.26 - < 5.65 mmol/L;
	Very High: TRIG ≥ 500 mg/dL.	Very High: TRIG ≥ 5.65 mmol/L;

Reference for conversion factors: *AMA Manual of Style*: *A Guide for Authors and Editors*. 11th ed. https://www.amamanualofstyle.com/page/si-conversion-calculator;jsessionid=8564876503a346c5c168bcd673e934c5;jsessionid=8564876503A346C5C168BCD673E934C5

Note: the conversion between conventional units and SI units are used ONLY for TFL display purpose. In the actual data sets conversion between conventional units and SI units, please use the strictest available conversional factors.

docusign

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

Signer Events	Signature	Timestamp
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
In Person Signer Events	Signature	Timestamp
Editor Delivery Events	Status	Timestamp
Agent Delivery Events	Status	Timestamp
Intermediary Delivery Events	Status	Timestamp
Certified Delivery Events	Status	Timestamp
Carbon Copy Events	Status	Timestamp
Witness Events	Signature	Timestamp
Notary Events	Signature	Timestamp

Envelope Summary Events	Status	Timestamps
Envelope Sent	Hashed/Encrypted	8/18/2025 1:41:45 PM
Certified Delivered	Security Checked	8/18/2025 1:41:55 PM
Signing Complete	Security Checked	8/18/2025 1:42:39 PM
Completed	Security Checked	8/18/2025 9:57:53 PM

Payment Events	Status	Timestamps