16.1.9 Documentation of Statistical Methods

Statistical Analysis Plan, v1.0

STATISTICAL ANALYSIS PLAN

Study Titles:	An International Phase 3 Double-Blind, Placebo-controlled, Randomized Withdrawal Study of Relugolix Co-administered with Estradiol and Norethindrone Acetate in Women with Heavy Menstrual Bleeding Associated with Uterine Fibroids
Investigational Product:	Relugolix
Protocol Number:	MVT-601-035
Indication:	Treatment of heavy menstrual bleeding associated with uterine fibroids
Sponsor:	Myovant Sciences GmbH Viaduktstrasse 8 CH-4051 Basel Switzerland
Regulatory Identifier(s):	Eudra CT # 2018-001368-43 IND # 131161
Version/Effective Date:	15 Mar 2021

STATISTICAL ANALYSIS PLAN APPROVAL SHEET

An International Phase 3 Double-Blind, Placebo-controlled, Randomized Withdrawal Study of Relugolix Co-administered with Estradiol and Norethindrone Acetate in Women with Heavy Menstrual Bleeding Associated with Uterine Fibroids

This statistical analysis plan has been approved by Myovant Sciences GmbH. The following signatures document this approval.

PPD
PPD PPD

Date

PPD
PPD PPD

Date

PPD
PPD PPD

Date

PPD
PPD PPD

Date

PPD
PPD PPD

Date

PPD
PPD

Date

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Term	Definition/Explanation
ALP	alkaline phosphatase
ALT	alanine aminotransferase
ANCOVA	analysis of covariance
AST	aspartate aminotransferase
ATC	Anatomical Therapeutic Chemical
BMD	bone mineral density
BMI	body mass index
BPD	bleeding and pelvic discomfort
bpm	beats per minute
CI	confidence interval
CSR	clinical study report
CTCAE	common terminology criteria for adverse events
DXA	dual-energy x-ray absorptiometry
E2	estradiol
ECG	electrocardiogram
eCRF	electronic case report form
EOT	end of treatment
FP	feminine product
FPRR	feminine product return rate
ICH	International Council on Harmonisation
LLN	lower limit of normal
LOCF	last observation carried forward
LS	least squares
MBL	menstrual blood loss
MCS	mental component summary
mITT	modified Intent-to-Treat
MedDRA	Medical Dictionary for Regulatory Activities
mmHg	millimeters of mercury
msec	millisecond
NETA	norethindrone acetate
NRS	numerical rating scale
LTE	long-term extension
PCS	physical component summary
PGA	Patient Global Assessment
PT	Preferred Term
SAP	statistical analysis plan
SD	standard deviation
SF-36v2	Short Form (36) version 2
SMQ	standard MedDRA query
SOC	System Organ Class

Term	Definition/Explanation
UFS-QoL	Uterine Fibroid Symptom and Health-Related Quality of Life (Questionnaire)
ULN	upper limit of normal
VAS	visual analogue score
WHO	World Health Organization
WPAI-UF	Work Productivity Activity Impairment-Uterine Fibroids

1. INTRODUCTION

The purpose of this statistical analysis plan (SAP) is to describe the efficacy and safety analyses planned for phase 3 study MVT-601-035, entitled "An International Phase 3 Double-Blind, Placebo-controlled, Randomized Withdrawal Study of Relugolix Co-Administered with Estradiol and Norethindrone Acetate in Women with Heavy Menstrual Bleeding Associated with Uterine Fibroids." Patients who completed the 24-week treatment period in a pivotal study (study MVT-601-3001 or study MVT-601-3002) and the 28-week treatment period of the long-term extension (LTE) (MVT-601-3003) study and met all eligibility criteria for study MVT-601-035 (including met the definition of responder) are enrolled in MVT-601-035. Enrolled patients are randomized to one of two treatment groups: once daily oral relugolix 40 mg with estradiol/norethindrone acetate (E2/NETA) 1 mg/0.5 mg for 52 weeks (referred to as the relugolix + E2/NETA group) or placebo for 52 weeks (referred to as the placebo group).

This SAP was developed in accordance with the International Council on Harmonisation (ICH) E9 guidelines. All decisions regarding statistical analysis of the study, as defined in this SAP, will be made prior to unblinding of the study data.

This SAP is based on:

- Protocol MVT-601-035, Amendment 1, dated 13 Aug 2018;
- ICH guidelines E3 (Clinical Study Reports) and E9 (Statistical Principles for Clinical Trials).

The methods used for analysis of the MVT-601-035 study are consistent with those used for the phase 3 studies (MVT-601-3001, MVT-601-3002 and MVT-601-3003). The SAP will be finalized, approved by the sponsor, and placed on file before the database is locked. Changes to the final approved plan will be noted in the clinical study report (CSR).

1.1. Study Objectives and Endpoints

The objectives of Study MVT-601-035 are to evaluate the long-term efficacy and safety of relugolix +E2/NETA, once daily, for up to 104 weeks in patients with uterine fibroids who have completed a total of 52 weeks of treatment including 24-week treatment period in a pivotal study (study MVT-601-3001 or study MVT-601-3002) and 28-week treatment period of the LTE study (study MVT-601-3003) and who meet the definition of responder. A responder is defined as a patient who demonstrates a menstrual blood loss of < 80 mL and at least a 50% reduction from pivotal study baseline menstrual blood loss volume on the alkaline hematin analysis of the feminine products (FPs) returned at Week 48 in the LTE study. The study objectives and endpoints are listed in Table 1. The endpoints in italics are not listed in the protocol, but they have been identified as important for the assessment of treatment effect based on emerging data and clinical relevance to the study objectives and therefore are prespecified in the SAP. The endpoints listed in Table 1 will also be analyzed in general using data collected during the combination treatment period (ie, while patients are on randomized treatment or re-treatment), unless otherwise specified (see details in Section 7, Section 8, and Section 9).

Primary and secondary assessment of efficacy and safety assessments will be made in the following two study treatment groups:

- 52 weeks of a relugolix tablet and an over-encapsulated tablet of E2/NETA, once daily;
- 52 weeks of a placebo relugolix tablet and a capsule of placebo E2/NETA acetate, once daily.

Week 52 of the LTE study will define the baseline for this randomized withdrawal study and will be used as the reference point for all changes from baseline-related endpoints unless otherwise specified. The Week 48 menstrual blood loss volume in the LTE study will establish responder status (results from Week 44 may be used if Week 48 data are unavailable at the time eligibility was being assessed).

Table 1: Study Objectives and Endpoints

Objective(s)	Endpoint(s)
Primary Efficacy	
• To evaluate the long-term effect of relugolix + E2/NETA, once daily, compared with placebo on menstrual blood loss at Week 76 (24 weeks after randomization).	• Proportion of women who maintained a menstrual blood loss volume of < 80 mL at Week 76 (24 weeks of the Randomized treatment period) as measured by the alkaline hematin method (referred to as sustained responder rate at Week 76), defined as the cumulative probability of MBL volume < 80 mL while on randomized treatment through Week 76.
Key Secondary Efficacy	
• To evaluate the long-term effect of relugolix + E2/NETA compared with placebo on the following	The following key secondary endpoints will be assessed at Week 76, at Week 104 and by visit (as appropriate) during the randomized treatment period:
• Resumption of heavy menstrual bleeding (HMB)	• Time to menstrual blood loss volume ≥ 80 mL;
• Menstrual blood loss	• Proportion of women who maintained a menstrual blood loss volume of < 80 mL at Week 104 (52 weeks of the randomized treatment period) as measured by the alkaline hematin method (named as sustained responder rate at Week 104), defined as the cumulative probability of MBL volume < 80 mL while on randomized treatment through Week 104;
• Achievement of amenorrhea	• Proportion of women achieving or maintaining amenorrhea at Week 76/EOT;

Objective(s)	Endpoint(s)
Other Secondary Efficacy	
• To evaluate the long-term effect of relugolix + E2/NETA, once daily compared with placebo on menstrual blood loss at Week 104 (52 weeks after randomization).	• Change from Week 52/Baseline to Week 76 and Week 104 in menstrual blood loss volume (during the randomized treatment period); • Percentage change from Week 52/Baseline to Week 76 and Week 104 in menstrual blood loss volume (during the randomized treatment period); • Proportion of women achieving or maintaining amenorrhea at Week 104/EOT.
• To evaluate the effect of retreatment with relugolix + E2/NETA on menstrual blood loss in patients whose menstrual blood volume returned to ≥ 80 mL during the 52-week randomized period.	• Proportion of patients who responded (menstrual blood loss volume of < 80 mL) to retreatment during the retreatment period among placebo patients whose menstrual blood loss volume returns to ≥ 80 mL during the 52-week randomized treatment period.
• To evaluate the long-term effect of relugolix + E2/NETA at 52 weeks after randomization on the following:	The following secondary endpoints will be assessed at Week 76, at Week 104 and by visit (as appropriate) during the randomized treatment period:
• Resumption of menses	• Proportion of women whose menses has resumed (among those who were amenorrheic at Week 52/Baseline); • Time to resumption of menses (among those who were amenorrheic at Week 52/Baseline).
• Resumption of heavy menstrual bleeding (HMB)	• Proportion of women with menstrual blood loss volume ≥ 80 mL at any timepoint during the 52-week randomized treatment period;
• Hemoglobin	• Change *and percent change* from Week 52/Baseline in hemoglobin; • *Proportion of patients who had hemoglobin level <=10.5g/dL over time* • *Proportion of patients who had hemoglobin level < 11.6g/dL over time*
• Health-related Quality of Life as measured by the Short Form (36) Health (SF-36) questionnaire	• Change from Week 52/Baseline in SF-36 domain and summary component scores;
• Patient Global Assessment (PGA) for function and symptoms	• Change from Week 52/Baseline in PGA for function and symptoms score;
• Work and Productivity impact as measured by the Work Productivity Activity Impairment-Uterine Fibroids (WPAI-UF) questionnaire	• Change from Week 52/Baseline in the WPAI-UF scores;

Objective(s)	Endpoint(s)
• Disease-specific quality of life as assessed by the Uterine Fibroid Symptom and Health-Related Quality of Life (UFS-QoL) questionnaire	• Change from Week 52/Baseline in the UFS-QoL scale and sub-scale scores as well as the total score.
Pharmacodynamic	
• To characterize the pharmacodynamic (PD) effect of withdrawal from relugolix + E2/NETA	• Pre-dose concentration of estradiol at Week 56
Safety	
• To evaluate the safety and tolerability of relugolix + E2/NETA, once daily, for up to an additional 52 weeks	The following safety endpoints will be assessed:
• Adverse events	• Incidence of adverse events; • Change in vital signs; • Clinical laboratory tests.
• Changes in bone mineral density	• Percent change from Week 52/Baseline to Week 104 in bone mineral density at the lumbar spine (L1-L4), femoral neck, and total hip as assessed by DXA. • *Change from Week 52/Baseline to Week 104 in bone mineral density Z-score at the lumbar spine (L1-L4), femoral neck, and total hip as assessed by DXA.* • *Percent change from pivotal study Baseline to Week 104 in bone mineral density at the lumbar spine (L1-L4), femoral neck, and total hip as assessed by DXA.*

Abbreviations: DXA = dual-energy x-ray absorptiometry; E2 = estradiol; EOT = End of Treatment; HMB = heavy menstrual bleeding; MBL = menstrual blood loss; NETA = norethindrone acetate; PGA = Patient Global Assessment; UFS-QoL = Uterine Fibroid Symptom and Health-Related Quality of Life; WPAI-UF = Work Productivity Activity Impairment-Uterine Fibroids.

2. STUDY DESIGN

2.1. Summary of Study Design

This randomized withdrawal study (MVT-601-035) is an international phase 3 double-blind, placebo-controlled study that enrolled eligible patients with uterine fibroids who have completed the 24-week treatment period in a pivotal study (study MVT-601-3001 or study MVT 601-3002) and the 28-week treatment period of the LTE study, MVT-601-3003. When including treatment during a pivotal study and the extension study, patients completing this randomized withdrawal study will have received up to a total of 104 weeks of treatment with relugolix.

All consenting patients completing MVT-601-3003 with a response to treatment were eligible to participate. A responder was defined as a patient who demonstrated a menstrual blood loss of < 80 mL and at least a 50% reduction from a pivotal study baseline menstrual blood loss volume on the alkaline hematin analysis of the FPs returned at Week 48 in MVT-601-3003. The objectives of the current study are to evaluate long-term efficacy and safety of treatment with relugolix + E2/NETA for up to 104 weeks (total treatment duration includes the pivotal study, the extension study, and the randomized withdrawal study). Approximately 360 patients completing the extension study and meeting the definition of a responder as well as all other eligibility criteria were expected to be randomized 1:1 to blinded treatment with oral relugolix + E2/NETA or placebo, once daily, for up to 52 weeks. For patients whose menstrual bleeding returned to ≥ 80 mL (heavy menstrual bleeding), treatment with open-label relugolix + E2/NETA was to be restarted.

Screening procedures were to be done on the same day as the Week 52 visit for MVT-601-3003. This visit will be referred to as the "Week 52/Baseline visit".

Patients were to receive the last dose of study drug in the study MVT-601-3003 on the day of the Week 52/Baseline visit (per the LTE study protocol) and to be dispensed study drug for this randomized withdrawal study in the clinic after being determined eligible for this study and providing informed consent to participate. The first dose of study drug for the randomized withdrawal study was to be self-administered the next morning following the Week 52/Baseline visit. Dosing was to be confirmed by telephone.

Patients were to be asked to collect FPs for alkaline hematin analysis and to return them at each visit until the analysis confirmed the return of heavy menstrual bleeding (defined as menstrual blood loss volume of ≥ 80 mL). The patients were then to be offered re-treatment with open-label relugolix +E2/NETA with the onset of the next menses. They were to continue collection of FPs for alkaline hematin analysis until two consecutive analyses confirmed resolution of heavy menstrual bleeding (menstrual blood loss of < 80 mL). Patients who experienced a return of heavy menstrual bleeding (or terminate early because they have a return of heavy menstrual bleeding) but who decline retreatment with open-label study drug were to complete an Early Termination visit and a 30-day Safety Follow-Up visit.

At the Week 104/Early Termination visit, all other patients were to have an assessment of bone mineral density (BMD) via dual-energy x-ray absorptiometry (DXA). Patients were to complete the bleeding diary daily, including compliance with study drug, vaginal bleeding, and use of FPs.

If a patient enrolls directly into another relugolix clinical study upon completion of the Week 104/Early Termination visit, then the Safety Follow-Up visit and the follow-up procedures performed under this protocol may be waived.

2.2. Sample Size Considerations

Because this is an extension of study MVT-601-3003, the sample size of this study was based on approximately the number of patients who completed a pivotal study (n = 780 patients in total), and who participated in the LTE study (n = 600 patients in total assuming about 20% dropout rate at 6 months from the pivotal studies). Assuming the proportion of responders at Week 52 in MVT-601-3003 LTE study is 60%, it was estimated that approximately 360 patients were to be randomized into this study.

With 360 patients, the study would have at least 90% power to detect a difference of 20% or greater between relugolix+ E2/NETA group and placebo group for the primary endpoint. The following assumptions were used to determine the sample size for this study:

- 2-sided type I error rate: 0.05
- Randomization: 1:1
- Responder rate for placebo group: 40%
- Responder rate for relugolix + E2/NETA group: 60%
- Dropout rate: 20%

3. PLANNED ANALYSES

3.1. Interim Analyses

There are no planned interim efficacy analyses.

3.2. Final Analyses

The final analysis of efficacy and safety data will occur after approximately 360 patients (projected to be responders from MVT-601-3003) who have been randomized and have had the opportunity to complete the study through either the Week 104 visit or the 30-day Safety Follow-up visit.

3.3. Safety Follow-Up Analyses

Patients who experience a BMD loss of ≥ 7% from the pivotal study baseline at any of the anatomical sites assessed will be discontinued from the study and will undergo another bone densitometry scan as described below. Patients with a BMD loss of ≥ 7% at lumbar spine, total hip, or femoral neck at their Week 104/Early Termination visit relative to pivotal study baseline measurement will undergo another bone densitometry scan at 6 (± 1) months after the last dose of study drug and will be contacted to obtain information about medications and conditions (eg, pregnancy) that might affect BMD through the time of the follow-up bone densitometry scan. Patients should be assessed for secondary causes of bone loss and followed-up further if the 6-month follow-up scan does not show improvement, unless an alternative etiology has been identified. The follow-up bone densitometry scan will be submitted for central reading.

Patients whose menses have not resumed as of the safety follow-up visit and for whom there is no explanation for the lack of resumption (eg, medical procedure or medications) will be contacted again by telephone 12 (± 2) weeks after the Safety Follow-up visit to determine if menses have resumed.

Complete data collected during the additional safety follow-up period will be summarized and reported in an addendum to the CSR.

4. GENERAL CONSIDERATIONS FOR DATA ANALYSES AND HANDLING OF MISSING DATA

4.1. Data Presentation Conventions

All statistical analyses will be conducted using SAS® Version 9.2 or higher.

All confidence intervals (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 treatment group. 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:

- 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 (eg, > 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 < 0.0001 will be presented as "< 0.0001" and p-values > 0.9999 will be presented as "> 0.9999";
- 1 month = 30.4375 days. Month is calculated as (days/30.4375) rounded to one decimal place;
- 1 year = 365.25 days. Year is calculated as (days/365.25) rounded to one decimal place;
- Age will be calculated using the date of randomization. If only year of birth is collected, 1 July of the year of birth will be used to calculate age.
- 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 analyses;
- For MBL volume reported as below the limit of quantification (eg, MBL below Quantification Level < 5.0 mL or < 2.5 mL), 0.0000000001 will be subtracted from the reported quantification threshold for the visit to derive a numeric result for analyses;
- For safety analyses, calculation of percentages will be calculated on the basis of the number of patients in the analysis population in each treatment group;
- For by-visit observed data analyses, calculation of percentages will be calculated on the basis of 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, and range (minimum and maximum);
- For time-to-event endpoints, the summary statistics will include median time to event-free survival, 25th and 75th percentiles and 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.

4.2. Analysis Populations

Four analysis populations are defined below. Number and percent of patients meeting the definition of each analysis population will be summarized by treatment group.

4.2.1. Modified Intent-to-Treat Population

Efficacy analyses will be performed using the modified Intent-to-Treat (mITT) population, unless otherwise specified. The mITT population consists of all patients randomized to treatment who have taken at least one dose of study treatment (relugolix/placebo or E2/NETA/placebo). Efficacy analyses will be performed by treatment group as randomized.

4.2.2. Per-Protocol Population

The Per-Protocol population consists of those patients in the mITT population who have no relevant major protocol deviations, defined as a subset of all major protocol deviations (see Section 5.2).

The Per-Protocol population will not be analyzed if this population comprises > 95% or < 50% of the mITT Population. This population will be used for supportive analysis of the primary efficacy endpoint. The Per-Protocol population will be identified prior to breaking the study blind.

4.2.3. Safety Population

The Safety population consists of all patients who were randomized and took at least one dose of study treatment. All safety analyses will be performed using the safety population, unless otherwise specified. Safety data will be analyzed by treatment group according to the actual treatment received (not the randomized treatment). Any patient who received at least one dose of relugolix during the randomized treatment period will be considered as a relugolix patient.

4.2.4. Retreatment Population

The Retreatment population consists of those patients in the mITT Population who started the retreatment. Selected efficacy and safety data will be analyzed using the retreatment population.

4.3. Definitions, Computation, and Convention

4.3.1. Definition of Date of First Dose and Date of Last Dose of Study Drug During the Randomized Treatment Period

The randomized treatment period is the period when a patient was on one of the randomized study drugs (relugolix + E2/NETA or placebo). The date of first dose of study drug is defined as the date when a patient receives the first dose of study drug for the randomized withdrawal study following randomization. The first dose of study drug for the randomized withdrawal study was to be self-administered the next morning following the Week 52/Baseline visit.

The date of the last dose of randomized study drug is defined as the date a patient receives the last dose of randomized study drug in this randomized withdrawal study, prior to re-treatment or on the Week 104 visit date. If the complete date of last dose of study drug is unknown, the last date the study drug was known to have been taken will be used.

4.3.2. Definition of Date of First Dose and Date of Last Dose of Retreatment Drug

The date of first dose of retreatment drug is defined as the date when a patient receives the first dose of retreatment drug (open-label relugolix + E2/NETA). When the return of heavy menstrual bleeding (defined as menstrual blood loss of ≥ 80 mL) was confirmed for a patient during the randomized treatment period, the patient was to be offered retreatment with open-label relugolix + E2/NETA with the onset of the next menses.

The date of the last dose of retreatment drug is defined as the date a patient receives the last dose of retreatment drug in this randomized withdrawal study, prior to or on the Week 104 visit date. If the complete date of last dose of retreatment drug is unknown, the last date the retreatment drug was known to have been taken will be used.

4.3.3. Study Day

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

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

For assessments conducted before the date (and time) of the first dose of randomized withdrawal study drug, study day will be calculated as:

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

For patients who do not receive any amount of randomized withdrawal study drug, study day will be calculated as above with respect to the date of randomization.

4.3.4. Definition of Treatment Duration During the Randomized Treatment Period

Treatment duration during the randomized treatment period 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 study drug was known to have been taken will be used to calculate treatment duration. For patients who did not return for the Early Termination visits during the randomized treatment period, the time after their last visit will not be included in calculations of treatment duration.

4.3.5. Definition of Treatment Duration During the Retreatment Period

The retreatment period is the period when a patient was on the open-label retreatment drug. Treatment duration of retreatment drug is defined as the duration of time from the date of the first dose of retreatment drug to the date of the last dose of retreatment drug as follows:

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

For patients without complete date of last dose of retreatment drug, the last date retreatment drug was known to have been taken will be used to calculate treatment duration. For patients who did not return for the Early Termination visits during the retreatment period, the time after their last visit will not be included in calculations of treatment duration.

The cumulative treatment period includes both the randomized treatment period and the retreatment period.

4.3.6. 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 in the randomized withdrawal study. This is for randomized treatment period and for cumulative treatment period, which includes the randomized treatment and retreatment periods. A post-Baseline value is defined as a measurement taken after the first dose of study drug. Change from Baseline is defined as (post-Baseline value – Baseline value). Both date and time of study drug administration and measurement will be considered when calculating Baseline value. If the time is not available, then the date alone will be used. For patients who receive no study medication, the date of randomization will be used in place of the date of first dose in determining Baseline and post-Baseline values.

Baseline for the retreatment period is defined as the last available measurement before the first dose of retreatment drug unless otherwise specified. A post-retreatment-baseline value is defined as a measurement taken after the first administration of retreatment drug. Change from retreatment baseline is defined as (post-retreatment-baseline value – retreatment baseline value).

The BMD measurement taken at Week 52 of the extension study will be used as the Week 52/Baseline in the randomized withdrawal study, as long as it was within 30 days of the first dose date of the randomized withdrawal study drug. If there are multiple BMD measurements at Week 52 in the extension study, the one closest to the first dose date of the randomized withdrawal study drug will be used as Week 52/Baseline in the randomized withdrawal study.

For menstrual blood loss volume, the baseline definition will be specified in Section 7.3.2.

4.3.7. Visit Windows

Visit windows, which will be used to associate assessments with a scheduled visit, will be used only for summarizing data by visit. The windows for scheduled assessments are shown in

Table 2 (for monthly assessment, SF-36v2, PGA, WPAI-UF), Table 3 (for BMD), Table 4 (for UFS-QoL), and Table 5 (for clinical laboratory tests and vital signs), respectively. For both efficacy and safety assessments, the study day will be used to determine the associated visit window.

The data collected in the paper diary related to bleeding and use of FPs will be assigned to visit windows as specified in Table 6 and will be used to calculate the feminine product return rate (FPRR) as specified in Section 7.3.2.5.

If the results from more than one assessment are within a given visit window, the non-missing result from the assessment closest to the target date will be used except for FP collection, which is specified in Section 7.3.2.4. If two assessments are equally close to the target day, the earlier assessment will be used. For summaries of shift from Baseline in safety parameters, all values will be considered for these analyses.

Table 2: Visit Windows for Monthly Assessments

Visit (Long label)	Visit (Short label)	Start Day	Target Day	End Day
Week 56 (4 weeks after randomization) [a]	Week 56	1	29	43
Week 60 (8 weeks after randomization)	Week 60	44	57	71
Week 64 (12 weeks after randomization)	Week 64	72	85	99
Week 68 (16 weeks after randomization)	Week 68	100	113	127
Week 72 (20 weeks after randomization)	Week 72	128	141	155
Week 76 (24 weeks after randomization)	Week 76	156	169	183
Week 80 (28 weeks after randomization)	Week 80	184	197	211
Week 84 (32 weeks after randomization)	Week 84	212	225	239
Week 88 (36 weeks after randomization)	Week 88	240	253	267
Week 92 (40 weeks after randomization)	Week 92	268	281	295
Week 96 (44 weeks after randomization)	Week 96	296	309	323
Week 100 (48 weeks after randomization)	Week 100	324	337	351
Week 104 (52 weeks after randomization)	Week 104	352	365	Date of last dose + 14 days [b]

[a] Start day of Week 56 for study day 1 includes only post-Baseline assessments that occurred after the first dose.
[b] For SF-36v2, PGA, WPAI-UF, 14 days after date of last dose is used as the end day of Week 104.

Table 3: Visit Windows for Bone Mineral Density Assessments

Visit (Long label)	Visit (Short label)	Start Day	Target Day	End Day
Week 104 (52 weeks after randomization)	Week 104	281	365	449

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. Assessments after date of last dose of study drug + 60 days will also be excluded from analysis.

Table 4: Visit Windows for UFS-QoL Assessments

Visit (Long label)	Visit (Short label)	Start Day	Target Day	End Day
Week 64 (12 weeks after randomization)	Week 64	64	85	106
Week 76 (24 weeks after randomization)	Week 76	148	169	196
Week 88 (36 weeks after randomization)	Week 88	197	253	308
Week 100 (48 weeks after randomization)	Week 100	309	337	351
Week 104 (52 weeks after randomization)	Week 104	352	365	421

Abbreviations: UFS-QoL = Uterine Fibroid Symptom and Health-Related Quality of Life.

Table 5: Visit Windows for Clinical Laboratory Tests and Vital Signs

Visit (Long label)	Visit (Short label)	Start Day	Target Day	End Day
Week 64 (12 weeks after randomization) [a]	Week 64	64	85	106
Week 76 (24 weeks after randomization)	Week 76	148	169	196
Week 104 (52 weeks after randomization)	Week 104	309	365	Date of last dose + 6 days
Safety Follow-Up [b]	Safety Follow-Up	Date of last dose + 7 days	Date of last dose + 30 days	Date of last dose + 60 days

[a] This visit is only for clinical laboratory tests.

[b] The safety follow-up visit window will be restricted to assessments prior to the date of initiation of another investigational agent or hormonal therapy affecting the hypothalamic-pituitary-gonadal axis or surgical intervention for uterine fibroids.

Table 6: Time Window for Feminine Product Collection

Visit (Long label)	Visit (Short label)	Feminine Product Collection Visit Date[a,b]	Time Window[a]
Week 56 (4 weeks after randomization) [a]	Week 56	X_1	(Date of Study Day 1) – < X_1
Week 60 (8 weeks after randomization)	Week 60	X_2	(X_1 +1) – ≤ X_2
Week 64 (12 weeks after randomization)	Week 64	X_3	(X_2 +1) – ≤ X_3
Week 68 (16 weeks after randomization)	Week 68	X_4	(X_3 +1) – ≤ X_4
Week 72 (20 weeks after randomization)	Week 72	X_5	(X_4 +1) – ≤ X_5
Week 76 (24 weeks after randomization)	Week 76	X_6	(X_5 +1) – ≤ X_6
Week 80 (28 weeks after randomization)	Week 80	X_7	(X_6 +1) – ≤ X_7
Week 84 (32 weeks after randomization)	Week 84	X_8	(X_7 +1) – ≤ X_8
Week 88 (36 weeks after randomization)	Week 88	X_9	(X_8 +1) – ≤ X_9
Week 92 (40 weeks after randomization)	Week 92	X_{10}	(X_9 +1) – ≤ X_{10}
Week 96 (44 weeks after randomization)	Week 96	X_{11}	(X_{10} +1) – ≤ X_{11}
Week 100 (48 weeks after randomization)	Week 100	X_{12}	(X_{11} +1) – ≤ X_{12}
Week 104 (52 weeks after randomization) /EOT	Week 104/EOT	X_{Last}[c]	(Previous Feminine Product Returned Visit +1) – ≤ X_{Last}

[a] If feminine products are collected at more than 1 visit within a given visit window (Table 2), the last feminine product collection date will be used to define the time window. If the patient missed the previous visit, a planned study visit date will be used to calculate the window.

[b] In the absence of feminine product collection due to amenorrhea the visit date when amenorrhea was reported will be used.

[c] Date of last non-missing feminine product collection within the interval from (previous visit end date + 1) to (last dose date + 7 days) (see Section 7.3.2).

4.4. General Rules for Missing Data

Handling of missing data for the primary efficacy analysis is described in Section 7.3.3.

4.4.1. 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.

4.4.2. 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 dates of the treatment-emergent period to determine if the event is treatment emergent. When in doubt, the adverse event will be considered treatment emergent by default. If the adverse events were recorded in pivotal or LTE studies, the rules used to impute the partial dates in corresponding study will be used.

The following rules will be applied to impute partial dates for adverse events when they were newly recorded in the randomized withdrawal study:

- 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, then set to treatment start date, as long as adverse event end date is not prior to treatment start date;
  - If both Month and Day are missing and Year ≠ Year of treatment start date, then set to January 1;
  - If Day is missing and Month and Year = Month and Year of treatment start date, then set to treatment start date, as long as adverse event end date is not prior to treatment start date;
  - If Day is missing and Month and Year ≠ Month and Year of treatment start date, then set to first of the month;
  - If start date is completely missing, set to treatment start date, as long as 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, then set to December 31;
  - If only Day is missing, then 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, then set to January 1;
  - If only Day is missing, then 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, then set to December 31;
  - If only Day is missing, then set to last day of the month.
- If start date or end date of a medication is completely missing, do not impute.

5. STUDY POPULATION

5.1. Subject Disposition

The number of patients for each of the following categories will be summarized by treatment group:

- All randomized patients;
- Patients included in the Safety population;
- Patients who completed 76-week treatment;
- Patients who completed 104-week treatment;
- Patients who discontinued early from the study and reasons for discontinuation;

Patient disposition will be summarized for all randomized patients. Summaries will include the number and percentage of patients in the mITT and Safety populations. The number and percentage of patients who prematurely discontinue study drug and the reasons for discontinuation will be summarized by treatment group.

5.2. Protocol Deviations

Protocol deviations that occurred during the randomized withdrawal study will be categorized as important or minor per the protocol deviation plan. Important protocol deviations will include, but will not be limited to, the following categories:

- Randomized patient who did not satisfy key entry criteria;
- Randomized patient who met withdrawal criteria during the study but was not withdrawn;
- Randomized patient who received the wrong treatment;
- Randomized patient who received a prohibited concomitant medication that met criteria for an important protocol deviation;
- Unintentional unblinding of treatment assignment.

Important protocol deviations will be summarized by deviation category for all patients in the mITT population. A patient listing of all important protocol deviations will be provided.

A selected subset of the major protocol deviations that are likely to affect analysis of efficacy will be identified to define the Per-Protocol population prior to the database lock. This subset will include but will not be limited to the following important protocol deviations:

- Did not satisfy key entry criteria (restricted to patients who did not meet inclusion criterion 4 of MVT-601-035, ie, patients who were not responders at the time eligibility was being assessed);
- Drug compliance during the randomized treatment period < 75%;
- Patient received prohibited concomitant medications that met criteria for important protocol deviation: restricted to patients who received prohibited concomitant medications that may cause significant drug-drug interaction;

- Unintentional unblinding of treatment assignment.

5.3. Demographic and Baseline Characteristics

Demographic and Baseline characteristics at pivotal study baseline (Pivotal Baseline) and at this randomized withdrawal study baseline (Week 52/Baseline) will be summarized by treatment group for the mITT population. Categorical data will be summarized using frequencies and percentages, by treatment group and overall (see Table 7 below) for each baseline. Summaries of continuous data will display the mean, SD, median, minimum, and maximum. The numbers of missing values will also be summarized.

Table 7: Categories for Demographic and Baseline Characteristics

Variable	Category
For both Pivotal Baseline and Week 52/Baseline	
Age (years)	< 40, ≥ 40
BMI (kg/m^2)	< 18.5, 18.5 to <25, 25 to <30, 30 to < 35, 35 to < 40, ≥ 40 < 30, ≥ 30
UFS-QoL Bleeding and Pelvic Discomfort Scale	0 to < 25, 25 to <50, 50 to <75, 75 to 100
Patient Global Assessment	
Function	No limitation at all, Mild limitation, Moderate limitation, Quite a bit of limitation, Extreme limitation
Symptoms	Not severe, Mildly severe, Moderately severe, Very severe, Extremely severe
Hemoglobin (g/dL)	Min to < 8, ≥ 8 to <10.5, ≥ 10.5 to <12, ≥ 12
For Week 52/Baseline Only	
Geographic region	North America, Other North America, Europe, Latin America, Rest of World
Duration of relugolix exposure prior to randomization (weeks)	28, 52
For Pivotal Baseline Only	
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
Menstrual blood loss volume at baseline (mL)	< 225, ≥ 225
Menstrual blood loss volume at baseline (mL)	< 160, ≥ 160

Variable	Category
History of prior pregnancy	Yes, No
Disease duration of uterine fibroid (years)	Min to <1, ≥ 1 to < 3, ≥3 to <5, ≥5 to <10, ≥ 10
Type of uterine fibroids	
Subserous fibroid	Yes, No
Intramural fibroid	Yes, No
Submucosal fibroid	Yes, No
Other	Yes, No
Any surgery for uterine fibroids	Yes, No
Volume of myoma at Pivotal Baseline (cm^3)	< 25, ≥ 25
Volume of uterus at Pivotal Baseline (cm^3)	< 300, ≥ 300
Maximum NRS score for uterine fibroid-associated pain at Pivotal Baseline	< 4, ≥ 4
Alcohol use	None, Moderate (2 - 6 drinks per week), Heavy (≥ 7 drinks per week)
Smoking history	Current smoker, Past smoker, Never

Abbreviations: BMI = body mass index; NRS = Numerical Rating Scale; UFS-QoL = Uterine Fibroid Symptom and Health-Related Quality of Life.

5.4. Medical History

Medical history collected at the time of entry into the pivotal studies will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) and will be summarized by system organ class (SOC) and preferred term (PT) for the Safety population. Additionally, summaries of uterine fibroid–specific medical and surgical treatment history collected at the time of entry into the pivotal study will be provided. A patient with multiple occurrences of medical history within a PT will be counted only once in that PT. Medical history is not collected in the eCRF of randomized withdrawal study. Adverse events that were collected in the pivotal studies or extension study and ended prior to the start of the randomized withdrawal study will be listed and summarized.

5.5. Prior Medications and Concomitant Medications

Prior medications and concomitant medications taken during this Randomized Withdrawal Study treatment period will be summarized for all patients in the Safety population by randomized treatment group. In addition, concomitant medications taken during the retreatment period will also be summarized for Safety Population by their randomized treatment group. Medications are considered concomitant for randomized treatment period if exposure occurs during the randomized treatment period. Medications are considered concomitant for retreatment period if exposure occurs during the retreatment period. Medications are considered prior if exposure started prior to the date of first dose of randomized withdrawal study drug.

The number and percentage of patients who took at least one dose of a prior or concomitant medication will be summarized by randomized treatment group and overall using the World Health Organization (WHO) Drug Dictionary and summarized according to the Anatomical Therapeutic Chemical (ATC) Classification System and generic medication name. A patient who has been administered several medications with the same preferred medication name will be counted only once for that preferred medication name.

6. STUDY DRUG EXPOSURE AND COMPLIANCE

All analyses of extent of exposure to study drug and compliance will be summarized by actual treatment received in this randomized withdrawal study. Extent of exposure to and compliance with relugolix (or relugolix placebo) and E2/NETA (or placebo) during this randomized treatment period and the retreatment period will be summarized for patients in the Safety population. Exposure to and compliance with relugolix and E2/NETA will be summarized separately and will be based on the drug accountability case report forms.

Study drug exposure summaries will include the total dosages of relugolix and E2/NETA taken in milligrams, the total number of tablets (or capsules) taken, and treatment duration. Total dosages of relugolix and E2/NETA taken during the randomized treatment period and the retreatment period will be summarized by randomized treatment group. Treatment duration during the randomized treatment period and during the retreatment period with any study drug (relugolix, E2/NETA, or placebo) will be summarized by randomized treatment group.

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

$$(\text{total tablets taken} / \text{total tablets expected to be taken}) \times 100$$

The total tablets taken will be calculated as:

$$(\text{total tablets dispensed} - \text{total tablets returned})$$

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 randomized treatment period of the study. Tablets that were dispensed and not returned will be assumed to have been taken. For patients who did not return for their last scheduled visit, tablets that were dispensed and not returned will not be included in the calculation of study drug compliance. For patients who did not return for any post-Week 52/Baseline visits and did not return dispensed study drug, 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 (eg, mean, median, etc.) will be presented, along with a categorical summary (eg, < 80%, 80% to 100%, > 100%).

The study drug compliance during the retreatment period will be calculated similarly and will be summarized for retreatment period by the randomized treatment group.

7. EFFICACY ANALYSES

7.1. General Considerations

Unless otherwise specified, efficacy analyses will be conducted using the mITT population according to the randomized treatment assignment. Stratified analyses will be stratified by the randomization stratification factors, which include pivotal study baseline menstrual blood loss (MBL) volume (< 225 vs. ≥225 mL), geographic region (North America vs. Other, where Europe, Latin America, and Rest of World will be pooled together as Other), and duration of relugolix exposure prior to randomization (28 weeks vs. 52 weeks). If the group of patients at any factor level from a randomization stratification factor (eg, patients with pivotal study Baseline MBL volume ≥ 225 mL) comprises < 10% of the entire mITT population, this stratification factor (eg, pivotal study Baseline MBL volume) will not be used for stratified analyses. In addition, based on the pooled baseline data, there are < 10 patients in 1 of the 8 strata (derived from the 3 stratification factors each with 2 levels, < 5 patients within a randomized treatment group), geographic region will not be used for stratified analysis. Stratification factors of pivotal study Baseline MBL volume (< 225 vs. ≥225 mL) and duration of relugolix exposure prior to randomization (28 weeks vs. 52 weeks) will be used in the stratified analysis for more robust strata-adjusted estimation of treatment effect. The stratification category used at the time of randomization (in the Interactive Web Recognition Service [IWRS] system) will be used for all analyses rather than data recorded on the electronic case report form (eCRF) unless otherwise specified. A sensitivity analysis of the primary endpoint will be performed if the data in the IWRS and eCRF for stratification factors differ by > 5%.

7.1.1. Analyses for 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.

For efficacy endpoints evaluating proportions descriptive statistics (point estimates and corresponding 95% CIs) will be provided by treatment group and visit as appropriate.

7.1.2. Analyses for Continuous Data

This section describes analysis for continuous endpoints in general, not including MBL volume related continuous endpoints. Analyses of change and percent change in MBL volume are described in Section 7.4.2.

Continuous variables will be summarized using descriptive statistics (eg, n, mean, median, SD, minimum, and maximum). For the analyses of change from Baseline, the mean at Baseline will be calculated for all patients with at least one post-Baseline value by treatment group. Additionally, the mean of observed results will also be calculated at each visit, including only the patients who are in the analysis population and who have data at that visit by treatment group.

For endpoints evaluating the change (absolute or percent change) from baseline, mean change as well as least squares (LS) mean change and 95% CI will be summarized. LS means and 95% CI will be derived using a mixed-effects model repeated measures approach with treatment,

randomization stratification factors, visit, and treatment by visit interaction included as fixed effects. The Baseline value will be included as a covariate, and an unstructured variance-covariance matrix will be assumed. If the mixed-effects model fails to converge, a first-order autoregressive variance-covariance matrix will be used. Calculation of the dependent variable (change from baseline) for each patient at each visit will be calculated based on the visit windows specified in Section 4.3.7. In addition, summary statistics (LS mean change or LS mean % change) will be graphically presented as appropriate. Note, the mixed-effects models will not be applied to the data analysis during the randomized treatment period due to informative censoring patient's data at the time of retreatment for patients who had MBL volume ≥ 80 mL during the randomized period. They will be applied to data analysis during the cumulative treatment period.

7.1.3. Analyses for Time to Event Data

Time-to-event endpoints will be summarized using the Kaplan-Meier method. The median, quartiles, and probabilities of an event at particular time points 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.

A stratified log-rank test will be used to compare relugolix+E2/NETA group to placebo group. Randomization stratification factors (pivotal study baseline MBL volume (<225 mL vs. 225 mL) and duration of exposure to relugolix (28 week vs. 52 weeks)) will be used for the stratified log-rank test..

7.2. Multiplicity Adjustment

The primary and the three key secondary efficacy analyses will be performed at an overall two-sided type I error of 0.05. A test will be deemed statistically significant if the two-sided p-value rounded to 4 decimal places is < 0.05. If the result of the primary endpoint analysis meets the respective evaluation criterion of the primary endpoint, the key secondary endpoints will then be tested with a fixed-sequence testing procedure to maintain the overall familywise error rate of 0.05 for the testing of primary and key secondary endpoints. The key secondary endpoints with alpha-protection for potential labelling purposes are described in Section 7.4.1. All p-values (if provided) aside from the endpoints listed in the testing order are not adjusted in multiplicity, thus are at a nominal level of 0.05.

7.3. Primary Efficacy Endpoint

The primary efficacy endpoint of the study is the proportion of patients who maintained a menstrual blood loss volume of < 80 mL at Week 76 (24 weeks/EOT of the randomized treatment period) as measured by the alkaline hematin method. The primary endpoint will be referred to as sustained responder rate at Week 76 and defined as the cumulative probability of MBL volume < 80 mL measured while on randomized treatment through Week 76. For the primary analysis, primary endpoint will incorporate the censoring rules described in Section 7.3.1. The primary endpoint of sustained responder rate will be evaluated using the mITT Population.

7.3.1. Primary Efficacy Analysis

The primary hypothesis for the primary efficacy endpoint to be tested in this study is relugolix+ E2/NETA group is superior to placebo group with respect to the sustained responder rate at Week 76 based on the MBL volume analyzed by alkaline hematin method:

$$\text{Null hypothesis } H_{01}: \ \pi^R \leq \pi^P \quad \text{vs.} \quad \text{Alternative hypothesis } H_{a1}: \ \pi^R > \pi^P$$

where π^R and π^P are the sustained responder rates for relugolix and placebo groups, respectively.

The primary efficacy analysis is the treatment comparison between the relugolix+E2/NETA group and the placebo group performed using a stratified version of the Mantel-Haenszel test statistic for sustained responder rates at Week 76 stratified by baseline mean menstrual blood loss in the pivotal study (< 225 mL vs. ≥ 225 mL), geographic region (North America vs. Other) and the duration of relugolix exposure prior to randomization (28 weeks vs. 52 weeks). In this analysis, the sustained responder rate will be calculated as the Kaplan-Meier estimate of the cumulative probability of MBL volume < 80 mL through Week 76 (Day 169) using the Kaplan-Meier method (see Section 7.3.1.1).

The difference in sustained responder rates between the relugolix+E2/NETA and placebo group and its two-sided 95% CI will be provided.

The study will be considered positive if the treatment effect for the primary endpoint is statistically significant with two-sided p-value < 0.05.)

For each patient, event status (yes/no) and time to first event (MBL volume ≥ 80 mL) or censoring (if no event) during the randomized treatment period will be derived and used in Kaplan-Meier analysis.

7.3.1.1. Kaplan-Meier Analysis

The primary endpoint of sustained responder rate at Week 76 will be estimated based on the Kaplan-Meier (KM) survival functions of time to 1st relapse (MBL volume ≥80 mL) along with 95% CI for sustained responder rate for each treatment group.

The 95% CI for the KM estimation of the sustained responder rate at Week 76 (Day 169) is calculated using the exponential Greenwood formula via log-log transformation of the survival function.

The 95% CI for treatment difference in sustained responder rates will be calculated via linear transformation of the difference in relapse free survival function with pooled variance, as follows:

$$[(\widehat{S_R}(t) - \widehat{S_P}(t)) - 1.96\sqrt{\hat{V}\left[\widehat{S_R}(t) - \widehat{S_P}(t)\right]}, (\widehat{S_R}(t) - \widehat{S_P}(t)) + 1.96\sqrt{\hat{V}\left[\widehat{S_R}(t) - \widehat{S_P}(t)\right]}\ ].$$

Treatment difference in sustained responder rates at Week 76 between relugolix and placebo will be estimated along with 95% CI using the method described below.

Variance of treatment difference will be calculated using the formula:

$$\hat{V}\left[\widehat{S_R}(t) - \widehat{S_P}(t)\right] = \hat{V}\left[\widehat{S_R}(t)\right] + \hat{V}[\widehat{S_P}(t)];$$

Where each of the variance of the KM estimate will be calculated using the Greenwood's formula

$$\hat{V}[\hat{S}(t)] = \hat{S}(t)^2 \sum_{t_i \le t} \frac{d_i}{n_i(n_i - d_i)}$$

where n_i denotes the number of patients at risk at time t_i and d_i denotes the number of events observed at time t_i.

P-value for testing treatment difference in sustained responder rates at Week 76 between relugolix+E2/NETA and placebo will be derived using an appropriate test statistic, as described in the following steps.

Step A: The stratified test statistics via log-log transformation of the difference in survival curve at a fixed time point will be used to derive the -value (Klein et al. 2007).

$$x^2 = \frac{\left\{\sum_{s=1}^{M} \left[\log\left(-\log\left(\widehat{S_{R_s}}(t)\right)\right) - \log\left(-\log\left(\widehat{S_{P_s}}(t)\right)\right)\right]\right\}^2}{\sum_{s=1}^{M} \left\{\frac{V(\widehat{S_{R_s}}(t))}{\left[\widehat{S_{R_s}}(t)\log\left(\widehat{S_{R_s}}(t)\right)\right]^2} + \frac{V(\widehat{S_{P_s}}(t))}{\left[\widehat{S_{P_s}}(t)\log\left(\widehat{S_{P_s}}(t)\right)\right]^2}\right\}}$$

Where s is the s[th] stratum in the total of M strata and t = Week 76 (Day 169) in this case. Each stratified test statistics at a fixed time under two samples follows a chi-square distribution with one-degree freedom.

Step B: If the stratified test statistics cannot be calculated due to event having not occurred in certain stratum, unstratified test will be used instead.

Step C: If unstratified test is still not estimable, the stratified version of the Mantel-Haenszel test using pooled KM estimators in each stratum will be constructed as the following. This test statistics follows a chi-square distribution with one-degree freedom.

$$\chi^2_{MH} = \frac{\left\{\sum_{s=1}^{M} \frac{n_{R_s} n_{P_s}}{n_{R_s} + n_{L_s}} \left(\widehat{S_{R_s}}(t) - \widehat{S_{P_s}}(t)\right)\right\}^2}{\sum_{s=1}^{M} n_{R_s} n_{P_s} V(\widehat{S_{pooled_s}}(t))}$$

It is noted that sustained responder rate at Week 104 (Day 365) as a key secondary endpoint will be defined as the KM estimates from the same KM curves and will be analyzed using the same method as described above.

7.3.1.2. Censoring Rules

The censoring rules in the Kaplan-Meier method are described as the following (see Table 8):

- For patients with relapse (MBL volume ≥ 80 mL) during the randomized treatment period, they will be considered as having an event at the time of first relapse. The time of first relapse is the date of the earliest menstrual bleeding days within the first FP collection visit window in which the MBL volume is ≥ 80 mL.
- For patients who did not have relapse during the randomized treatment period and had complete FP collection at each scheduled visit during the randomized treatment period,

they will be considered as not having an event and will be censored at the last available FP collection visit date during the randomized treatment period.

- For patients who did not have relapse during the randomized treatment period and had incomplete or missing FP collection during the randomized treatment period, they will be considered as censored using the following censoring rules:
  - If patients did not have FP collection after Week 64 (ie, have up to 3 FP collection visits through Week 64), they will be censored at the date of the last available FP collection visit if there are no consecutive visits with incomplete or missing FP collection; they will be censored at the date of the last available FP collection visit prior to the consecutive visits with incomplete or missing FP collections if there are two or more consecutive visits with incomplete or missing FP collections.
  - If patients had FP collection after Week 64
    - did not have incomplete or missing FP collections at consecutive visits, they will be censored at the last available FP collection date during the randomized treatment period.
    - had two or more consecutive visits with incomplete or missing FP collections, they will be censored at the last available FP collection visit date prior to the two consecutive visits during the randomized treatment period.

The time to event or censoring will be summarized by the Kaplan-Meier method. If the event time or censoring time is after the target (Day 365), Day 365 will be used as the event time or censoring time.

A sensitivity analysis will be performed considering different censoring rules (see Section 7.3.4.3).

Table 8: Censoring Rules for Kaplan-Meier Analysis During Randomized Treatment Period

Type of Patients	Event	Event Time[a]	Censored	Censoring Time[a]
Patients who experienced a relapse (MBL ≥ 80 mL)	Yes	Date of first relapse	No	N/A
Patients who did not experience a relapse and had complete FP collection	No	N/A	Yes	Date of last available FP collection visit
Patients who did not have relapse and had incomplete or missing FP collection a. Patients did not have FP collection after Week 64 b. Patients had FP collection after Week 64 i. No incomplete or missing FP collection at consecutive visits ii. Incomplete or missing FP collection at two consecutive visits iii. Incomplete or missing FP collection at three or more consecutive visits	No	N/A	Yes	a. Use the date of last available FP collection visit if no consecutive visits with incomplete or missing FP collection; Use the date of last available FP collection visit prior to the consecutive visits with incomplete or missing FP collection. b. See below i. Date of last FP collection visit ii. Date of last available FP collection visit prior to the two consecutive visits with incomplete or missing FP collection iii. Date of last available FP collection visit prior to the three or more consecutive visits with incomplete or missing FP collection

Abbreviations: N/A = not applicable.

[a] The reference point for event (or censoring) time is defined as from date of first dose of randomized study drug.

7.3.2. Definitions Related to Menstrual Blood Loss for Each Feminine Product Collection Interval

The data sources that will be used to support derivation of responder status for each FP collection interval include:

- Menstrual blood loss volume determined by the alkaline hematin method;
- Daily patient report of bleeding (yes/no) and use of FP (yes/no) captured in the paper patient diary;
- The status of FP collection return (yes/no) recorded on the eCRF page at each visit with specific reasons captured when no product collection was returned.

The total MBL volume is reported from the analysis of FP returned for each collection interval. An inventory of days (with dates) for which FP was collected and returned is also available. This inventory is aligned with patients' reports of bleeding and FP use in the paper patient diary. The status of FP collection return, and specifically the reason for non-return of FP reported on the Feminine Product Collection eCRF page is used to support derivation of responder status (see Section 7.3.3 for details).

7.3.2.1. Menstrual Blood Loss Volume

All returned FPs (validated, validated but unauthorized, or unvalidated products) collected at each visit will be analyzed by the alkaline hematin method to obtain the MBL volume. The MBL volume measured over each FP collection interval will be used to determine responder status at that visit (see details below). The vendor, KCAS, reports when unauthorized FPs (products not dispensed for use in the trial) have been returned. KCAS also reports whether the unauthorized products have previously been validated for their analysis. The report details MBL volumes for authorized, unauthorized but validated, and unauthorized and unvalidated products.

7.3.2.2. Week 52/Baseline Menstrual Blood Loss Volume

Week 52/Baseline MBL volume is defined as the last available MBL volume which was collected on or after the Week 44 in MVT-601-3003 and prior to the date of the first dose of randomized withdrawal study drug as assessed by the alkaline hematin method. If a patient did not return any FPs and reported amenorrhea, Week 52/Baseline MBL volume is set to 0 mL.

Although the responder status used to determine the patient's eligibility for this study is based on Week 48 data (or Week 44 data if Week 48 data were not available at the time when eligibility was being assessed) in MVT-601-3003, the Week 52/Baseline MBL volume defined above will be used to establish the patient's baseline for evaluation of menstrual blood loss of this study.

In addition, baseline MBL volume for the retreatment period is defined as the highest value prior to the first dose of the retreatment drug. This definition will be used for retreatment period related summaries.

7.3.2.3. Feminine Product Collection Interval

The FP collection interval at each visit is driven by types of bleeding patterns experienced by the patients, as described below:

- For patients who continue to have cyclic bleeding, the length of the interval depends on the duration of the patient's natural cycle;
- Patients who report irregular, non-cyclic bleeding are instructed to collect and return all FP used between study visits, up to 35 days, as per the schedule of events;
- For patients who report amenorrhea on the Feminine Product Collection eCRF page, an interval of last 35 days of treatment will be reviewed to ensure that reported amenorrhea is not due to incomplete collection.

For patients who are in the midst of an episode of cyclic bleeding at the time of the FP collection visit (or the last visit in the randomized treatment period which is prior to Week 104), the visit window may be extended up to 7 days after Week 104 visit (or last visit) to ensure patients return all used FPs over that bleeding episode.

Per protocol, all used FPs are to be collected at each visit and returned for analysis using the alkaline hematin method. For patients who continue to have menstrual bleeding, study visits are timed such that the FPs used in the entire menstrual bleeding cycle are collected in one container provided at each visit.

7.3.2.4. Menstrual Blood Loss Volume for Each Feminine Product Collection Interval

Menstrual blood loss volume for each FP collection interval is defined as the MBL volume obtained from the FP returned over the FP collection interval (or the FP collection interval between the two consecutive visits during the randomized treatment period which is prior to Week 104). Sum of MBLs from multiple FP collection kegs (if any) within the same FP collection interval will be used as the total MBL for that visit. The total MBL volume over each FP collection interval will be used to derive the responder status at that visit.

If a patient did not return FP over the FP collection interval prior to retreatment and reported amenorrhea on the Feminine Product Return eCRF page, she will be considered as amenorrheic and her MBL volume will be assigned as 0 mL.

If a patient did not return FP over the FP collection interval prior to retreatment and reported spotting on the Feminine Product Return eCRF page, her MBL volume will be assigned as <5 mL.

7.3.2.5. Feminine Product Return Rate by Visit

To quantify degree of compliance with FP collection during the randomized treatment period, the FPRR will be calculated based on the inventory of FP returned by day (dates) provided by the vendor, KCAS, the reason why the FP was not collected from the Feminine Product Collection eCFR page, and responses to the paper patient diary questions regarding bleeding experience and the use of FP obtained for the corresponding diary window (see Table 6). Specifically:

- For those who returned FP at a FP collection visit, the FPRR was calculated as the observed number of days with returned FPs (based on the inventory of FP received by KCAS) divided by the expected number of days with bleeding and use of product as reported on the patient's paper diary within the FP collection interval (as defined above).

$$\text{FPRR} = \frac{\text{observed (No. of days with returned FP [per KCAS])}}{\text{expected (No. of days reported bleeding and use of FP [per Diary])}} \times 100$$

- For those who did not return any FPs:
    - If the reason was "Amenorrheic" or "Patient only had Spotting that did not require the use of Sanitary Products" on the eCRF, their FPRR will be set to 100% because the lack of menstruation obviates the need for FP collection.
    - If the reason was "Patient failed to collect Used Products per Protocol" on the eCRF, their FPRR was set to 0.
    - If the reason was "Retreatment Visit; No Product Collection Required" or "Successful Retreatment; No Product Collection Required" on the eCRF, their FPRR was set to 0.
    - If the reason is "Other", their FPRR was set to 0.

The texts specified for reason "Other" will be reviewed by the medical team to identify those which belong to any pre-defined reasons listed in the first three bullets above and the reason will thus be treated accordingly.

7.3.3. Identifying Responder Status and Incomplete Feminine Product Collection at Each Visit

For each schedule visit, responder (MBL volume < 80 mL) status and incomplete FP collection will be identified for each patient by checking compliance with FP collection against the patient's paper diary, as measured by FPRR, and reasons for no FP collection (as displayed in Table 9).

Table 9: Identifying Incomplete Feminine Product Collection and Derivation of Responder Status at Each Visit During Randomized Treatment Period

FP Collection (Yes/No)	Reason for No FP Collection (CRF page)	Observed MBL Volume	FPRR	Incomplete FP collection	Responder Status
Yes	N/A	< 80 mL	>80-100%	Complete	Yes
	N/A	≥ 80 mL	N/A	N/A	No
	N/A	< 80 mL	<= 80%	Incomplete, Not acceptable	Missing
No	Amenorrheic	Set as 0 mL	100%	Complete	Yes
	Patient only had spotting that did not require the use of sanitary products	Set as 4.99 mL	100%	Complete	Yes
	Patient failed to collect used products per protocol	Missing	0%	Incomplete	Missing
	Retreatment visit; No product collection required	N/A	N/A	N/A	N/A
	Successful retreatment; No product collection required	N/A	N/A	N/A	N/A
	Other	Missing	0%	Incomplete	Missing

Abbreviations: FP = feminine product; FPRR = feminine product return rate; MBL = menstrual blood loss; N/A = not available.

7.3.4. Sensitivity Analyses

To assess the robustness of the primary analysis, the following sensitivity analyses of the primary endpoint (sustained responder rate) will be conducted at Week 76.

7.3.4.1. Sensitivity Analysis 1

To assess the potential impact of the length and full exposure of the treatment, the primary endpoint (sustained responder rate) will be analyzed for the Week 76 completers population as a sensitivity analysis as appropriate. The Week 76 completers population is defined as patients in the mITT population who completed 24 weeks of treatment.

7.3.4.2. Sensitivity Analysis 2

The primary endpoint will be analyzed on the Per-Protocol population as a sensitivity analysis, using the methods specified for the primary analysis (see definition of Per-Protocol population in Section 4.2.2).

7.3.4.3. Sensitivity Analysis 3

The primary endpoint will be analyzed on the mITT population as a sensitivity analysis, using different censoring rules for some of the patients who did not have relapse during the randomized treatment period and had incomplete or missing FP collection during the randomized treatment period, which are listed below. Except the differences in the two situations listed below, the censoring rules in Table 8 will be used. The differences are:

- If patients did not have FP collection after Week 64 (ie, have <=3 FP collection visits), they will be considered as having an event at the date of the first visit with incomplete or missing FP collection during the randomized treatment period.
- If patients had FP collection after Week 64 and had three or more consecutive visits with incomplete or missing FP collection, they will be considered as having an event at the date of the first visit of the three of more consecutive visits with incomplete or missing FP collection during the randomized treatment period.

7.3.5. Subgroup Analyses

Subgroup analyses of the primary efficacy endpoint of sustained responder rate comparing relugolix+E2/NETA and placebo will be performed to determine whether treatment effect is consistent across clinically meaningful subgroups. The differences in sustained responder rate and corresponding 95% CIs (calculated via linear transformation of the difference in relapse free survival function with pooled variance) will be displayed in a forest plot. Subgroups will include but not be limited to the following: baseline MBL volume in the pivotal study (< 225 mL vs. ≥ 225 mL), geographic region (North America vs. Other), duration of exposure to relugolix prior to randomization (28 weeks vs. 52 weeks) as well as other baseline subgroups such as age, and race. Details are provided in Table 10.

Table 10: Planned Subgroup Analyses

Subgroup Name	Subgroup Level
Geographic region	North America vs. Other
Menstrual blood loss volume at Pivotal Baseline (mL)	< 225 vs. ≥ 225 < 120, 120 to < 160, 160 to < 225, ≥ 225
Duration of exposure to relugolix prior to randomization (weeks)	28 vs. 52
Age category (years) at Pivotal Baseline	< 40 vs. ≥ 40
Race	Black or African American vs. Other ; Black or African American, White, Other
Volume of myoma at Pivotal Baseline (cm^3)	< 25 vs. ≥ 25
Volume of uterus at Pivotal Baseline (cm^3)	< 300 vs. ≥ 300
BMI (kg/m^2) at Pivotal Baseline	< 30 vs. ≥ 30 < 25, 25 to < 30, 30 to < 35, 35 to <40, ≥ 40
Maximum NRS score for uterine fibroid–associated pain at Pivotal Baseline	< 4 vs. ≥ 4

Abbreviations: BMI = body mass index; NRS = Numerical Rating Scale.

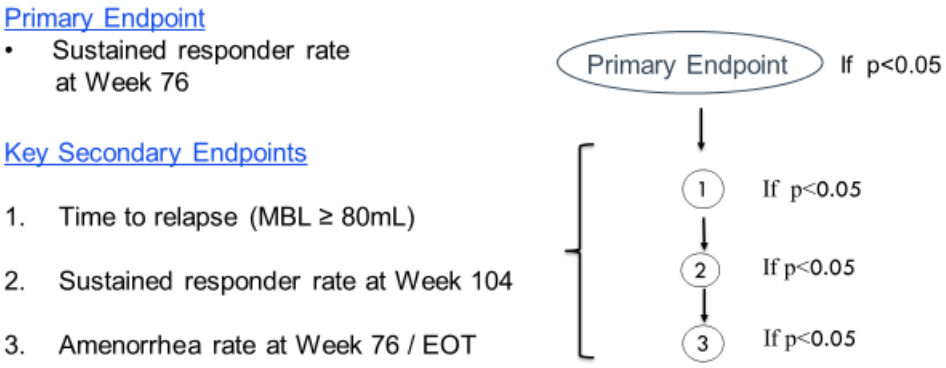
7.4. Secondary Efficacy Endpoints

Comparative statistics (p values, 95% CIs for differences) will be provided for treatment comparison of the relugolix group with the placebo group for secondary efficacy endpoints during randomized treatment period.

7.4.1. Key Secondary Endpoints

For testing whether relugolix+E2/NETA is statistically significantly superior to placebo for the primary efficacy endpoint as well as the three key secondary endpoints listed below, a fixed sequence testing procedure will be applied to maintain the family wise type I error rate. Under this testing procedure, the primary endpoint will be tested first at a 2-sided 0.05 significance level. If the p value for primary endpoint is < 0.05, the three key endpoints will be tested sequentially in the order depicted in Figure 1.

Figure 1: Fixed Sequence Testing Procedure for Primary and Key Secondary Endpoints



Abbreviations: EOT = end of treatment; MBL = menstrual blood loss.

7.4.1.1. Time to Relapse (Menstrual Blood Loss Volume ≥ 80 mL)

For the endpoint of time to MBL volume ≥ 80 mL (event) during the randomized treatment period, time-to-event will be defined as weeks from first dose of the randomized withdrawal study drug to first occurrence of MBL volume of ≥ 80 mL. The corresponding menstruation start date will be considered as the event date. Patients without an event will be censored by following the same censoring rules described in Section 7.3.1.2. Kaplan-Meier methods will be used to describe the time to event (relapse) distributions for each study treatment group from which median time to event will be derived. Hazard ratio (95%CI) of relugolix + E2/NETA to placebo for risk of relapse (MBL volume ≥ 80mL) will be provided using a proportional hazard model stratified by the stratification factors (pivotal study baseline MBL volume (< 225 ml vs. ≥ 225 mL) and duration of exposure to relugolix (28 weeks vs. 52 weeks). p-value for comparison of relugolix + E2/NETA to placebo based on the stratified log-rank test will be provided.

7.4.1.2. Proportion of Patients Who Maintained a Menstrual Blood Loss Volume of < 80 mL at Week 104

The proportion of patients who maintained a menstrual blood loss volume of < 80 mL by Week 104 (52 weeks of the randomized treatment period) as measured by the alkaline hematin method will be analyzed using the same method as that for primary efficacy endpoint described in Section 7.3. It is noted that sustained responder rate at Week 104 (Day 365) will be defined as the KM estimates from the same KM curves and will be analyzed using the same method as for the primary endpoint (see Section 7.3.1).

No subgroup analysis will be performed for this secondary efficacy endpoint.

The proportion of patients who maintained a menstrual blood loss volume of < 80 mL will also be summarized during the cumulative treatment period (including both randomized treatment period and the retreatment period).

The proportion of patients who had a menstrual blood loss volume of < 80 mL at a visit will also be summarized descriptively by visit during the cumulative treatment period (including both randomized treatment period and the retreatment period).

7.4.1.3. Proportion of Patients Achieving or Maintaining Amenorrhea

7.4.1.3.1. Determination of Amenorrhea

Rules for determining amenorrhea in the randomized treatment period is defined as those who meet one of the following requirements for two consecutive visits (approximately 56 consecutive days). Patients will be deemed to be amenorrheic during a particular visit window according to the following rules (see Table 11):

- No FP returned due to reported amenorrhea;

OR

- No FP returned due to reported spotting or FP collection with a negligible observed MBL volume coupled with other data indicating infrequent non-cyclic bleeding/spotting as described in Table 11.

If no FP collection because patient failed to collect used products per protocol or due to "Other" reason, the amenorrhea status will be missing. If no FP collection due to retreatment visit or successful retreatment, the amenorrhea status is not applicable.

Missing responses for menstrual bleeding questions in the paper diary will be treated as "No Bleeding" if paper diary compliance rate is > 70%.

Table 11: Rules for Determining Reported Amenorrhea by Visit

Feminine Product Collection [a]	Patient's Paper Diary	Amenorrhea Status
No feminine product collection due to reported amenorrhea	N/A, since paper diary entry is not required per protocol.	Yes
No feminine product collection due to reported spotting or feminine product collection with negligible observed MBL volume defined as <5 mL	• Data indicating infrequent, non-cyclic bleeding/spotting defined as bleeding/spotting with feminine product use for no more than 3 consecutive days and no more than 5 days bleeding total per visit window • Diary entry rate > 70%	Yes

Abbreviations: MBL = menstrual blood loss; N/A = not applicable.
[a] There is no requirement for feminine product return rate, as the determination of amenorrhea is based on the Feminine Product Collection eCRF page and the diary response.

7.4.1.3.2. Amenorrhea During the Randomized Treatment Period

Patients achieving or maintaining amenorrhea during the randomized treatment period at Week 76/EOT (or at Week 104/EOT) are defined as those who meet the definition of amenorrhea at last two consecutive post-baseline visits prior to date of last dose of study drug during the randomized treatment period through Week 76 (or through Week 104) visit.

A key secondary endpoint is the proportion of patients who achieved or maintained amenorrhea at Week 76/EOT during the randomized treatment period.

Patients missing amenorrhea reporting for all visits in the study will not be considered as amenorrheic by Week 76 or Week 104.

Proportion of patients achieving or maintaining amenorrhea at Week 76/EOT and at Week 104/EOT will be summarized by treatment and treatment comparisons will be performed using a stratified Cochran-Mantel-Haenszel test.

7.4.2. Other Secondary Endpoints

7.4.2.1. Proportion of Patients Who Responded to Retreatment During the Retreatment Period

Proportion of patients who responded (MBL volume of < 80 mL) to retreatment with open-label relugolix 40 mg +E2/NETA during the retreatment period among placebo patients whose MBL volume had returned to ≥ 80 mL during the randomized treatment period will be summarized descriptively. The point estimate and its corresponding 95% CIs will be provided.

The change and percent change in MBL volume from retreatment initiation will also be summarized descriptively as a continuous variable every 4 weeks after retreatment initiation during the retreatment period.

7.4.2.2. Change and Percent Change from Week 52/Baseline in Menstrual Blood Loss Volume During the Randomized Treatment Period

The change (for all patients) and percent change (for patients who did not have amenorrhea at baseline) in MBL volume from Week 52/baseline will be summarized descriptively as a continuous variable at each visit during the randomized treatment period including Week 76 and Week 104 visits. Observed MBL volume for each timepoint will be used to calculate change and percent change from Week 52/Baseline. In addition, the change and percent change in MBL volume from Week 52/baseline will also be summarized descriptively as a continuous variable at each visit during the cumulative treatment period.

For change and percent change in MBL volume, mixed-effects model will not be used to derive LS means (and 95% CIs) due to skewed distribution of MBL volumes (majority of patients experiencing amenorrhea with MBL volume set as 0 mL at baseline). Descriptive statistics (medians and simple means) will be provided for each treatment group.

Categorical summary of change in MBL volume will be provided by grouping patients into three categories: who had increase, no change, and decrease in MBL volume from Week52/Baseline over time by visit for cumulative treatment period. A figure will also be provided to display the percentages for each category by treatment group every 12 weeks from Week 52/Baseline up to Week 104.

Besides, the observed MBL values will be categorized into three categories: 0 mL, > 0 to < 80 mL, and ≥ 80 mL. Shift tables will be provided to summarize Week 52/Baseline versus each post-baseline visit for the randomized treatment period and cumulative treatment period. Figures will also be provided to display the percentages for each category by treatment group at Week 52/Baseline, Week 76, and Week 104.

7.4.2.3. Proportion of Patients Whose Menses Has Resumed

This endpoint is defined only for patients who were amenorrheic (ie, had MBL volume of 0 mL at baseline).

7.4.2.3.1. Determination of Resumption of Menses

Patients who were previously amenorrheic will be deemed to resume the menses during a visit window according to the following rules:

- FPs were collected and MBL volume was ≥ 5 mL;

OR

- FPs were collected and MBL volume was < 5 mL; however, there were more than 5 days and more than 3 consecutive days of bleeding with FP use during the visit window per the diary;

OR

- No FPs were returned because patient failed to collect used products per protocol; however, there were more than 5 days and more than 3 consecutive days of bleeding with FP use during the visit window per the diary;

OR

- No FPs were returned due to other reasons that indicate menstruation occurred; also, there were more than 5 days and more than 3 consecutive days of bleeding with FP use during the visit window per the diary.

These were also described in Table 12.

Table 12 Rules for Determining Resumption of Menses by Visit

Feminine Product Collection	Patient's Paper Diary	Resumption of Menses?	Date of Resumption of Menses
Feminine products were collected with observed MBL volume ≥ 5 mL	N/A	Yes	Earliest menstruation date recorded in KCAS[a] data for the corresponding feminine product collection
Feminine products were collected with observed MBL volume < 5 mL	> 5 days and > 3 consecutive days of bleeding with feminine product use reported in paper diary	Yes	Earliest menstruation date recorded in KCAS[a] data for the corresponding feminine product collection
Feminine product collection not performed per protocol or no feminine product collection performed due to other reasons which indicate menstruation occurred.	> 5 days and > 3 consecutive days of bleeding with feminine product use reported in paper diary	Yes	Paper diary date of first episode of > 3 consecutive days of bleeding with feminine product use in timeframe corresponding to missed collection.

[a] This date has been reconciled against the Menstruation Status data in EDC and is more reliable if irregular bleeding has been reported.
Abbreviations: MBL = menstrual blood loss; N/A = not applicable.

7.4.2.3.2. Resumption of Menses During the Randomized Treatment Period

Proportion of patients whose menses has resumed during the randomized treatment period among those who were amenorrheic at Week 52/Baseline will be summarized by treatment at Week 76 and at Week 104. The point estimate and its corresponding 95% CIs will be provided based on the Kaplan-Meier method as described below in Section 7.4.2.4.

7.4.2.4. Time to Resumption of Menses

For time to resumption of menses among those who were amenorrheic at Week 52/Baseline, time-to-event will be defined as weeks from first dose of the Randomized Withdrawal study drug to first occurrence of the menses resumption during the randomized treatment period. Patients without an event will be censored at last assessment date prior to the last dose of randomized withdrawal study drug before Week 104.

Kaplan-Meier methods will be used to describe the time to event distributions by Week 76 and by Week 104 for each study treatment arm. Difference (95% CI) in the rate of menses resumption between the two treatment groups will be calculated at Week 76 and Week 104 based on the KM estimates. P-value from the unstratified log-rank test for treatment comparison in the KM rates at Week 76 and Week 104 will be provided. See Section 7.1.3 for more details.

7.4.2.5. Proportion of Patients with Menstrual Blood Volume ≥ 80 mL at Any Timepoint During the 52-week Randomized Treatment Period

Proportion of patients with menstrual blood volume ≥ 80 mL at any timepoint during the randomized treatment period will be summarized by visit based on the Kaplan-Meier estimates. The treatment differences (95% CI) in the proportions at Week 76 and at Week 104 will be provided using the analysis method described above in Section 7.4.1.1.

7.4.2.6. Change and Percent Change from Week 52/Baseline in Hemoglobin

The change and percent change from Week 52/Baseline in hemoglobin during the randomized treatment period will be summarized at each scheduled visit as described in Section 7.1.2.

The change and percent change from Week 52/Baseline in hemoglobin during the cumulative treatment period will be summarized descriptively by visit.

7.4.2.7. Proportion of Patients with Anemia or Below Lower Limit of Normal

The proportion of patients with anemia (hemoglobin values ≤ 10.5 g/dL) during the randomized treatment period will be summarized by visit. Similarly, the proportion of patients with hemoglobin values below the lower limit of normal (ie, < 11.6 g/dL) during the randomized treatment period will also be summarized by visit.

The proportion of patients with hemoglobin ≤ 10.5 g/dL and with hemoglobin < 11.6 g/dL will also be summarized by visit for the cumulative treatment period.

7.4.2.8. Change from Week 52/Baseline in SF-36v2 Scores

Change from Week 52/Baseline in SF-36v2 scores of eight scales and two summary components during the randomized treatment period will be summarized using descriptive statistics. The scores after the time of retreatment initiation with open-label relugolix + E2/NETA will not be included in the analysis for randomized treatment period.

Change from Week 52/Baseline in SF-36v2 scores of eight scales and two summary components during the cumulative treatment period will be summarized by visit.

Calculation of SF-36v2 Scale Scores

The eight SF-36v2 scales, including Physical Functioning, Role-Physical, Bodily Pain, General Health, Vitality, Social Functioning, Role-Emotional, Mental Health, and Reported Health Transition will be calculated using the PRO CoRE™ Scoring Software provided by QualityMetric Incorporated, LLC. The physical component summary (PCS) and mental component summary (MCS) scores will also be calculated using this software.

PRO CoRE™ is a validated tool that an organization can use to score health-related quality of life data and feel confident that the data was scored correctly. The eight SF-36v2 scales are normalized using mean scores and their SDs from the 2009 general US population, and then they are aggregated using weights from the 2009 general US population to obtain PCS and MCS. Finally, the normalized scores of the eight scales and two component summaries provided by the software will be used for analysis. Higher scores indicate higher health related quality of life.

7.4.2.9. Change from Week 52/Baseline in PGA Score

Change from Week52/Baseline in PGA for functions and symptoms during the randomized treatment period will be summarized using descriptive statistics. The scores after the time of retreatment initiation with open-label relugolix + E2/NETA will not be included in the analysis for randomized treatment period.

Categorical change from Week 52/Baseline at each visit will also be summarized for randomized treatment period and cumulative treatment period separately. The PGA for function and symptoms will be evaluated using a 5-point response scale (eg, absent, mild, moderate, severe, and very severe). To calculate change from Week 52/Baseline to Week 76 or Week 104, the numerical scores shown in Table 13 will be assigned to each response level.

Table 13: Patient Global Assessment Numerical Scores

Response Scale (Function)	Response Scale (Symptoms)	Numerical Score
No limitation at all	Not severe	1
Mild limitation	Mildly severe	2
Moderate limitation	Moderately severe	3
Quite a bit of limitation	Very severe	4
Extreme limitation	Extremely severe	5

For each item, the count and proportion of improvement by level or at least one level will be tabulated by treatment group and by visit. The denominator for the proportion will be based on the number of patients who provided non-missing responses to the items.

The change from Week 52/Baseline in PGA scores during the cumulative treatment period will be summarized descriptively by visit.

7.4.2.10. Change from Week 52/Baseline in the WPAI-UF Scores

Change from Week 52/Baseline in WPAI-UF scores during the randomized treatment period will be summarized using descriptive statistics. The scores after the time of retreatment initiation with open-label relugolix + E2/NETA will not be included in the analysis for randomized treatment period.

Change from Week 52/Baseline in WPAI-UF scores during the cumulative treatment period will be summarized by visit.

Calculation of WPAI-UF Scores

WPAI-UF outcomes are expressed as impairment percentages, with higher numbers indicating greater impairment and less productivity, ie, worse outcomes. There are six questions in this questionnaire:

1 = currently employed

2 = hours missed due to uterine fibroid symptoms

3 = hours missed other reasons

4 = hours actually worked

5 = degree uterine fibroid symptoms affected productivity while working in past 7 days

6 = degree uterine fibroid symptoms affected regular activities in past 7 days

If a patient is not employed (ie, answered no to the first question), only question 6 needs to be answered.

Four WPAI-UF scores will be calculated using the following algorithms (Reilly et al. 1993; Reilly Associates Health Outcomes Research website [http://www.reillyassociates.net]) and analyzed. Multiply the scores by 100% to express them in percentages. Higher percentages indicate that impairment due to Uterine Fibroids is worse.

- Percent work time missed due to uterine fibroid symptoms: $Q2/(Q2 + Q4) \times 100\%$
- Percent impairment while working due to uterine fibroid symptoms: $Q5/10 \times 100\%$
- Percent overall work impairment due to uterine fibroid symptoms:
$$\{Q2/(Q2 + Q4) + [1 - Q2/(Q2 + Q4)] \times (Q5/10)\} \times 100\%$$
- Percent activity impairment due to uterine fibroid symptoms: $Q6/10 \times 100\%$

7.4.2.11. Change from Week 52/Baseline in the UFS-QoL Score

Change from Week 52/Baseline in the UFS-QoL Bleeding and Pelvic Discomfort (BPD) Scale score, UFS-QoL Symptom Severity Scale score, UFS-QoL Revised Activities Scale, UFS-QoL subscale scores, and total scores during the randomized treatment period will be summarized using descriptive statistics by treatment group and by visit. The scores after the time of retreatment initiation with open-label relugolix + E2/NETA will not be included in the analysis for randomized treatment period.

The change from Week 52/Baseline in UFS-QoL scores during the cumulative treatment period will be summarized by visit.

In addition, the proportion of responders as measured by the BPD Scale score and Revised Activities Scale will be summarized, as described below. This analysis will be performed for randomized treatment period and cumulative treatment period separately.

Calculation of UFS-QoL Symptom Severity Scale Score

To calculate the Symptom Severity Scale score, a summed score is created for the items listed below and then the formula below the table is used to transform raw scores to a normalized score with a range of possible values from 0 to 100. This provides Symptom Severity Scale scores, where higher scores are indicative of greater symptom severity and lower scores indicate lower symptom severity.

Scale	Sum Item Values	Lowest and Highest Possible Raw Scores	Possible Raw Score Range
Symptom Severity	Sum 1 – 8	8, 40	32

Formula for Transformation of Symptom Severity Raw Scores:

$$\text{Transformed Scale} = \left[\frac{(\text{Actual raw score} - \text{lowest possible raw score})}{\text{Possible raw score range}}\right] \times 100$$

Calculation of UFS-QoL Bleeding and Pelvic Discomfort Scale Score

The UFS-QoL BPD Scale has been derived from the UFS-QoL Symptoms Scale; the derivation and validation of this new scale can be found in the MVT-601-3001/3002 SAP. The new scale consists of the following three symptoms proximal to uterine fibroids:

- Heavy bleeding during your menstrual period (Q1)
- Passing blood clots during your menstrual period (Q2)
- Feeling tightness or pressure in your pelvic area (Q5)

To calculate the score for the BPD Scale, a summed score of the items listed below is created and then the formula below the table is used to transform the raw score to a normalized score. This provides BPD Scale scores, where higher score values are indicative of greater symptom severity and lower scores will indicate minimal symptom severity (high scores = bad).

Sub-Scale	Sum Item Values	Lowest and Highest Possible Raw Scores	Possible Raw Score Range
Q1, Q2 and Q5	Sum 1,2,5	3, 15	12

Formula for Transformation of BPD Raw Scores:

$$\text{Transformed Scale} = \left[\frac{(\text{Actual raw score} - \text{lowest possible raw score})}{\text{Possible raw score range}}\right] \times 100$$

On the basis of transformed score for BPD Scale, change from Week 52/Baseline in the transformed score for BPD Scale at Week 76 and Week 104 will be defined as a secondary endpoint. The proportion of patients who are responders (defined as meeting a meaningful change threshold from Week 52/Baseline in the BPD Scale) at Week 76 and Week 104 on the transformed score for the BPD Scale will be summarized for each treatment group. The proposed responder threshold is a 20-point change. Details in the determination of the meaningful change in the BPD Scale can be found in the MVT-601-3001/3002 SAP Appendix 4.

As a descriptive assessment on robustness of the responder analysis, a plot of the cumulative distribution function (CDF) will be provided for each treatment group to display the change from

Week 52/Baseline to Week 76 and Week 104 in the transformed score for BPD Scale on the x-axis and cumulative percentage of patients experiencing up to that change on the y-axis.

Calculation of Other UFS-QoL Scale Scores and UFS-QoL Total Score

For the other UFS-QoL scales (concern, activities, revised activities, energy/mood, control, self-conscious, and sexual function), a summed score of the items listed below is created for each individual scale. To calculate the UFS-QoL total score, the values for each individual scale are summed. See Table 14 for details. Using the formula below the table, all raw scores are transformed to normalized scores. Higher scores are indicative of better health-related quality of life (high = good).

For endpoints evaluating a single question, the raw score is used in the analysis. The activity and revised activity domain scores will be summarized by treatment group.

Table 14: Formulas for Scoring and Transforming UFS-QoL Scale Scores and UFS-QoL Total Score

Scale	Sum Item Values	Lowest and Highest Possible Raw Scores	Possible Raw Score Range
Concern	9+15+22+28+32	5, 25	20
Activities	10+11+13+19+20+27+29	7, 35	28
Revised activities	11+13+19+20+27	5,25	20
Energy/mood	12+17+23+24+25+31+35	7, 35	28
Control	14+16+26+30+34	5, 25	20
Self-conscious	18+21+33	3, 15	12
Sexual function	36+37	2, 10	8
HRQL TOTAL	Sum of 6 Subscale Scores[a]	29, 145	116

Abbreviations: HRQL = health-related quality of life; UFS-QoL = Uterine Fibroid Symptom and Health-Related Quality of Life (Questionnaire).

[a] HRQL Total includes following scales: concern, activities, energy/mood, control, self-conscious, and sexual function.

Formula for Transformation of Raw Scores of Other Scale Scores:

$$\text{Transformed Scale} = \left[\frac{(\text{Highest possible raw score} - \text{Actual raw score})}{\text{Possible raw score range}}\right] \times 100$$

For revised activities, the proportion of patients who are responders (defined as meeting a meaningful change from Week 52/Baseline in the revised activity score) at Week 76 and Week 104 will be analyzed similarly to that for the change in BPD Scale score for each treatment group. The proposed responder threshold is a 20-point increase. Details of the determination of the meaningful change in the Revised Activities Scale score can be found in the MVT-601-3001/3002 SAP Appendix 5.

Missing Item Handling in UFS-QoL Score Calculation

For any scale analyses, if < 50% of the scale items are missing, the scale should be retained using the mean scale score of the items present. If ≥ 50% of the items are missing, no scale score should be calculated; the subscale score will be considered missing.

8. PHARMACODYNAMIC ANALYSES

Pharmacodynamic data for the randomized withdrawal study consists of pre-dose estradiol serum concentration at Week 52/Baseline and Week 56 only. Estradiol concentration data will be listed and summarized using descriptive statistics (including absolute and change from Week 52/Baseline) by randomized withdrawal study treatment group at baseline and Week 56. The number and percentage of patients with individual estradiol concentration values will be summarized by treatment group at baseline and Week 56 according to the following categories.

- < 10 pg/mL, 10 to < 20 pg/mL, 20 to < 30 pg/mL, 30 to < 40 pg/mL, 40 to < 50 pg/mL, 50 to < 60 pg/mL, 60 to < 70 pg/mL, and ≥ 70 pg/mL
- < 20 pg/mL, 20 to < 50 pg/mL, and ≥ 50 pg/mL
- < 20 pg/mL, 20 to < 60 pg/mL, and ≥ 60 pg/mL

Estradiol concentration values below the limit of quantification will be set to half of the lower limit of quantification.

Follicle-stimulating hormone concentrations were collected at Week 76 in patients who have been amenorrheic since randomization and the concentration data will be listed and summarized using descriptive statistics.

9. SAFETY ANALYSES

Unless otherwise specified, safety analyses will be conducted using the safety population according to the actual treatment received by the patients in this randomized withdrawal study.

9.1. Adverse Events

Adverse events will be collected from the time of the first dose of randomized withdrawal study drug through the safety follow up visit approximately 30 days after the last dose of randomized withdrawal study drug (if no retreatment was initiated) or approximately 30 days after the last dose of retreatment drug (if the retreatment was initiated), or the date of initiation of another investigational agent or hormonal therapy (eg, elagolix, acotol) or surgical intervention (eg, uterine artery embolization), 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.

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), v.5.0 dated 27 Nov 2017, and will be coded to preferred term and system organ class using MedDRA 24.0 or higher.

A treatment-emergent adverse event for cumulative treatment period (including the randomized treatment and the retreatment periods) is defined as any adverse event that onsets or worsens after administration of the first dose of randomized withdrawal study drug.

For patients who stayed on the randomized treatment till end of the study, a treatment-emergent adverse event for the randomized treatment period is defined as any adverse event that onsets or worsens after administration of the first dose of randomized withdrawal study drug. For patients who had the retreatment drug initiated, any adverse event that onsets or worsens after administration of the first dose of randomized withdrawal study drug and prior to the initiation of the retreatment drug is defined as treatment-emergent adverse event for the randomized treatment period.

For patients who started the retreatment drug, a treatment-emergent adverse event for the retreatment period is defined as any adverse event that onsets after the initiation of retreatment drug.

Adverse event summaries will be based on treatment-emergent adverse events in the randomized withdrawal study, unless otherwise specified. The summaries will mainly be done for the cumulative treatment period. Some selected summaries may be provided for the randomized withdraw period and the retreatment period separately. All adverse events will be listed in by-patient listings.

The following tabular summaries that include the number and percentage of patients will be provided for the cumulative treatment period:

- Overview of adverse events;
- All adverse events;
  - By SOC and PT;

  - By decreasing frequency of PT;
  - By SOC, PT, and maximum severity;
  - Drug-related per investigator by SOC and PT;
  - By time to onset, SOC and PT;
- Grade 3 or higher adverse events;
  - By SOC and PT;
  - By decreasing frequency of PT;
  - Drug-related per investigator by SOC and PT;
- Adverse events leading to study drug withdrawal;
  - By SOC and PT;
  - By decreasing frequency of 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 special interest or clinical interest (ALT or AST $\geq 3 \times$ ULN);
  - By SOC, PT, and maximum severity;
  - By decreasing frequency of PT.
- Adverse events with first onset during the randomized withdrawal study;
  - By SOC and PT;
  - By decreasing frequency of PT.

Additionally, adverse event categories defined in Table 15 will be summarized by decreasing frequency of PT for the cumulative treatment period.

The following tabular summaries that include the number and percentage of patients will be provided for the retreatment period:

- Overview of adverse events;

- All adverse events;
  - By SOC and PT;
  - By decreasing frequency of PT;
  - By time to onset, SOC and PT;
- Grade 3 or higher adverse events;
  - By SOC and PT;
  - By decreasing frequency of PT;
- Adverse events leading to study drug withdrawal;
  - By SOC and PT;
  - By decreasing frequency of PT;
- Serious adverse events;
  - By SOC and PT;
  - By decreasing frequency of PT;
- Adverse events of special clinical interest (ALT or AST $\geq 3 \times$ ULN);
  - By decreasing frequency of PT.

The adverse events which were ongoing at the time of randomization of the randomized withdrawal study will be summarized as specified below. Such adverse events were collected during the pivotal study and LTE study and they are considered pre-existing conditions for the randomized withdrawal study.

- All adverse events;
  - By SOC and PT;
  - By decreasing frequency of PT;
- Grade 3 or higher adverse events;
  - By SOC and PT;
  - By decreasing frequency of PT;

9.1.1. Relationship to Study Drug

Adverse events will be classified as "related" to study treatment if the relationship was rated by the investigator as possibly related or probably related. Adverse events related to any study drug component (relugolix or placebo and E2/NETA or placebo) will be considered as related to study drug.

9.1.2. Severity of Adverse Event

Grade 3 or higher adverse events will be summarized by SOC, PT, and/or maximum severity, relationship to study treatment.

9.1.3. Serious Adverse Event

Serious adverse events will be summarized by SOC, PT, and/or maximum severity, relationship to study treatment.

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 randomized withdrawal study;
- Deaths occurring after a patient leaves the 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.

9.1.4. Adverse Event Leading to Withdrawal of Study Drug

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

Adverse events with “drug withdrawn” as action taken due to any one of the components of study drug will be considered as leading to withdrawal of study drug.

9.1.5. Adverse Events Leading to Dose Interruption

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

Adverse events with “drug interrupted” as action taken due to any one of the components of study drug will be considered as leading to dose interruption.

9.1.6. Adverse Events Resulting to Fatal Outcome

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

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

9.1.7. Other Safety Parameters of Interest

In addition, other safety parameters of interest defined in Table 15 summarized by decreasing frequency of PT under each safety population.

Table 15: Other Safety Parameters of Interest

Category	Search Criteria
Bone health related events	Osteoporosis/Osteopenia SMQ (broad) Fracture (custom SMQ): All preferred terms including the term "fracture," excluding "Tooth fracture" and "Fracture of penis"
Hepatic transaminase elevations	Drug-related hepatic disorders – comprehensive SMQ (narrow)
Carbohydrate and lipid metabolic effects	Dyslipidemia SMQ Hyperglycemia/new onset diabetes mellitus SMQ (narrow)
Vasomotor symptoms	The following five Preferred Terms will be included: • Hyperhidrosis; • Feeling hot; • Hot flush; • Night sweats; • Flushing.
Mood disorders	MedDRA Depression and Suicide/Self-Injury SMQ (broad)

Abbreviations: MedDRA = Medical Dictionary for Regulatory Activities; SMQ = Standardized MedDRA Query.

9.1.8. Adverse Events of Clinical Interest

Adverse Events of clinical interest in the randomized withdraw study are defined as any increase in ALT or AST $\geq 3 \times$ ULN, which are reported as "Yes" for adverse events of clinical interest in the adverse event CRF. Adverse events of clinical interest will be summarized.

9.2. Laboratory Data

Laboratory parameters, including chemistry and hematology panels, specified as per protocol for the randomized withdrawal study, and collected from the central laboratory will be tabulated and presented in by-patient listings. Local laboratory parameters will not be used for summary. Urinalysis and hepatitis virus serological test results will be provided in by-patient listing only.

The National Cancer Institute CTCAE Grading Scale with numeric component will be used to categorize toxicity grade for laboratory parameters (CTCAE v5.0, dated 27 Nov 2017). Parameters that have criteria available for both low and high values (eg, hypercalcemia 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 she has laboratory values meeting each criterion. Shift tables will be provided for each parameter, which is gradable by the CTCAE, to summarize Week 52/Baseline toxicity grade versus worst post-Baseline toxicity grade during this randomized withdrawal study. For laboratory parameters that are not gradable by the CTCAE, a shift table based upon the normal range (low, normal, and high) will be provided for each parameter to summarize the Week 52/Baseline versus worst post-Baseline results during this randomized withdrawal study. A shift table for selected parameters based on predefined

categories (refer to Appendix 1) will be provided to summarize Week 52/Baseline category versus post-baseline category in the randomized withdrawal study. The parameters include glucose, fasting glucose, HbA1c, low density cholesterol, and high-density cholesterol.

Boxplots of laboratory values over time in the randomized withdrawal study, will be plotted for key laboratory parameters. These laboratory parameters include, but are not limited to, hematology (hemoglobin, platelets, leukocytes, neutrophils), creatinine, glomerular filtration rate, and hepatic function panel (alanine aminotransferase [ALT], aspartate aminotransferase [AST], alkaline phosphatase [ALP], and total bilirubin).

The change from Week 52/Baseline to each post-Baseline study visit will be presented by treatment group for each laboratory test in both tables and figures.

The number and proportion of patients with liver test elevations will be presented by treatment group. Liver test elevations are assessed by using post-Baseline results for ALT, AST, ALP, and total bilirubin based on the definitions presented in Table 16.

Table 16: Categories of Liver Test Elevations

Laboratory Test	Category
ALT or AST	ALT or AST > ULN < 3 × ULN ALT or AST ≥ 3 × to < 5 × ULN ALT or AST ≥ 5 × to < 10 × ULN ALT or AST ≥ 10 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 + total bilirubin > 2 × ULN
ALT or AST, total bilirubin, and ALP	ALT or AST ≥ 3 x ULN + total bilirubin > 2 × ULN + ALP < 2 × ULN

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

The number and percentage of patients with concurrent (defined as measurements on the same day) ALT or AST ≥ 3 × ULN and total bilirubin > 2 × ULN will also be presented.

Selected chemistry and hematology test results meeting pre-defined limits of change at any time and based on last observation on treatment will be also summarized per pre-defined threshold (refer to Appendix 2) by treatment group.

9.3. Other Safety Analyses

9.3.1. Electrocardiograms

ECG was collected at unscheduled visit only besides at Week 52 (which was collected as part of the LTE study). Overall ECG assessments will be listed.

9.3.2. Vital Signs and Body Weight

Blood pressure (systolic and diastolic), heart rate, weight, and BMI will be summarized at Week 52/Baseline and each subsequent scheduled assessment by treatment group. Change from Week 52/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 data will also be provided in by-patient listings.

Potentially clinically significant abnormalities in vital signs are defined in Table 17, 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 17: Categories of Potentially Clinically Significant Abnormalities in Vital Signs

Parameter	Category
Systolic blood pressure	≥ 140 mmHg and greater than Week 52/Baseline ≥ 180 mmHg and greater than Week 52/Baseline
	≤ 90 mmHg and less than Week 52/Baseline
	Increase of ≥ 20 mmHg from Week 52/Baseline Decrease of ≥ 20 mmHg from Week 52/Baseline
Diastolic blood pressure	≥ 90 mmHg and greater than Week 52/Baseline ≥ 105 mmHg and greater than Week 52/Baseline
	≤ 50 mmHg and less than Week 52/Baseline
	Increase of ≥ 15 mmHg from Week 52/Baseline Decrease of ≥ 15 mmHg from Week 52/Baseline
Heart rate	≥ 120 bpm and greater than Week 52/Baseline
	< 45 bpm and less than Week 52/Baseline
	Increase of ≥ 15 bpm from Week 52/Baseline Decrease of ≥ 15 bpm from Week 52/Baseline

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

Potentially clinically significant abnormalities in post-Baseline weight are to be summarized using the categories specified in Table 18. Potentially clinically significant abnormalities in weight will also be flagged in by-patient listings.

Table 18: Categories of Potentially Clinically Significant Abnormalities in Weight

Parameter	Category
Body weight	Increase of ≥ 5% from Week 52/Baseline
	Increase of ≥ 10% from Week 52/Baseline
	Decrease of ≥ 5% from Week 52/Baseline
	Decrease of ≥ 10% from Week 52/Baseline

9.3.3. Bone Mineral Density

Corrected BMD data will be used for analysis as determined by the central radiology laboratory in the three prespecified anatomical locations: lumbar spine (L1–L4), total hip, and femoral neck.

9.3.3.1. Analysis Using Week 52/Baseline Bone Mineral Density as Baseline

As defined in Section 4.3.6, the BMD at Week 52 of the extension study will be used as the Week 52/Baseline in the randomized withdrawal study as long as it was within 30 days of the first dose date of the randomized study drug. If there are multiple BMD measurements at Week 52 in the extension study, the one closest to the first dose date of the study drug will be used as Week 52/Baseline in the randomized withdrawal study.

BMD at Week 52/Baseline and Week 104 visit will be summarized descriptively by treatment group and each measured anatomical location for all patients in the safety population. Percentage changes from Week 52/Baseline along with 95% CIs of mean percentage changes also will be summarized by treatment group and anatomical location. Mean percentage change from Week 52/Baseline with its corresponding 95% CI will be plotted by visit, treatment group, and anatomical location.

An analysis of covariance (ANCOVA) model will be used to assess BMD at Week 104. The model will include treatment, stratification factors (menstrual blood loss volume at pivotal study baseline, and duration of relugolix exposure prior to randomization), race (African American versus Other) as fixed factors, and age at Week 52/Baseline, Week 52/Baseline BMD value, and BMI at Week 52/Baseline as covariates. Least square means on each anatomical location will be presented and plotted with associated 95% CIs. Taking into consideration of the BMD data collected for patients who prematurely discontinued prior to Week 104, the ANCOVA model analysis will also be performed similarly to assess BMD at Week 104/EOT.

Categorical representation of percentage change from Week 52/Baseline to Week 104 (also Week 104/EOT) will be presented by the number and proportion of patients who had BMD increase > 0%, no change (0%), declines of ≤ 2%, >2% to 3%, > 3% to 5%, > 5% to 8%, and > 8% by pivotal study treatment group and anatomical location. The 95% CIs will be provided for the respective proportions.

Categorical changes from Week 52/Baseline in overall BMD (defined as lumbar spine and total hip) also will be assessed.. Femoral neck evaluates a smaller area of bone mass than the total hip and is prone to lower precision in the measurement (Leslie et al. 2007; ISCD 2015). Since

femoral neck BMD may be associated with discordant readings compared with the total hip or lumbar spine due to technical considerations, it is not expected to add meaningful interpretation of overall BMD changes in response to treatment, though will be included for completeness.

Corrected BMD Z-scores and its change from Week 52/Baseline will be summarized by treatment group, visit, and anatomical location with descriptive statistics including 95% CIs, and the number and percentage of patients with a corrected BMD Z-score < -2.0 will be presented by treatment group, visit, and anatomical location. Additionally, categorical change from Week 52/Baseline in Z-score will be summarized as ≥ 0, < 0 to -0.25, < -0.25 to -0.5, < -0.5 to -0.75, < -0.75 to -1, < -1 to -2 and < -2, by treatment group, visit and anatomical location.

BMD percentage changes from Week 52/Baseline will also be summarized by intrinsic factors (eg, age, race, body mass index) and extrinsic factors (eg, geographic region). Subgroups will include, but will not be limited to, the subgroups outlined in Table 19.

Table 19: Planned Subgroup Analyses for Bone Mineral Density

Subgroup Name	Subgroup Level
Geographic region	North America, Other
Age category (years) at pivotal study baseline	< 40, ≥ 40
Race	Black or African American, Other
Ethnicity	Hispanic or Latino, Not Hispanic or Latino
BMI (kg/m^2) at pivotal study baseline	<30, ≥ 30
Duration of exposure to relugolix prior to randomization (weeks)	28, 52

Abbreviations: BMI = body mass index.

9.3.3.2. Analysis Using the Pivotal Study Baseline Bone Mineral Density as Baseline

For the visits in pivotal and extension studies, the visit window specified in the extension study SAP Table 3 will be used where the first dose of pivotal study drug was used as the reference.

Percentage changes from pivotal study baseline along with 95% CIs of mean percentage changes since first dose of the pivotal study drug at all the timepoints when BMD was scheduled to be collected in a pivotal study, extension study and the randomized withdrawal study also will be summarized by treatment group and anatomical location. Mean percentage change from pivotal study baseline with its corresponding 95% CI will be plotted by visit, treatment group, and anatomical location.

Categorical representation of percentage change from pivotal study baseline to all the timepoints where BMD was scheduled to be collected since pivotal study baseline will be presented by the number and proportion of patients who had BMD increase > 0%, no change (0%), declines of ≤ 2%, >2% to 3%, > 3% to 5%, > 5% to 8%, and > 8% by treatment group and anatomical location. The 95% CIs will be provided for the respective proportions.

Categorical changes from pivotal baseline in overall BMD (defined as lumbar spine and total hip) also will be assessed at all the timepoints where BMD was scheduled to be collected since pivotal study baseline.

Corrected BMD Z-scores and its change from pivotal study baseline will be summarized by treatment group, visit, and anatomical location with descriptive statistics including 95% CIs, and the number and percentage of patients with a corrected BMD Z-score < -2.0 will be presented by treatment group, visit, and anatomical location. Additionally, categorical change from pivotal study baseline in Z-score will be summarized as ≥ 0, < 0 to -0.25, < -0.25 to -0.5, < -0.5 to -0.75, < -0.75 to -1, < -1 to -2 and < -2, by treatment group, visit and anatomical location.

These analyses will be carried out by six combined analysis groups where the pivotal study treatment group, LTE study treatment and the randomized withdrawal study treatment group are considered (see the Table 20 below).

In addition, to observe the long-term impact of relugolix + delayed E2/NETA on BMD loss, these analyses will be performed by the relugolix monotherapy status groups: patients who were previously treated with any relugolix + delayed E2/NETA (from Groups 3 and 4) and the rest of the patients (from Groups 1, 2, 5 and 6).

9.3.3.3. Analysis Using Pivotal Week 12 BMD as Baseline

To observe the long-term impact of relugolix+E2/NETA on BMD loss in dependent of hypoestrogenic effect caused by relugolix monotherapy, BMD at Week 12 in the pivotal studies will be used as baseline to calculate percent change in BMD for each patient. Descriptive statistics will be provided for percent change from pivotal Week 12 in BMD by the six combination groups and as well as by the monotherapy status groups.

Table 20: Combined Analysis Groups for Bone Mineral Density

Combined Analysis Group	Pivotal Study Treatment Group	Extension Study Treatment	Randomized Withdrawal Study Treatment Group
Group 1 Relugolix + E2/NETA Relugolix + E2/NETA Relugolix + E2/NETA Relugolix + E2/NETA	Relugolix + E2/NETA	Relugolix + E2/NETA	Relugolix + E2/NETA
Group 2 Relugolix + E2/NETA Relugolix + E2/NETA Relugolix + E2/NETA Placebo	Relugolix + E2/NETA	Relugolix + E2/NETA	Placebo
Group 3 Relugolix + Delayed E2/NETA Relugolix + E2/NETA Relugolix + E2/NETA Relugolix + E2/NETA	Relugolix + Delayed E2/NETA	Relugolix + E2/NETA	Relugolix + E2/NETA
Group 4 Relugolix + Delayed E2/NETA Relugolix + E2/NETA Relugolix + E2/NETA Placebo	Relugolix + Delayed E2/NETA	Relugolix + E2/NETA	Placebo
Group 5 Placebo Relugolix + E2/NETA Relugolix + E2/NETA Relugolix + E2/NETA	Placebo	Relugolix + E2/NETA	Relugolix + E2/NETA
Group 6 Placebo Relugolix + E2/NETA Placebo Placebo	Placebo	Relugolix + E2/NETA	Placebo

Patients who prematurely discontinued from the study treatment after Week 60 will undergo another bone densitometry scan at the safety follow-up visit. Patients with a BMD loss of ≥ 7% at lumbar spine, total hip, or femoral neck at their Week 104/Early Termination visit relative to pivotal study baseline measurement will undergo another bone densitometry scan at 6 (± 1) months and follow-up to evaluate recovery per the protocol. All available post-treatment follow-up BMD data will be listed. BMD data collected during the additional safety follow-up period may be summarized and reported in an addendum to the CSR if deemed necessary.

9.3.4. Endometrial Biopsy and Transvaginal Ultrasound

The endometrial biopsy procedure was performed only if required in this study and consensus readings were not performed. The biopsy related information collected in the eCRF will be listed.

Transvaginal ultrasound was to be performed at Week 52/Baseline visit and unscheduled visit in this study, and the results collected in the eCRF will be listed.

10. REFERENCES

ISCD. ISCD 2015 Official Positions Brochure.pdf | Powered by Box [Internet]. 2015 [cited 2019 Jun 11]. Available from: https://iscd.app.box.com/v/OP-ISCD-2015-Adult

Klein JP, Logan B, Harhoff M, Andersen PK. Analyzing survival curves at a fixed point in time. Stat Med. 2007;26(24):4505–19.

Leslie WD, Lix LM, Tsang JF, Caetano PA, Manitoba Bone Density Program. Single-site vs multisite bone density measurement for fracture prediction. Arch Intern Med. 2007 Aug 13;167(15):1641–7.

Reilly MC, Zbrozek AS, Dukes EM. The validity and reproducibility of a work productivity and activity impairment instrument: PharmacoEconomics. 1993 Nov;4(5):353–65.

APPENDIX 1: PRE-DEFINED CATEGORIES FOR LAB SHIFT TABLE

Clinical Laboratory Parameter	Category
Fasting Glucose/Glucose	Category 1 (< 100 mg/dL) Category 2 (100 - 125 mg/dL) Category 3 (126 - 200 mg/dL) Category 4 (> 200 mg/dL)
HbA1C	Category 1 (< 5.7%) Category 2 (5.7 – 6.4%) Category 3 (6.5 – 8%) Category 4 (8.1 – 9.4%) Category 5 (9.5 – 11%) Category 6 (> 11%)
Total Cholesterol (CTCAE Grade)	Grade 0 (≤ ULN) Grade 1 (> ULN – 300 mg/dL) Grade 2 (> 300 – 400 mg/dL) Grade 3 (> 400 – 500 mg/dL) Grade 4 (> 500 mg/dL)
Low Density Cholesterol (NCEP ATP III Guidelines)	Category 1 (< 100 mg/dL) Category 2 (100 – 129 mg/dL) Category 3 (130 – 159 mg/dL) Category 4 (160 – 189 mg/dL) Category 5 (≥ 190 mg/dL)
High Density Cholesterol (NCEP ATP III Guidelines)	Category 1 (< 40 mg/dL) Category 2 (40 – 59 mg/dL) Category 3 (≥ 60 mg/dL)
Triglycerides (CTCAE Grade)	Grade 0 (< 150 mg/dL) Grade 1 (150 – 300 mg/dL) Grade 2 (301 – 500 mg/dL) Grade 3 (501 – 1000 mg/dL) Grade 4 (> 1000 mg/dL)

Abbreviations: HbA1c = Hemoglobin A1c; CTCAE = Common Terminology Criteria for Adverse Events; NCEP ATP III= National Cholesterol Education Program Adult Treatment Panel III; ULN= upper limit of normal.

APPENDIX 2: LIST OF PRE-DEFINED THRESHOLDS IN SELECTED CHEMISTRY AND HEMATOLOGY TEST RESULTS

Chemistry Laboratory	
Liver Function	
ALT > ULN and < 3× ULN	Total BILI > ULN
ALT ≥ 3× ULN and < 5× ULN	Total BILI >2× ULN
ALT ≥ 5× ULN and < 10× ULN	
ALT ≥ 10× ULN and < 20× ULN	ALT or AST ≥ 3× ULN and Total BILI > 2× ULN
ALT ≥ 20× ULN	ALT or AST ≥ 3× ULN and Total BILI > 2× ULN and ALP < 2× ULN
AST > ULN and < 3× ULN	GGT > ULN and < 3× ULN
AST ≥ 3× ULN and < 5× ULN	GGT ≥ 3× ULN and < 5× ULN
AST ≥ 5× ULN and < 10× ULN	GGT ≥ 5× ULN and < 10× ULN
AST ≥ 10× ULN and < 20× ULN	GGT ≥ 10× ULN and < 20× ULN
AST ≥ 20× ULN	GGT ≥ 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	
Renal Function	
CR > 1.5 mg/dL and > BL	GFR < 15 mL/min per 1.73 m^2
CR > 50% increase from BL	GFR ≥ 15 - < 30 mL/min per 1.73 m^2
	GFR ≥ 30 - < 60 mL/min per 1.73 m^2
	GFR ≥ 60 - < 90 mL/min per 1.73 m^2
	GFR ≥ 90 mL/min per 1.73 m^2

Chemistry Laboratory	
Metabolic Parameters	
Fasting Glucose	Highest Postbaseline Glucose
< 100 mg/dL at BL	Gluc ≥ 200 mg/dL and > BL
< 100 mg/dL	Gluc ≥ 200 mg/dL and ≥ 126 mg/dL at BL
≥ 100 - < 126 mg/dL	Gluc ≥ 500 mg/dL and > BL
≥ 126 mg/dL	Gluc ≥ 500 mg/dL and ≥ 126 mg/dL at BL
≥ 100 - < 126 mg/dL at BL	Total CHOL > 200 mg/dL and > BL
< 100 mg/dL	Total CHOL increase > 30 mg/dL from BL
≥ 100 - < 126 mg/dL	
≥ 126 mg/dL	
≥ 126 mg/dL at BL	HDL < LLN and < BL
< 100 mg/dL	
≥ 100 - < 126 mg/dL	
≥ 126 mg/dL	
	LDL > ULN and > BL
TRIG > ULN and > BL	LDL 100 - < 130 mg/dL and > BL
TRIG 150 - < 200 mg/dL and > BL	LDL 130 - < 160 mg/dL and > BL
TRIG 200 - < 500 mg/dL and > BL	LDL 160 - < 190 mg/dL and > BL
TRIG ≥ 500 mg/dL and > BL	LDL ≥ 190 mg/dL and > BL
Electrolytes and other Chemistry Parameters	
ALB < LLN and < BL	MG < LLN and < BL
ALB > ULN and > BL	MG > ULN and > BL
ALP > 2× ULN and > BL	CK > 2× ULN and > BL
ALP > 5× ULN and > BL	CK > 5× ULN and > BL
ALP > 10× ULN and > BL	CK > 10× ULN and > BL

CA < LLN and < BL	PHOS < LLN and < BL
CA > ULN and > BL	PHOS > ULN and > BL
K < LLN and < BL	NA < LLN and < BL
K > ULN and > BL	NA > ULN and > BL
Hematology Laboratory	
HCT < LLN and < BL	NEUT < LLN and < BL
HCT decrease ≥ 10 % from BL	NEUT > ULN and > BL
HGB ≤ 10.5 g/dL and < BL	BASO < LLN and < BL
HGB decrease > 1 g/dL from BL	BASO > ULN and > BL
MCV < LLN and < BL	EOS < LLN and < BL
MCV > ULN and > BL	EOS > ULN and > BL
	EOS > 5% and > BL
WBC < LLN and < BL	
WBC > ULN and > BL	PLT < LLN and < BL
	PLT < 100 × 10^9/L and < BL
LYM < LLN and < BL	PLT > ULN and > BL
LYM > ULN and > BL	
	HbA1c ≤ 5.6% and > BL
MONO < LLN and < BL	HbA1c 5.7 - 6.4% and > BL
MONO > ULN 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; 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; NEUT = neutrophils; PHOS = phosphate; PLT = platelets; TRIG = triglycerides; ULN = upper limit of normal; WBC = white blood cell.