CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

214846Orig1s000

STATISTICAL REVIEW(S)

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Sciences
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/BLA #:	NDA214846
Product Name:	Relugolix pegol (MYFEMBREE)
Indication(s):	Management of heavy menstrual bleeding associated with uterine fibroids
Applicant:	Myovant Sciences GmbH
Dates:	Date Submitted: 06/26/2010
	PDUFA Goal Date: 06/26/2021
Review Priority:	Standard
Biometrics Division:	DBIV
Statistical Reviewer:	Abel Eshete, PhD
Concurring Reviewers:	Guoxing Soon, PhD, Statistical Team Leader
	Daphne Lin, PhD, Deputy Division Director
Medical Division:	Division of Urology, Obstetrics, and Gynecology (DUOG)
Clinical Team:	Jennifer Lawrence, MD, Medical Officer
	Linda Jaffe, MD, Medical Officer
	Gerald Willett, MD, Medical Team Leader
Project Manager:	Maria Wasilik, PhD

Key words: Uterine fibroids, Menstrual bleeding.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

1 EXECUTIVE SUMMARY

This is the statistical review of the New Drug Application (NDA) submitted by Myovant Sciences GmbH (Applicant) for MYFEMBREE, Relugolix co-administered with low-dose estradiol (E2) and norethindrone acetate (NETA). The proposed indication is for the management of heavy menstrual bleeding associated with uterine fibroids. The primary objective of this review is to evaluate whether the safety and efficacy results in the two pivotal Phase 3 studies submitted in this NDA, Study MVT-601-3001 and Study MVT-601-3002, support the proposed indication. The two studies were randomized, double-blind, placebo-controlled studies. Per the Applicant, placebo was selected as a control because there was no standard of care medical therapy for the long-term treatment of women with uterine fibroids and heavy menstrual bleeding. A total of 388 subjects in Study MVT-601-3001 and 382 subjects in Study MVT-601-3002 were randomized in a 1:1:1 ratio to one of the three treatment groups:

- R+E2/NETA: Relugolix 40 mg tablet co-administered with E2 1 mg and NETA 0.5 mg capsule for 24 weeks
- R+delE2/NETA: Relugolix 40 mg tablet with E2/NETA placebo tablet for 12 weeks followed by Relugolix 40 mg tablet co-administered with E2 1 mg and NETA 0.5 mg capsule for 12 weeks (Note: delayed arm: E2/NETA is delayed by 12 weeks)
- Placebo: Relugolix placebo tablet co-administered with E2/NETA placebo capsule for 24 weeks.

Randomization was stratified by mean screening menstrual blood loss (MBL) volume (< 225 mL versus ≥ 225 mL) and by geographic region (North America versus Rest of the World). The studies had scheduled visits occurring at screening, baseline (Day 1) and every four weeks until Week 24. The primary efficacy endpoint of these studies was the proportion of responders. Responders are women with an MBL volume of < 80 mL and at least a 50% reduction from baseline MBL volume over the last 35 days of treatment.

The primary efficacy analysis provided the treatment difference and the corresponding 95% confidence interval based on the modified intent to treat (mITT: randomized and treated) population. The confidence interval is computed using the stratum adjusted Cochran-Mantel-Haenszel (CMH) method. In all arms, subjects who had less than 4 weeks of treatment exposure and subjects who discontinued the study for lack of efficacy, or due to surgical procedures for uterine fibroids were considered as non-responders. For subjects with treatment exposure exceeding 4 weeks, the responder status was determined based on the observed MBL data (for those who had complete information) or using imputed data. Please refer to the statistical methods section for the detailed summary of the missing data and responder status derivation methods.

The following seven key secondary efficacy endpoints were analyzed in these studies:

1. Proportion of women who achieved amenorrhea over the last 35 days of treatment
2. Percent change from baseline to Week 24 in MBL volume

3. Change from baseline to Week 24 in Bleeding and Pelvic Discomfort (BPD) scale score as measured by the UFS-QoL (Question 1, Question 2, Question 5)
4. Proportion of women with a hemoglobin ≤10.5 g/dL at baseline who achieved an increase of ≥2 g/dL from baseline at Week 24
5. Proportion of women who achieved a maximum Numerical Rating Scale (NRS) score ≤ 1 for uterine fibroids associated pain over the last 35 days of treatment in the subset of women with a maximum pain score ≥ 4 during the 35 days prior to randomization
6. Percent change from baseline to Week 24 in uterine fibroid volume
7. Percent change from baseline to Week 24 in uterine volume.

The analysis of the binary secondary efficacy endpoints (1, 4 and 5) was conducted using the CMH method applied for the primary efficacy analysis. For the secondary efficacy endpoints 2 and 3, treatment comparisons were performed using a mixed model for repeated measures approach, whereas, an analysis of covariance (ANCOVA) model was used for the secondary efficacy endpoints 6 and 7. To adjust for multiplicity due to the treatment comparisons of the primary and multiple secondary efficacy endpoints, the studies used a gate-keeping mixed-sequence testing procedure. In this testing procedure, the primary efficacy endpoint was tested first at a two-sided 0.05 significance level. If the p-value for the primary endpoint was < 0.05, the first four key secondary endpoints were to be tested sequentially until a non-significant effect is detected. If the two-sided p-value for the fourth endpoint was < 0.05, the Hochberg step-up procedure would be used to adjust for multiplicity due to the test of the remaining three endpoints (the fifth, sixth and seventh ranked endpoints).

Reviewer's remark: On May 7, 2019, the Applicant submitted a common statistical analysis plan (SAP) for the analysis of the two pivotal studies. This SAP was reviewed and accepted by the Agency. However, on June 14, 2019, the Applicant submitted an amendment to their SAP for MVT-601-3002 alone. This amendment was finalized prior to unblinding of study MVT-601-3002 but after unblinding of MVT-601-3001. The amendment changed the order of testing of the hemoglobin and the NRS endpoints in Study MVT-601-3002. In the original SAP, the hemoglobin and NRS endpoints were to be tested in both studies as the 4th and 5th ranked endpoints, respectively. After the amendment, the order of testing is reversed, with NRS tested as the 4th endpoint and hemoglobin as the 5th endpoint, in study MVT-601-3002. Although major changes to an SAP after a substantial number of study subjects have completed safety and efficacy evaluations should be avoided, because this amendment was made before unblinding of Study MVT-601-3002, in accordance with the ICH guidelines, the amendment was accepted by the Agency.

Reviewer's remark: On June 10, 2020, the Applicant informed the Agency that they closed study Site 1152 due to multiple observations of noncompliance with Good Clinical Practice (GCP). Site 1152 enrolled 27 patients in Study MVT-601-3001 [R+E2/NETA (N=6), R+delE2/NETA (N=7), placebo (N=14)]. (b) (4) *The Agency* (b) (4) *decided that the analysis results excluding data from Site 1152 will be used to evaluate the primary and secondary efficacy endpoints in Study MVT-601-3001. Therefore, the efficacy results in this and subsequent sections are based on the analysis excluding data from this site. However, safety data from Site 1152 will be included in*

all safety analyses. Note, although there is a slight change in the magnitude of the treatment effects, the overall conclusions remain consistent when the analysis of the primary and secondary efficacy endpoints are conducted with subjects from Site 1152 included.

The Applicant's findings for the primary and secondary efficacy endpoints in the two studies established the efficacy of R+E2/NETA for the management of heavy menstrual bleeding associated with uterine fibroids (Table 1, Table 2, Figure 4 and Figure 5). The Applicant also presented the analyses of the primary efficacy endpoint across various patient subgroups, different analysis populations and defining the responder status under different assumptions. Results from these analyses are generally consistent with the primary analysis findings. In addition, the Applicant's findings for the R+delE2/NETA arm are also consistently favorable (Table 1, Figure 6, Figure 7 and Table 36).

Regarding safety, the percentage of subjects who reported at least one adverse event was 61.0% in the R+E2/NETA group, 72.1% in the R+delE2/NETA group, compared to 62.5% in the placebo group. The percentage of subjects who reported at least one serious adverse event was slightly higher in the R+E2/NETA group (3.1%) compared to both the R+delE2/NETA group (1.9%) and the placebo group (2.3%). No death was reported in either study. Adverse events leading to study discontinuation were reported for 3.9% (10/254) subjects in the R+E2/NETA group, 11.6% (30/258) subjects in the R+delE2/NETA group, and 4.3% (11/256) subjects in the placebo group. The three most frequently reported adverse events in the R+E2/NETA group were headache (9.8%), hot flash (8.3%), and hypertension (4.7%). The corresponding figures in the placebo group were 13.3% [headache], 5.9% [hot flash] and 1.6% [hypertension]. The incidence rate of these events in the R+delE2/NETA group were (16.3% [headache], 35.3% [hot flash] and 3.9% [hypertension]).

Table 1: Proportion of responders at Week 24 (mITT)

	Treatments			**Difference (95% CI)**	
Study	**R+E2/NETA**	**R+delE2/NETA**	**Placebo**	**R+E2/NETA vs Placebo**	**R+delE2/NETA vs Placebo**
MVT-601-3001	88 /122 (72.1%)	98/125 (78.4%)	19 /113 (16.8%)	55.3% (44.2%, 65.6%)	61.6% (51.0%, 71.0%)
MVT-601-3002	89 /125 (71.2%)	93/127 (73.2%)	19/129 (14.7%)	56.5% (46.6%, 66.5%)	58.5% (48.6%, 68.4%)

Source: Adapted from Table 6 of the Sponsor's response to information request sent on 02/08/2021 (MVT-301-601-3001; the 95% confidence intervals are adjusted for the stratifying variables); Table 7.2.1.1 of the Study report (MVT-601-3002).

Table 2: Summary of key secondary efficacy endpoints

	Study MVT-601-3001			**Study MVT-601-3002**		
Alpha -adjusted Endpoints	**R+E2/NETA**	**Placebo**	**Difference (95% Cl)**	**R+E2/NETA**	**Placebo**	**Difference (95% CI)**
Proportion of women who achieved amenorrhea over the last 35 days of treatment	61/122 (50.0%)	7/113 (6.2%)	43.8% (33.9%, 53.7%)	63/125 (50.4%)	4/129 (3.1%)	47.3% (38.0%, 56.6%)
Percent change from baseline to Week 24 in MBL volume [mean (SD)]	-82.0 (5.09)	-19.1 (5.2)	-62.9 (-76.4, -49.4)	-84.3 (5.45)	-15.1 (5.47)	-69.2 (-84.1, -54.3)
Change from baseline to Week 24 in UFS-QoL BPD scale score as measured by UFS-QoL (Q1, Q2, Q5) [mean (SD)]	-44.1 (3.13)	-16.4 (3.24)	-27.7 (-35.7, -19.7)	-51.7 (2.89)	-18.3 (2.93)	-33.4 (-41.2, -25.5)
Proportion of women with hemoglobin level ≤10.5g/dL at baseline who achieved an increase of >2 g/dL from baseline at Week 24	15/30 (50.0%)	5/22 (22.7%)	27.3% (2.24%, 52.3%)	19/31 (61.3%)	2/37 (5.4%)	55.9% (37.2%, 74.5%)

Proportion of women who achieved a maximum NRS score ≤1 for uterine fibroid-associated pain over the 35 days of treatment in the subset of women with a maximum pain score ≥4 during the 35 days prior to randomization	23/55 (41.8%)	7/61 (11.5%)	30.3% (15.1%, 45.6%)	32/68 (47.1%)	14/82 (17.1%)	30.0% (15.6%, 44.4%)
Percent change from baseline to Week 24 in primary uterine fibroid volume [mean (SD)]	-12.6 (5.95)	-0.01 (6.19)	-12.6 (-27.7, 2.52)	-17.4 (5.9)	-7.4 (5.92)	-10.0 (-25.8, 5.8)
Percent change from baseline to Week 24 in uterine volume [mean (SD)]	-12.9 (3.08)	2.2 (3.01)	-15.1 (-23.0, -7.3)	-13.8 (3.4)	-1.5 (3.37)	-12.2 (-21.3, -3.2)

Source: Adapted from Table 6 of the Sponsor's response to information request sent on 02/08/2021. The summary of the hemoglobin endpoint is provided by this reviewer. Mean: Least-squares means; SD: Standard deviation.

2 INTRODUCTION

This is the statistical review of the New Drug Application (NDA) submitted by Myovant, referred to as the Applicant, on June 26, 2020 for Relugolix co-administered with low-dose estradiol (E2) and norethindrone acetate (NETA). The proposed indication is for the management of heavy menstrual bleeding associated with uterine fibroids. The primary evidence of efficacy and safety for this NDA comes from two pivotal Phase 3 studies (Studies MVT-601-3001 and MVT-601-3002).

The two studies were conducted across multiple sites in the United States (US), and globally including sites in Brazil, Italy, Poland, South Africa, and the United Kingdom. Study MVT-601-3001 enrolled 388 subjects across 80 sites, of which, 309/388 (79.6%) subjects were enrolled in 63/80 (78.7%) sites located within the US. Similarly, Study MVT-601-3002 enrolled 382 subjects across 99 sites, of which, 284/382 (74.3%) subjects were enrolled in 67/99 (67.7%) sites located within the US.

The Applicant proposes to include findings from Study MVT-601-3001 and Study MVT-601-3002 into the "Clinical Studies" (Section 14) of the US Prescribing Information (USPI) to describe the efficacy of R+E2/NETA for the management of heavy menstrual bleeding associated with uterine fibroids. This review investigates whether the findings from these studies support the proposed indication and provides recommendations for the USPI to be considered by the division of urology, obstetrics and gynecology (DOUG), if the product is approved.

2.1 Overview

This section provides a brief overview of the class and indication of the studied drug, the history of the drug development and outlines the Applicant's summary of the specific studies reviewed.

2.1.1 Drug Class and Indication

The Applicant is developing a once daily Relugolix 40 mg in combination with estradiol (1.0 mg) and a progestin (0.5 mg norethindrone acetate) for the management of heavy menstrual bleeding associated with uterine fibroids. Per the Applicant, Relugolix is an oral gonadotropin-releasing hormone (GnRH) receptor antagonist designed to reduce the production of ovarian

estradiol, a hormone involved in uterine fibroids growth. The Applicant notes that estradiol and progestins are hormones similar to those naturally occurring in the body. Per the Applicant, low-estrogen levels can weaken bone health and cause unwanted side effects such as hot flashes. Low doses of estradiol and progestins help offset these effects. To this effect, the Applicant is investigating whether their combination product maintains estrogen in the low normal range to balance the potential benefits of Relugolix, while also maintaining bone health and mitigating the side effects that can result from a low-estrogen state.

2.1.2 History of Drug Development

The Applicant had a pre-IND meeting with the Agency on 6-15-2007 (b) (4) During the meeting, agreement was reached on the need for a thorough QT/QTc study, ocular surveillance, the plan to open the IND with a first-in-human study in healthy pre-menopausal women, and the conduct of two Phase 3 efficacy trials. FDA stated that a reasonable proportion of the subjects should be from the U.S. and Canada. The Applicant was informed that reduction in fibroid volume is not an acceptable primary endpoint and that ICH guidelines stipulate a safety database with 1500 total exposed subjects, 300-600 subjects exposed for 6 months, and 100 subjects exposed for 1 year. FDA also recommended that the Applicant (b) (4)

The Applicant had an End of Phase 2 meeting with the Agency on 10-5-2016. During this meeting, agreement was reached on the need for a food effect study using the to-be-marketed formulation, the use of 3 arms in the Phase 3 studies (Relugolix 40mg, Relugolix 40mg plus add-back, and placebo), the inclusion and exclusion criteria, the primary endpoint as the proportion of subjects with a menstrual blood loss volume of less than 80ml and at least a 50% reduction in menstrual blood loss volume from baseline to the last month of treatment, as measured by the alkaline hematin method, and the safety assessment and monitoring plan. FDA informed the Applicant that the difference in responder rates between treatment arms must be both statistically significant and clinically meaningful. The Applicant informed the FDA that a majority of subjects enrolled would be from the U.S.

On 8-22-2018, the Applicant had a teleconference with the FDA. During this meeting, the FDA advised the Applicant to use the to-be-marketed product in the Phase 3 program. A Type C guidance meeting was held on 1-31-2019 to discuss the statistical analysis plan (SAP) for the Phase 3 trials. On 4-26-2019, FDA issued an advice letter regarding the SAP. In the letter, the FDA recommended that subjects who withdraw for lack of efficacy or the need for surgery to be considered non-responders in the primary efficacy analysis. The FDA also noted that the proposed mixed effect model for imputing missing data might underestimate menstrual blood loss volume for subjects with partial feminine product information and requested the Applicant to clarify. On 5-7-2019, the Applicant responded to the 4-26-2019 letter by agreeing that subjects who withdraw for lack of efficacy or the need for surgery would be considered non-responders but disagreed that the mixed-effects model may underestimate menstrual blood loss volume.

The Applicant had a pre-NDA meeting with the FDA on 3-25-2020. On this meeting, agreement was reached on the need to address the potential interaction between calcium and/or

Vitamin D supplementation and bone mineral density (BMD), that the original NDA submission would include final safety data from the extension trial, and that case report forms and narratives would be included for several adverse events, including alopecia. FDA also outlined how to present post-treatment BMD results and that the 120-day safety update should include information from any studies (including endometriosis trials). The Applicant explained that they planned to submit a 505(b)(2) application relying on the Agency's finding of safety and effectiveness for Activella (estradiol/norethindrone acetate) tablets and to leverage the Guidance for Industry: Labeling for Combined Hormonal Contraceptives for language in the Pregnancy subsection of labeling. The FDA agreed to consider this proposal. The FDA informed the Applicant that the NDA would not be granted a priority review. FDA offered additional advice on the content and format of the NDA which was not discussed at the meeting.

2.1.3 Studies Reviewed

The Applicant's overall efficacy summary for Studies MVT-601-3001 and MVT-601-3002 is presented in Table 3.

Table 3: Efficacy summaries of MVT-601-3001 and MVT-601-3002

Design	Treatment	Endpoint/Analysis	Applicant's findings
MVT-601-3001 [1]MC, RD, DB, PG, PC	• R+E2/NETA (N=122) • Placebo (N=113) • R+delE2/NETA (N=125)	**Primary Endpoint**: Proportion of responders, defined as subjects with a menstrual blood loss (MBL) volume of < 80 mL and at least a 50% reduction from baseline MBL volume over the last 35 days of treatment. The primary efficacy analysis provided the difference in responder rates between the R+E2/NETA group and the placebo group and its two-sided 95% CI using stratum-adjusted Cochran-Mantel-Haenszel (CMH) method stratified by the baseline mean screening MBL volume (< 225 mL versus ≥ 225 mL) and by the geographic region (North America versus Rest of World). The primary efficacy analysis is conducted based the modified intent to treat (mITT) population (all randomized and treated).	The study met its primary objective of demonstrating the efficacy of Relugolix co-administered with low-dose estradiol and norethindrone acetate for the treatment of heavy menstrual bleeding associated with uterine fibroids. The number of responders in the R+E2/NETA group was 88 (72.1%) compared with 19 (16.8%) in the placebo group. The observed difference between the two groups was 55.3% (95% CI: 44.2% to 65.6%) in favor of the R+E2/NETA group.
MVT-601-3002	• R+E2/NETA (N=125)	**Primary Endpoint**: Proportion of responders,	The study met its primary objective of

[1]MC, RD, DB, PG, PC	• Placebo (N=129) • R+delE2/NETA (N=127)	defined as subjects with a menstrual blood loss (MBL) volume of < 80 mL and at least a 50% reduction from baseline MBL volume over the last 35 days of treatment. The primary efficacy analysis provided the difference in responder rates between the R+E2/NETA group and the placebo group and its two-sided 95% CI using stratum-adjusted Cochran-Mantel-Haenszel (CMH) method stratified by the baseline mean screening MBL volume (< 225 mL versus ≥ 225 mL) and by the geographic region (North America versus Rest of World). The primary efficacy analysis is conducted based the modified intent to treat (mITT) population (all randomized and treated).	demonstrating the efficacy of Relugolix co-administered with low-dose estradiol and norethindrone acetate for the treatment of heavy menstrual bleeding associated with uterine fibroids. The percentage of responders in the R+E2/NETA group was 89 (71.2%) compared with 19 (14.7%) in the placebo group. The observed difference between the two groups was 56.5% (95% CI: 46.6% to 66.5%) in favor of the R+E2/NETA group.

Source: Applicant's Study Reports. [1]MC: multicenter, RD: randomized, DB: double-masked, PG: parallel-group, PC: placebo-controlled. The missing data handling method and responder status definition is summarized in Table 4.

2.2 Data Sources

This NDA was submitted electronically and includes full study reports as well as standardized datasets using SDTM and ADaM formats that are relevant for the analyses of studies MVT-601-3001 and MVT-601-3002 presented in this review. Datasets and corresponding definition files can be found at the following location:

\\cdsesub1\evsprod\NDA214846\0001\m5\datasets

For each study, the following datasets submitted by the Applicant are used in this statistical review:

- adsl.xpt contains the demographic and disposition data
- admbl.xpt contains the menstrual blood volume data (primary efficacy outcome)
- adbmd.xpt contains the bone mineral density data
- adae.xpt contains the adverse event data
- adbase.xpt contains the baseline characteristics data

3 STATISTICAL EVALUATION

3.1 Data and Analysis Quality

The quality of the datasets and analyses conducted by the Applicant are acceptable. The data definition files, and reviewer's guide submitted in the NDA were sufficiently detailed to facilitate replication of the findings from the Applicant's primary analysis and other major analyses using the submitted datasets.

3.2 Evaluation of Efficacy

This section summarizes the design of studies MVT-601-3001 and MVT-601-3002 and the corresponding efficacy results submitted by the Applicant and produced by the reviewer's analyses.

3.2.1 Design of Clinical Trials Intended to Demonstrate Benefit to Patients

3.2.1.1 Trial Design

Both MVT-601-3001 and MVT-601-3002 were multicenter, double-masked, randomized, parallel-group, placebo-controlled studies. These studies were designed to evaluate the safety and efficacy of Relugolix co-administered with low-dose estradiol (E2) and norethindrone acetate (NETA) compared with placebo in subjects with heavy menstrual bleeding associated with uterine fibroids.

3.2.1.2 Eligibility Criteria

To be eligible for these studies, patients who had signed informed consent had to meet the following inclusion criteria:

- Premenopausal female aged 18 to 50 years old (inclusive) on the day of signing and dating the informed consent form;

- Had regularly occurring menstrual periods of ≤ 14 days duration with a cycle of 21 to 38 days from the start of one menstrual period until the start of the next, by patient history for at least three months prior to the screening 1 visit;

- Had a diagnosis of uterine fibroids that was confirmed by a transvaginal ultrasound performed during the screening period; at least one uterine fibroid had to be verified by a central reader to meet at least one of the following criteria:

- Subserosal, intramural, or < 50% intracavitary submucosal fibroid with a diameter ≥ 2 cm (longest diameter), or

- Multiple small fibroids with a total uterine volume of ≥ 130 cm^3

- Had heavy menstrual bleeding associated with uterine fibroids as evidenced by an MBL volume of ≥ 160 mL during 1 cycle or ≥ 80 mL per cycle for 2 menstrual cycles as measured by the alkaline hematin method

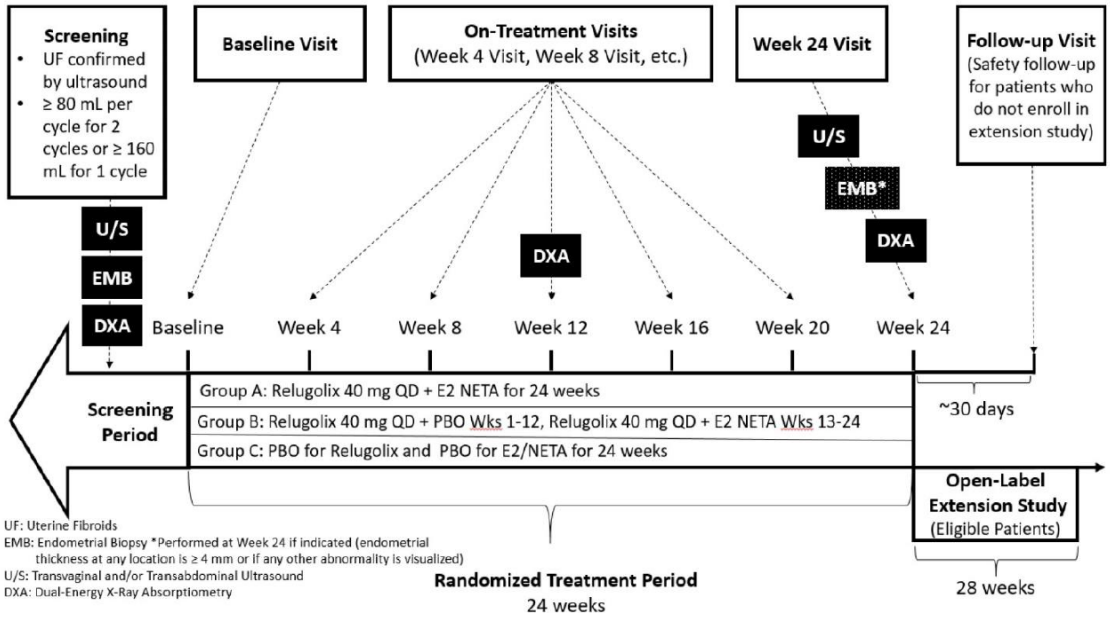
3.2.1.3 Randomization and Treatment

Both studies used a 1:1:1 randomization ratio for allocating eligible patients to the three study treatments:

- R+E2/NETA: Relugolix 40 mg tablet co-administered with E2 1 mg and NETA 0.5 mg capsule for 24 weeks
- R+delE2/NETA: Relugolix 40 mg tablet with E2/NETA placebo tablet for 12 weeks followed by Relugolix 40 mg tablet co-administered with E2 1 mg and NETA 0.5 mg capsule for 12 weeks (*delayed arm: The E2/NETA part is delayed by 12 weeks*)
- Placebo: Relugolix placebo tablet co-administered with E2/NETA placebo capsule for 24 weeks.

The total duration of both studies was 24 Weeks. The studies had scheduled visits occurring at screening, baseline (Day 1), and every 4 weeks till Week 24. Both studies also had an open-label safety follow-up of up to 28 Weeks. During the open-label safety extension period, all subjects were to receive R+E2/NETA (See Study Design Schema in Figure 1).

Figure 1: Study design schema



Randomization was stratified by mean screening menstrual blood loss (MBL) volume (< 225 mL versus ≥ 225 mL) and by geographic region (North America versus Rest of World).

3.2.1.4 Efficacy Endpoints

The primary efficacy endpoint of these studies was the proportion of responders. Responders are women with an MBL volume of < 80 mL and at least a 50% reduction from baseline MBL volume over the last 35 days of treatment.

The two studies had the following key secondary efficacy endpoints:

1. Proportion of women who achieved amenorrhea over the last 35 days of treatment
2. Percent change from baseline to Week 24 in MBL volume
3. Change from baseline to Week 24 in Bleeding and Pelvic Discomfort (BPD) scale score as measured by the UFS-QoL (Question 1, Question 2, Question 5)
4. Proportion of women who achieved a maximum NRS score ≤ 1 for uterine fibroids associated pain over the last 35 days of treatment in the subset of women with a maximum pain score ≥ 4 during the 35 days prior to randomization
5. Proportion of women with a hemoglobin ≤10.5 g/dL at baseline who achieved an increase of ≥2 g/dL from baseline at Week 24
6. Percent change from baseline to Week 24 in uterine fibroid volume
7. Percent change from baseline to Week 24 in uterine volume.

3.2.2 Statistical Methods

This section describes the statistical hypotheses, sample size calculations, analyses populations and the efficacy analyses presented in this review that are performed by the Applicant, as described in the statistical analysis plan (SAP) for studies MVT-601-3001 and MVT-601-3002, as well as independent analyses performed by the statistical reviewer.

3.2.2.1 Statistical Hypotheses and Sample size

Hypotheses

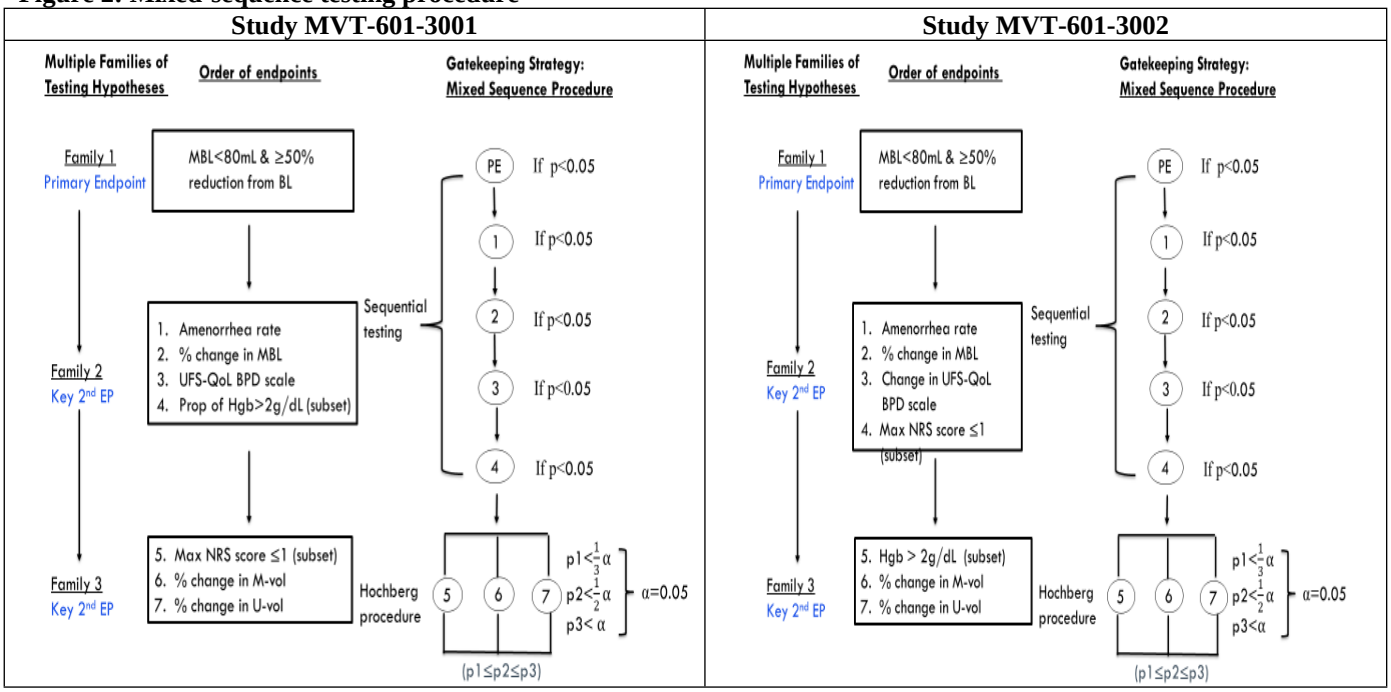
The primary null and alternative hypotheses can be mathematically stated as follows:

$$H_{o1}: P_R - P_p \leq 0\%$$
$$H_{a1}: P_R - P_p > 0\%,$$

where P_R, and P_p are the probabilities of responders for the R+E2/NETA and placebo treatment arms, respectively. A conclusion that R+E2/NETA is superior to placebo is made if the lower bound for the 2-sided 95% confidence interval (CI) for the difference in proportions is greater than 0%.

If the R+E2/NETA arm has demonstrated superiority over placebo for the primary endpoint, additional comparisons to placebo were to be made with respect to the seven key secondary efficacy endpoints based on the predefined mixed-sequence testing procedure presented in Figure 2.

Figure 2: Mixed-sequence testing procedure



Sample Size Calculation

In each trial, the Applicant planned to enroll a total of 390 subjects. This sample size was calculated to provide approximately 99% power to detect a 30% difference in responder rates between R+E2/NETA and placebo. This calculation assumes that the tests will be conducted at 2-sided alpha level of 0.05 based on a chi-square test; a responder rate of 25% in the placebo group and a dropout rate of 20%. MVT-601-3001 enrolled 388 subjects [R+E2/NETA (N=128), R+delE2/NETA (N=132), and Placebo (N=128)]. Similarly, in MVT-601-3002, 382 subjects [R+E2/NETA (N=126), R+delE2/NETA (N=127), and Placebo (N=129)] were enrolled.

3.2.2.2 Analysis Populations

The SAP defined the following analysis populations:

- The intent-to-treat (ITT): Includes all randomized subjects.
- The modified intent-to-treat (mITT): Includes all randomized and treated subjects.
- The per-protocol (PP): Includes all randomized and treated subjects who do not have protocol deviations that impact the primary efficacy variable.
- The safety population: Includes all treated subjects (subjects who received at least one dose of the study medications).

3.2.2.3 Analysis Methods

A. Primary Efficacy Analysis

The primary efficacy analysis in both studies provided the treatment difference in the proportions of subjects with an MBL volume of < 80 mL and at least a 50% reduction from baseline MBL volume over the last 35 days of treatment and the corresponding 2-sided 95% confidence interval. The confidence intervals were computed using the CMH weighted method with the baseline (< 225 mL versus ≥ 225 mL) and geographic region (North America versus Rest of World) used as stratification factors. The primary efficacy analysis was conducted based on the mITT population.

Subjects who discontinued the study due to lack of efficacy or those who required surgical intervention for uterine fibroids were considered as non-responders in the primary efficacy analysis. The Applicant devised the approach summarized in Table 4 to deal with missing data and to define the responder status. The approach takes duration of treatment exposure, feminine product (FP) return rate and information availability in the e-diary into consideration.

Table 4: Derivation of responder status and handling of missing data

Treatment Exposure	FP Collection	Observed MBL Volume	Reason for No FP collection	Responder Status
<4 weeks	N/A	N/A	N/A	Imputed as non-responder
≥4 weeks	100% FP compliance	N/A	N/A	Based on observed MBL volume
	<100% FP compliance	MBL≥80mL or <50% reduction from baseline	N/A	Imputed as non-responder based on observed MBL volume
		MBL<80mL and ≥50% reduction from baseline	N/A	Based on Imputed MBL volume**
	No FP Collection	N/A	Reported "Amenorrhea"	Imputed as responder
			Reported "Spotting or negligible bleeding" and confirmed by eDiary	Imputed as responder
			Reported "Spotting or negligible bleeding" although not confirmed by eDiary or any other reason, had at least 8 weeks of MBL data	Based on the imputed MBL volume**
			The entries in the eDiary did not verify Spotting or negligible bleeding" or any other reason and if had less than 8 weeks of MBL data	Imputed as non-responder

Source: Table 7 of the Applicant's study reports. ** Responder status is determined based on predicted MBL volume from a mixed-effects model. The model included baseline MBL volume, geographic region, visit (Weeks 4, 8, 12, 16, 20 and 24), treatment, and treatment by visit interaction as fixed effects. An unstructured variance-covariance matrix was used to account for within-subject correlation.

B. Analysis of Key Secondary Efficacy Endpoints

As stated earlier, the studies had seven key secondary efficacy endpoints. The analysis of the binary secondary efficacy endpoints (1, 4 and 5) was conducted using the CMH method that was used for the primary efficacy analysis. For the secondary efficacy endpoints 2 and 3, which evaluated the change (absolute or % change) from baseline to Week 24, treatment comparisons were performed using a mixed model repeated measures approach. The model included treatment, visit, randomization stratification factors and treatment by visit interactions as fixed effects, and an unstructured variance-covariance matrix was used to account for within subject correlation. For the secondary efficacy endpoints 6 and 7, an analysis of covariance (ANCOVA) model with treatment, randomization stratification factors and baseline value included as covariates was used.

3.2.3 Patient Disposition, Demographic and Baseline Characteristics

3.2.3.1 Patient Disposition

The patient disposition provided in Table 5 shows that, among the three arms, the R+E2/NETA arm had the highest incidence of treatment discontinuation due to lack of efficacy in both studies. In addition, in Study MVT-601-3001, a higher percentage of R+E2/NETA treated subjects discontinued the study due to adverse events compared to the placebo arm; but the incidence rate was lower than the R+delE2/NETA arm. On the other hand, in Study MVT-601-3002, the incidence of treatment discontinuation due to adverse events was the lowest in the R+E2/NETA arm compared to both R+delE2/NETA and placebo. Overall, "withdrawal by patient" was the main reason for study discontinuation for subjects in the R+E2/NETA arm.[1]

Table 5: Patient disposition (Week 24)

	Study MVT-601-3001			**Study MVT-601-3002**		
	R+E2/NETA	**R+delE2/ NETA**	**Placebo**	**R+E2/NETA**	**R+delE2/ NETA**	**Placebo**
Randomized	128	132	128	126	127	129
Treated[1]	128 (100%)	132 (100%)	127 (99.4%)	125 (99.2%)	127 (100%)	129 (100%)
Completed the Week 24	100 (78.1%)	103 (78.0%)	105 (82.0%)	102 (81.0%)	98 (77.2%)	102 (79.1%)
Discontinued the Study[2]	28 (21.9%)	29 (22.0%)	22 (17.2%)	23 (18.3%)	29 (22.8%)	27 (20.9%)
Reason for Subject Withdrawal						
Adverse Event(s)	7 (5.5%)	18 (13.6%)	5 (3.9%)	2 (1.6%)	15 (11.8%)	6 (4.7%)
Lack of Efficacy	4 (3.1%)	0 (0.0%)	3 (1.0%)	2 (1.6%)	1 (0.8%)	1 (0.8%)
Lost to Follow-Up	1 (0.8%)	5 (3.8%)	5 (3.9%)	4 (3.2%)	2 (1.6%)	7 (5.4%)
Withdrawal by Patient	10 (7.8%)	3 (2.3%)	7 (5.5%)	13 (10.3%)	6 (4.7%)	6 (4.7%)
Protocol Deviation	1 (0.8%)	0 (0.0%)	0 (0.0%)	1 (0.8%)	2 (1.6%)	1 (0.8%)
Other	5 (3.9%)	3 (2.3%)	1 (0.8%)	1 (0.8%)	3 (2.4%)	5 (3.9%)
Pregnancy	0 (0.0%)	0 (0.0%)	1 (0.8%)	0 (0.0%)	0 (0.0%)	1 (0.8%)

Source: Adapted from Figure 3 of the Applicant's Study Reports. [1] Included in the mITT population. [2]Discontinued the study prior to Week 24.

[1]Reviewer's remark: A closer look into subjects whose reason for discontinuation was reported as "withdrawal by patient" shows that, although some subjects discontinued the study for reasons that appear to be treatment related (for e.g. planning to have a surgery or not seeing any improvement), the majority discontinued the study for reasons that appear to be personal in

nature and unrelated to treatment. However, subjects whose reasons are reported as "Other" are mostly subjects who did not comply with the study conduct (Table 38).

The summary of subjects included in the different analysis populations is provided in Table 6. Except for one subject each from the placebo and R+E2/NETA arms, all randomized subjects are included in the mITT population. In both studies, slightly more subjects were excluded from the per-protocol population in the two Relugolix arms compared to the placebo arm.

Table 6: Summary of subjects included in different analysis populations

	Study MVT-601-3001			**Total**
	R+E2/NETA	**R+delE2/ NETA**	**Placebo**	
All Randomized	128 (100%)	132 (100%)	128 (100%)	388
mITT	128 (100%)	132 (100%)	127 (99.2%)	387
Per-protocol	119 (93.0%)	123 (93.2%)	121 (94.5%)	363
Safety	128 (100%)	132 (100%)	127 (99.2%)	387
	Study MVT-601-3002			**Total**
	R+E2/NETA	**R+delE2/ NETA**	**Placebo**	
All Randomized	126 (100%)	127 (100%)	129 (100%)	382
mITT	125 (99.2%)	127 (100%)	129 (100%)	381
Per-protocol	113 (89.7%)	120 (94.5%)	122 (94.6%)	355
Safety	125 (99.2%)	126 (99.2%)	129 (100%)	381

Source: Figure 4 of the Applicant's Study reports.

3.2.3.2 Demographic and Baseline Characteristics

The majority (over 74%) of participants are from the north American region (US and Canada). A larger proportion of subjects in all arms were older than 40 years, with the average age in each group being around 42 years. Study participants were predominantly white or black with very few Asian participants and people of mixed race.

Table 7: Baseline and demographic characteristics

	Study MVT-601-3001			**Study MVT-601-3002**		
	R+E2/NETA N=128	**R+delE2/NETA N=132**	**Placebo N=127**	**R+E2/NETA N=125**	**R+delE2/NETA N=127**	**Placebo N=129**
Age (years)						
Mean (SD)	42.5 (4.99)	41.3 (5.39)	42.2 (5.7)	42.4 (5.38)	42.1 (5.25)	41.8 (5.26)
Age Categories						
<40 years	30 (23.4%)	49 (37.1%)	36 (28.3%)	32 (25.6%)	39 (30.7%)	42 (32.6%)
≥40 years	98 (76.6%)	83 (62.9%)	91 (71.7%)	93 (74.4%)	88 (69.3%)	87 (67.4%)
Geographic Region						
North America	98 (76.6%)	101 (76.5%)	98 (77.2%)	93 (74.4%)	94 (74.0%)	96 (75.4%)
Rest of World	30 (23.4%)	31 (23.5%)	29 (22.8%)	32 (25.6%)	33 (26.0%)	33 (25.6%)
Race						
American Indian or Native	2 (1.6%)	5 (3.8%)	1 (0.8%)	0 (0.0%)	2 (1.6%)	1 (0.8%)
Asian	0 (0.0%)	2 (1.5%)	2 (1.6%)	0 (0.0%)	3 (2.4%)	1 (0.8%)
Black or African American	59 (46.1%)	67 (50.8%)	65 (51.2%)	62 (49.6%)	66 (52.0%)	74 (57.4%)
Native Hawaiian	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
White	64 (50.0%)	53 (40.2%)	56 (44.1%)	58 (46.4%)	50 (39.4%)	49 (38.0%)
Other	2 (1.6%)	1 (0.8%)	1 (0.8%)	1 (0.8%)	2 (1.6%)	3 (2.3%)
Multiple	1 (0.8%)	4 (3.0%)	2 (1.6%)	1 (0.8%)	1 (0.8%)	0 (0.0%)
Not reported	0 (0.0%)	0 (0.0%)	0 (0.0%)	3 (2.4%)	3 (2.4%)	1 (0.8%)
Ethnicity						
Not Hispanic or Latino	92 (71.9%)	99 (75.0%)	103 (81.1%)	105 (84.0%)	91 (71.7%)	96 (74.4%)

Hispanic or Latino	34 (26.6%)	33 (25.0%)	23 (18.1%)	18 (14.4%)	34 (26.8%)	32 (24.8%)
Not reported	2 (1.6%)	0 (0.0%)	1 (0.8%)	2 (1.6%)	2 (1.6%)	1 (0.8%)
BMI (kg/m2)						
Mean (SD)	31.4 (7.6)	31.3 (7.3)	32.3 (7.5)	31.0 (6.6)	30.8 (5.7)	32.1 (7.6)
Median	30.3	31.0	31.1	31.4	30.4	30.8
Min, Max	17.0, 62.8	19.3, 52.0	19.4, 54.7	17.3, 48.8	19.1, 50.4	16.0, 60.1
MBL volume (mL)						
Mean (SD)	239.4 (180.2)	228.8 (159.6)	218.8 (125.0)	246.7 (186.0)	227.4 (134.3)	211.7 (129.9)
Median	193.5	193.5	185.7	184.1	192.8	177.4
Min, Max	88.3, 1365.1	81.2, 1036.5	84.7, 754.5	88.5,1045.5	87.6, 800.2	85.5,839.6
MBL volume (mL)						
<225	84 (65.6%)	86 (65.2%)	85 (66.9%)	80 (64.0%)	80 (63.0%)	86 (66.7%)
≥225	44 (34.4%)	46 (34.8%)	42 (33.1%)	45 (36.0%)	47 (37.0%)	43 (33.3%)
Index uterine fibroid volume						
Mean (SD)	71.9 (128.1)	93.8 (143.8)	71.8 (123.9)	73.7 (126.7)	78.9 (157.5)	74.1 (123.0)
Median	24.2	37.9	27.9	29.2	37.2	31.4
Min, Max	1.5, 989.0	1.0, 904.2	0.8, 1031.2	2.2, 944.0	0.8, 1378.3	1.3, 866.5
Uterine volume						
Mean (SD)	379.1 (316.8)	469.9 (427.9)	397.8 (324.8)	387.7 (344.0)	402.6 (371.1)	407.8 (402.0)
Median	265.1	329.9	293.2	274.3	266.8	295.9
Min, Max	56.6, 1580.1	46.5,2555.6	65.5,2015.1	79.4, 1907.6	52.3,2381.8	58.2,2625.0

Source: Table 9 of the Applicant's Study Reports. Subjects from the closed site, Site 1152 are included in these summaries.

3.2.4 Results and Conclusions

3.2.4.1 Efficacy Results

Per the Applicant, Site 1152 from MVT-601-3001 was closed due to non-compliance with GCP. The Agency decided to present the efficacy analyses results conducted excluding the 27 subjects enrolled in this site [R+E2/NETA (n=6), R+delE2/NETA (n=7), placebo (n=14)]. The summary results in this section are therefore based on the analysis in which data from Site 1152 is excluded. Note that, although there were minor changes in the magnitude of treatment effects, the overall conclusion of the efficacy results remains unchanged when the analysis is conducted with data from this site included.

3.2.4.2 Primary Efficacy Results

Recall that the primary efficacy endpoint is the proportion of responders defined as subjects with an MBL volume of < 80 mL and at least a 50% reduction from baseline MBL volume over the last 35 days of treatment at Week 24. This reviewer was able to replicate the Applicant's analyses of the primary efficacy endpoint. In both trials, the lower bounds of the 95% CIs for the treatment differences (R+E2/NETA-Placebo) are greater than 0%, leading to conclusions of superiority of R+E2/NETA over placebo (Table 8). For MVT-601-3001, the responder rate is 72.1% for R+E2/NETA and 16.8% for placebo resulting in a treatment difference (95% CI) of 55.3% (44.2%, 65.6%). For MVT-601-3002, the responder rate is 71.2% for R+E2/NETA and 14.7% for placebo resulting in a treatment difference (95% CI) of 56.5% (46.6%, 66.5%).

Table 8: Proportion of responders at Week 24 (mITT)

	Treatments		**Difference (95% CI)**
Study	**R+E2/NETA**	**Placebo**	
MVT-601-3001	88 /122 (72.1%)	19/113 (16.8%)	55.3% (44.2%, 65.6%)
MVT-601-3002	89 /125 (71.2%)	19/129 (14.7%)	56.5% (46.6%, 66.5%)

Source: Adapted from Table 6 of the Sponsor's response to information request sent on 02/08/2021 (MVT-301-601-3001); Table 7.2.1.1 of the Study reports (MVT-601-3002).

3.2.4.3 Sensitivity Analyses for the Primary Efficacy Endpoint

This section summarizes the Applicant's sensitivity analyses results and the supplemental analysis conducted by this reviewer. The summaries are provided for the comparison of R+E2/NETA (the to be marketed treatment regimen) against placebo only.

A. Applicant's Sensitivity Analyses

The Applicant conducted sensitivity analyses under different assumptions for missing data and incomplete feminine product collection. The detailed descriptions of the Applicant's sensitivity analyses are provided in Appendix B. The results from these analyses are consistent with the findings from the primary efficacy analyses (Table 9).

Table 9: Applicant's sensitivity analyses for the primary efficacy endpoint

	Study MVT-601-3001		
Methods	**R+E2/NETA N=122**	**Placebo N=113**	**Difference (95% CI)**
Analysis using MBL volume from feminine product collections containing unvalidated materials	87 (71.3%)	19 (16.8%)	54.5% (43.9%, 65.1%)
Analysis using MBL volume from feminine product collections deemed inadequate	91 (74.6%)	22 (19.5%)	55.1% (44.5%, 65.7%)
Analysis treating early discontinuation as non-responders	87 (71.3%)	20 (17.7%)	53.0% (42.9%, 64.0 %)
Analysis using 24-week completers	78/94 (83.0%)	18/92 (19.6%)	63.4% (52.3%, 74.5%)
Analysis using per-protocol population	85/114 (74.6%)	18/108 (16.7%)	57.9% (47.2%, 68.5%)
Analysis using multiple imputation for handling missing MBL	(78.8%)	(25.3%)	53.5% (40.8%, 65.1%)
	Study MVT-601-3002		
Methods	**R+E2/NETA N=125**	**Placebo N=129**	**Difference (95% CI)**
Analysis using MBL volume from feminine product collections containing unvalidated materials	90 (72.0%)	18 (13.9%)	58.0% (48.1%, 67.9%)
Analysis using MBL volume from feminine product collections deemed inadequate	91 (72.8%)	21 (16.3%)	56.5% (46.4%, 66.6%)
Analysis treating early discontinuation as non-responders	88 (70.4%)	21 (16.3%)	54.1% (43.8%, 64.3%)
Analysis using 24-week completers	82/102 (80.4%)	18/102 (17.6%)	62.7% (52.1%, 73.4%)
Analysis using per-protocol population	85/113 (75.2%)	18/122 (14.7%)	60.5% (53.3%, 70.6%)
Analysis using multiple imputation for handling missing MBL	(70.6%)	(16.9%)	53.7% (42.6%, 64.7%)

Source: Reviewer's analysis (Study MVT-601-3001); Adapted from Table 17 of the Study report (Study MVT-601-3002).

B. Reviewer's Supplemental Analyses

As discussed in the statistical analysis section, for each subject, the Applicant defined the responder status by taking the duration of treatment exposure, feminine product return rate and information availability from the e-diary into consideration. The breakdown of the number of responders/non-responders based on the criteria used is summarized in Table 10.

Table 10: Proportion of responders at Week 24 grouped by criteria used (mITT)

	Study MVT-601-3001		Study MVT-601-3002	
Responder Status	**R+E2/NETA N=122**	**Placebo N=113**	**R+E2/NETA N=125**	**Placebo N=129**
Responders	88 (72.1%)	19 (16.8%)	89 (71.2%)	19 (14.7%)
Confirmed amenorrhea	57 (46.7%)	7 (6.2%)	57 (45.6%)	5 (3.9%)
Confirmed spotting or negligible bleeding	3 (2.5%)	0 (0.0%)	2 (1.6%)	0 (0.0%)
Observed MBL volume < 80 mL and >= 50% reduction from baseline	15 (12.3%)	12 (10.6%)	17 (13.6%)	13 (10.1%)
Imputed MBL volume < 80 mL and >= 50% reduction from baseline	*13 (10.7%)*	*0 (0.0%)*	*13 (10.4%)*	*1 (0.8%)*

Non-Responders	34 (28.9%)	94 (83.2%)	36 (28.8%)	110 (85.3%)
Observed MBL volume >= 80 mL or < 50% reduction from baseline	15 (12.3%)	76 (67.3%)	19 (15.2%)	92 (71.3%)
Imputed MBL volume >= 80 mL or < 50% reduction from baseline, regardless of FPRR	*4 (3.3%)*	*6 (5.3%)*	*6 (4.8%)*	*12 (9.3%)*
Treatment duration < 4 weeks	7 (5.7%)	6 (5.3%)	5 (4.0%)	3 (2.3%)
Withdrew from study due to lack of efficacy or need of surgical intervention	4 (3.3%)	3 (2.7%)	2 (1.6%)	1 (0.8%)
Missed Visit <= Week 8	2 (1.6%)	0 (0.0%)	1 (0.8%)	1 (0.8%)
Feminine product not collected for reason other than amenorrhea and no MBL data Week 8 or later	2 (1.6%)	3 (2.7%)	3 (2.4%)	1 (0.8%)

Source: Reviewer's analysis. The *italicized entries* represent subjects whose responder status is determined based on imputed MBL data using a mixed effects model.

As can be seen, subjects in the R+E2/NETA arm were considered responders mainly because they had confirmed amenorrhea. It is also evident that the imputed data based on the mixed effects model appears to favor the R+E2/NETA arm: a higher number of subjects in the R+E2/NETA arm are set as responders based on the imputed data, while the imputed data resulted in higher number of non-responders in the placebo arm.

Note that, during the review of the SAP for two pivotal studies, the statistical reviewer noted that the proposed mixed effect model for imputing missing data might underestimate MBL volume for subjects with partial feminine product information. The following comment was communicated with the Applicant:

Please clarify how mixed effect model approach will be used to impute the responder status for the following scenario. You stated that for subjects who had "treatment exposure ≥4 weeks, FPRR <100%, MBL<80mL and >=50% reduction from baseline", they will be imputed for partial or complete missing MBL volume (see Page 37) and the responder status is based on imputed MBL volume on in Table 7 (page 39) in the SAP. However, the mixed effect model that is used to impute MBL volume in section 7.3.6 could yield an imputed MBL value smaller than the observed value for the subject at the same visit when the data is partially missing. For example: a patient who reported 5 days of bleeding and feminine product return rate (FPRR)=60% and observed MBL was 50ml and had >50% reduction from baseline. If the imputed MBL value is 40 ml based on the mixed effect model, then using this imputed value to determine responder status is not appropriate as specified in the SAP for this situation.

The Applicant disagreed with the statistical reviewer's statements that the imputation model will underestimate the MBL volume. They further stated that, even if this scenario occurs, it would only be in few subjects. To evaluate the effect of the imputation approach on the study results, this reviewer conducted the analysis of the primary efficacy endpoint by re-classifying the responder status determined based on the imputed data as follows:

- subjects in the R+E2/NETA arm who were set as "responders" based on imputed data will be considered as non-responders (13 subjects in each study).

- subjects in the placebo arm whose status was considered "non-responder" based on imputed data will be considered as responders (6 subjects in Study MVT-601-3001 and 12 subjects in Study MVT-601-3002).

Per this analysis, a little over 60% of subjects in the R+E2/NETA group were responders compared to around 24% in the placebo group; and the treatment difference is still favorable to the R+E2/NETA arm.

Table 11: Proportion of responders at Week 24 with imputed data treated as failure (mITT)

Study	Treatments		Difference (95% CI)
	R+E2/NETA	Placebo	
MVT-601-3001	75/122 (61.5%)	25/113 (22.1%)	39.4% (27.8%, 50.8%)
MVT-601-3002	76 /125 (60.8%)	31/129 (24.3%)	36.7% (25.5%, 48.1%)

Source: Reviewer's analysis.

In the protocol defined primary efficacy analysis, subjects who discontinued the study for lack of efficacy or those who required surgical intervention for uterine fibroids were consider as non-responders. In line with this logic, and under the assumption that inability to tolerate the treatment could be an indication of treatment failure, the reviewer conducted the analysis of the primary efficacy endpoint with subjects who discontinued the study due to adverse events (AE) treated as non-responders. Missing data due to other reasons are handled as in the primary efficacy analysis. In this analysis, 3 subjects in the R+E2/NETA arm (MVT-601-3001) and 1 subject in the placebo arm (MVT-601-3002) were set as non-responders because they discontinued the study due to AE. The results are consistent with the findings of the primary efficacy analysis (Table 12).

Table 12: Proportion of responders at Week 24 with AE treated as failure (mITT)

Study	Treatments		Difference (95% CI)
	R+E2/NETA	Placebo	
MVT-601-3001	85/122 (69.7%)	19/113 (16.8%)	52.9% (42.2%, 63.5%)
MVT-601-3002	89 /125 (71.5%)	18/129 (13.9%)	57.6% (47.8%, 67.5%)

Source: Reviewer's analysis.

3.2.4.4 Analysis of Key Secondary Efficacy Endpoints

This section provides a detailed analysis of the seven key secondary efficacy endpoints; including additional analyses conducted to address missing data issues. Unless stated otherwise, the summary tables presented in this section are prepared by the statistical reviewer and replicate the Applicant's results.

[redacted] (b) (4) endpoints in the drug labeling: Amenorrhea, percent change in MBL volume and hemoglobin.

A. Amenorrhea

Amenorrhea refers to the absence of menstrual periods. In this study, a subject is considered to have confirmed amenorrhea if one of the following criteria is met:

1. No feminine product returned due to reported amenorrhea for two consecutive visits

2. No feminine product returned due to other reasons coupled with data indicating infrequent non-cyclic bleeding/spotting

3. Feminine product collection with a negligible observed MBL volume coupled with data indicating infrequent non-cyclic bleeding/spotting.

The proportion of subjects with confirmed amenorrhea at the last 35 days of treatment was the first key secondary efficacy endpoint of the studies. The treatment comparison for this endpoint was conducted using the same CMH approach used for the primary efficacy endpoint. The Applicant's findings for this efficacy endpoint is presented in Table 13. The results are significantly favorable to the R+E2/NETA arm.

Table 13: Proportion of subjects with confirmed amenorrhea

Study	Treatments		Difference (95% Cl)
	R+E2/NETA	Placebo	
MVT-601-3001	61/122 (50.0%)	7/113 (6.2%)	43.8% (33.9%, 53.7%)
MVT-601-3002	63 /125 (50.4%)	4/129 (3.1%)	47.3% (38.0%, 56.6%)

Source: Table 6 of the Sponsor's response to information request sent on 02/08/2021.

Reviewer's remark: As described in the breakdown of the responder status, most subjects who were classified as responders had confirmed amenorrhea. Therefore, given this endpoint is a subcategory of the primary efficacy endpoint, the results are not surprising.

Reviewer's remark: As supporting evidence for the efficacy of the drug for this endpoint, the Applicant compared the treatment arms with respect to time to first amenorrhea and the proportion of subjects with sustained amenorrhea (defined as the maintenance of amenorrhea at every subsequent visit after the initial achievement of amenorrhea). The results from these analyses were consistently favorable to the R+E2/NETA arm (see Figure 9 and 10 of the Applicant's study reports).

B. Percent change from baseline in MBL volume

The percentage change from baseline in MBL volume at Week 24 was the second key secondary efficacy endpoint. The Applicant conducted the analysis of this endpoint using a repeated measures analysis. The results are statistically significantly in favor of the R+ E2/NETA arm.

Table 14: Percent change from baseline in MBL volume at Week 24

Study	Treatments		Difference (95% CI)
	R+E2/NETA	Placebo	
MVT-601-3001	-82.0 (5.09)	-19.1 (5.2)	-62.9 (-76.4, -49.4)
MVT-601-3002	-84.3 (5.45)	-15.1 (5.47)	-69.2 (-84.1, -54.3)

Source: Table 6 of the Sponsor's response to information request sent on 02/08/2021.

Additional analysis conducted by the statistical reviewer

Table 15 presents the reviewer's analyses conducted using the multiple imputation and trimmed mean approaches. The results from these analyses provided results that are consistent with the Applicant's finding.

Table 15: Sensitivity analysis of the percent change from baseline in MBL volume at Week 24

Methods	Study MVT-601-3001			Study MVT-601-3002		
	Relugolix +E2/NETA	**Placebo**	**Diff (95% CI)**	**Relugolix +E2/NETA**	**Placebo**	**Diff (95% CI)**
Multiple Imputation	-83.0 (7.22)	-21.0 (6.81)	-61.9 (-80.6, -43.3)	-82.7 (8.17)	-17.1 (7.81)	-65.6 (-86.9, -44.3)
Trimmed mean	-101.0 (3.61)	-51.2 (3.35)	-49.9 (-57.9, -41.9)	-98.0 (3.32)	-43 (3.28)	-55.1 (-63.4, -46.7)
Trimmed mean*	-101.0 (3.61)	-51.2 (3.35)	-49.9 (-56.2, -43.8)	-98.0 (3.32)	-43 (3.28)	-55.1 (-61.6, -49.0)

Source: Reviewer's analysis: *confidence interval based on bootstrap samples. For the trimmed mean approach, the analysis is conducted using an ANCOVA model adjusting for baseline MBL, BMI, and stratifying variables. P<0.0001 for all comparisons.

Reviewer's remark: Both the Applicant's analysis and the multiple imputation analysis are based on the missing at random (MAR) assumption and hence their values are comparable. The trimmed mean analysis assumes missing data as "bad outcomes" and hence is more in line with the missing not at random (MNAR) assumption. The magnitude of the treatment effects in each arm is slightly higher in the trimmed mean as this is based on the top 50% best performers in each arm; and because this is more pronounced in the placebo arm, a lower treatment difference is observed.

C. Change from baseline bleeding and pelvic discomfort score (BPD)

This endpoint is derived based on a patient reported outcome (PRO). The PRO has 29 questions, of which, 8 questions were intended to measure disease severity and the rest to measure quality of life. For each question, patients were asked to rate the level of distress they felt during the previous 3 months using a 5-unit scale (1=not at all, 2=a little bit, 3=somewhat, 4=great deal 5=a very great deal). The Applicant used a Factor Analysis to group seven of the eight questions into three Factors:

- Factor 1: Bleeding and pelvic discomfort (BPD) (Q 1, Q2, Q5)
    - Q1: Heavy bleeding during your menstrual period
    - Q2: Passing blood clots during your menstrual period
    - Q5: Feeling tightness or pressure in your pelvic area
- Factor 2: Fluctuation in Menstruation (Q3 and Q4).
    - Q3: Fluctuation in duration of menstruation
    - Q4: Fluctuation in length of monthly cycle
- Factor 3: Urinary Symptoms (Q6 and Q7)
    - Q6: Frequency urination in daytime
    - Q7: Frequent nighttime urination

[redacted] (b) (4) Per the Applicant, this scale consists of three items proximal to uterine fibroids that are experienced by most patients. This score is derived as follows:

- Take the sum of the reported raw scores for the three questions

- Convert the raw sum into a normalized score using the following formula:

  *Transformed Score = [(Actual raw score – 3)/ (12)] * 100*

For the raw score, the minimum and maximum possible values for a given subject are 3 and 15, respectively. For the transformed score, the possible values are between 0 and 100. For both scales, higher scores are indicative of greater symptom severity (high scores = bad). For each patient, the Applicant derived the transformed score at baseline and at Week 24 and compared the treatment arms with respect to the mean change from baseline in BPD score (Table 16).

Table 16: Mean change from baseline at Week 24 BPD score

Study	Treatments		Difference (95% CI)
	R+E2/NETA	Placebo	
MVT-601-3001	-44.1 (3.13)	-16.4 (3.24)	-27.7 (-35.7, -19.7)
MVT-601-3002	-51.7 (2.89)	-18.3 (2.93)	-33.4 (-41.2, -25.5)

Source: Table 6 of the Sponsor's response to information request sent on 02/08/2021.

Additional analysis conducted by the statistical reviewer

This reviewer conducted additional analysis of this endpoint using different approaches for missing data and adjusting for baseline. These analyses include trimmed mean (considering the top 50% performing subjects from both groups) and multiple imputation approaches (including placebo-controlled multiple imputation in which observed data from the placebo group is used to impute missing data for subjects in the treatment group). The results of these analyses are consistent with the Applicant's findings and are favorable to the R+E2/NETA arm (Table 17).

Table 17: Sensitivity analysis for the mean change from baseline at Week 24 BPD score

	MVT-601-3001			MVT-601-3002		
	R+E2/NETA	Placebo	Diff (95% CI)	R+E2/NETA	Placebo	Diff (95% CI)
Multiple imputation [1]	-46.7 (2.41)	-14.4 (2.42)	-32.3 (-38.5, -26.1)	-52.5 (2.59)	-19.1 (2.66)	-33.5 (-40.5, -26.5)
Multiple imputation [2]	-41.4 (2.76)	-14.7 (2.71)	-26.7 (-33.7, -19.7)	-46.5 (2.79)	-19.4 (2.75)	-27.2 (-34.2, -20.0)
Trimmed Mean (50%) [3]	-62.0 (2.77)	-32.5 (2.97)	-29.6 (-36.8, -22.4)	-65.5 (2.21)	-30.4 (2.19)	-35.2 (-41.0, -29.3)
Trimmed Mean (Adap) [4]	-44.1 (3.15)	-19.9 (3.38)	-24.3 (-32.5, -16.0)	-54.6 (2.56)	-54.6 (2.56)	-36.3 (-43.2, -29.5)

Source: Reviewer's analysis.

[1]: Multiple imputation conducted under the MAR assumption
[2]: Reference-based multiple imputation in which the multiple imputation for subjects with missing data in both arms was based on data from subjects in the placebo arm. This imputation assumes that subjects in the treatment arm who discontinue the study treatment will act like those in the placebo arm who have the same baseline characterizes. In the missing data taxonomy, this falls under the MNAR assumption.
[3]: Trimmed mean (50%): Comparison between the top 50% performs from the two treatment groups
[4]: Trimmed mean (Adap): Comparison between equal percentage of completers determined based on the percentage of completers in the treatment arm. Note that, because there was comparable percentage of completers in both arms, this analysis is comparable to the complete case analysis.

<u>Reviewer's remark</u>: The Agency has determined that sufficient evidence of content validity has not been provided by the Applicant for the UFS-QoL BPD subscale (b) (4) .

D. Uterine pain numerical rating scale (NRS) score

The NRS is a self-reported instrument assessing pain on an 11-point scale [0 (no pain), 1-3 (mild pain), 4-6 (moderate pain), and 7-10 (severe pain)]. The proportion of patients with a maximum NRS score ≤1 during the last 35 days of treatment among subjects who had a maximum NRS score ≥4 prior to treatment was one of the key secondary efficacy endpoints. In MVT-601-3001, the proportion of subjects who had a maximum NRS score ≥4 at baseline was 66.4% (81/122) in R+E2/NETA arm and 75.2% (85/113) in the placebo arm. The corresponding figures in MVT-601-3002 were 74.4% (93/125) and 73.6% (95/129), respectively (Table 18).

Per the SAP, to ensure a robust estimate of patient response, patients need to record at least 28 days (80% of the last 35 days of treatment) of pain scores in the e-Diary. Consequently, the Applicant defined the pain evaluable population as subjects with a maximum NRS score ≥4 prior to treatment and have at least 28 days of pain scores in the e-Diary. In MVT-601-3001, 55/122 (45.1%) and 61/113 (54.0%), subjects in the R+E2/NETA and placebo arms, respectively, were included in the pain evaluable population. In MVT-601-3002, 68/125 (54.4%) and 82/129 (63.6%) subjects in the R+E2/NETA and placebo arms, respectively, were included in the pain evaluable population (Table 18).

Table 18: Patient disposition for the NRS scale

	MVT-601-3001		MVT-601-3002	
	R+E2/NETA	**Placebo**	**R+E2/NETA**	**Placebo**
Randomized	122	113	125	129
NRS<4	41/122 (33.6%)	28/113 (24.8%)	32 (25.6%)	34 (26.3%)
NRS≥4[1]	81/122 (66.4%)	85/113 (75.2%)	93 (74.4%)	95 (73.6%)
Pain Evaluable [2]	55/122 (45.1%)	61/113 (54.0%)	68/125 (54.4%)	82/129 (63.6%)

Source: Reviewer's analysis: [1] Subjects with baseline NRS score of≥4; [2] Subjects who had NRS≥4 at baseline and had at least 28 days of e-dairy pain data of the last 35 days of treatment.

The Applicant's analysis based on the pain evaluable population in both trials provided statistically significant treatment differences in favor of the R+E2/NETA arm. Among the pain evaluable subjects in MVT-601-3001, 23/55 (41.8%) in the R+E2/NETA arm compared to only 7/61 (11.5%) in the placebo arm had a maximum NRS score ≤1 during the last 35 days of treatment [Difference (95% CI): 30% (15.1%, 45.6%)]. In MVT-601-3002, 32/68 (47.1%) pain evaluable subjects in the R+E2/NETA arm compared to 14/82 (17.0%) subjects in the placebo arm had a maximum NRS score ≤1 during the last 35 days of treatment [Difference (95% CI): 30% (15.6%, 44.4%; Table 19].

Table 19: Proportion of patients with a maximum NRS score ≤1 during the last 35 days of treatment (pain evaluable population)

Study		**Treatments**		**Difference (95% CI)**
		R+E2/NETA	**Placebo**	
MVT-601-3001	**MVT-601-3001**	23/55 (41.8%)	7/61 (11.5%)	30% (15.1%, 45.6%)
MVT-601-3002	**MVT-601-3002**	32/68 (47.1%)	14/82 (17.1%)	30% (15.6%, 44.4%)

Source: Reviewer's analysis (MVT-601-3001); Table 6 of the Sponsor's response to information request sent on 02/08/2021 (MVT-601-3002).

As stated above, the analysis for this endpoint was conducted based on the pain evaluable population which excluded subjects with less than 28 days of e-diary records. As sensitivity analysis, the Applicant conducted the analysis of this endpoint using data from all women who

had a maximum pain score ≥ 4 during the 35 days prior to randomization regardless of the number of days of pain scores recorded in the e-diary. The results from this analysis in both studies are also favorable to the R+E2/NETA arm (Table 20).

Table 20: Proportion of patients with a maximum NRS score ≤ 1 during the last 35 days of treatment (All subjects with pre-randomization maximum NRS score≥4)

Study	Treatments		% Difference (95% CI)
	R+E2/NETA	Placebo	
MVT-601-3001	31/81 (38.3%)	11/85 (12.9%)	25.3% (12.6%, 38.1%)
MVT-601-3002	34/93 (36.6%)	17/95 (17.9%)	18.7% (6.2%, 31.1%)

Source: Reviewer's analysis (MVT-601-3001); Table 7.2.8.1 of the Applicant's Study report (MVT-601-3002)

The Applicant stated that uterine fibroid pain mostly occurs during bleeding days. Consequently, women take pain medications more often during the bleeding days. To evaluate the effect of concomitant pain medication on the pain outcome, the Applicant conducted additional analyses of the NRS endpoint on what they referred to as the dysmenorrhea evaluable patients. This population consist of subjects in the pain evaluable population who had a maximum NRS score of ≥4 on menstrual bleeding days prior to baseline.

In this analysis, subjects who had increased use of pain medication after baseline were considered as non-responders. Based on this analysis, 32/52 (61.5%) subjects in R+E2/NETA arm compared to 9/60 (15.0%) subjects in the placebo arm achieved a maximum NRS score ≤1 at the last 35 days of treatment in MVT-601-3001. In MVT-601-3002, the corresponding numbers are 40/62 (64.5%) and 18/77 (23.4%) in the R+E2/NETA and placebo arms, respectively (Table 21). Note that, the Applicant did not provide what constitutes an increased use of pain medication.

Table 21: Proportion of patients with a maximum NRS score ≤ 1 during the last 35 days of treatment (Dysmenorrhea evaluable patient)

Study	Treatments		% Difference (95% CI)
	R+E2/NETA	Placebo	
MVT-601-3001	32/52 (61.5%)	9/60 (15.0%)	46.4% (30.5%, 62.5%)
MVT-601-3002	40/62 (64.5%)	18/77 (23.4%)	41.1% (25.9%, 56.3%)

Source: Reviewer's analysis. Subjects with increased use of pain medication were considered non-responders. If the patient did not have any bleeding days during the last 35 days of treatment, NRS score is set to 0.

Additional analysis conducted by the statistical reviewer

Per the concomitant medication dataset, the number of subjects in MVT-601-3001 who received at least one concomitant medication of any kind for uterine fibroid pain in the R+E2/NETA arm was 77/122 (63.1%) and was 75/113 (66.6%) in the placebo group. Similarly, in MVT-601-3002, 86/125 (68.8%) subjects in the R+E2/NETA arm and 87/129 (67.4%) subjects in the placebo group received at least one concomitant medication for uterine fibroid pain (Table 22). The most frequently used medication was ibuprofen.

Table 22: Summary of concomitant pain medication use for uterine fibroid pain

	MVT-601-3001		MVT-601-3002	
	R+E2/NETA	Placebo	R+E2/NETA	Placebo
Randomized	122	113	125	129
Uterine fibroid pain concomitant medication use				

Safety population [1]	77/122 (63.1%)	75/113 (66.6%)	86/125 (68.8%)	87/129 (67.4%)
NRS>=4 Score [2]	57/81 (70.2%)	61/85 (71.8%)	69/93 (74.2%)	72/95 (75.8%)
Pain Evaluable [3]	40/55 (81.8%)	45/61 (73.8%)	51/68 (75%)	66/82 (80.5%)

[1] % computed using number of subjects who received at least one study medication (safety population). [2] % computed using number of subjects who had NRS≥4 at baseline. [3] % computed using number of subjects in the pain evaluable population.

Reviewer's remark: In the advice letter sent to the Applicant on February 3, 2017, the following comment was included regarding the list of permitted medications, "subjects are allowed to take ibuprofen or other Nonsteroidal Anti-Inflammatory Drugs (NSAIDs) for fibroid-related pain and allowed to use acetaminophen for non-fibroid pain. Justify this distinction and describe how subjects could distinguish pelvic pain/menstrual pain that is or is not fibroid-related."

This reviewer conducted the analysis of the NRS endpoint by treating subjects who received at least one concomitant medication post-randomization for uterine fibroid pain as non-responders. This analysis is conducted based on the pain evaluable population as well as considering all subjects with a maximum NRS score ≥4 at baseline. Per these analyses, only a few subjects in both arms would qualify as responders (≤1 maximum NRS score at Week 24). However, the treatment difference is still numerically favorable to the R+E2/NETA arm in both trials; and nominally significant in one of the two trials (Study MVT-601-3001; Table 23 and Table 24).

Table 23: Proportion of patients with a maximum NRS score ≤ 1 during the last 35 days of treatment (Concomitant pain medication set as non-responder: Pain evaluable population)

Study	Treatments		% Difference (95% Cl)
	R+E2/NETA	Placebo	
MVT-601-3001	8/55 (14.5%)	0/61 (0.0%)	14.5% (5.2%, 23.9%)
MVT-601-3002	8/68 (11.7%)	4/82 (4.9%)	6.9% (-2.1%, 15.8%)

Source: Reviewer's analysis.

Table 24: Proportion of patients with a maximum NRS score ≤ 1 during the last 35 days of treatment (Concomitant medication set as non-responder: All subjects with maximum NRS≥4 at baseline)

Study	Treatments		% Difference (95% Cl)
	R+E2/NETA	Placebo	
MVT-601-3001	12/81 (14.8%)	1/85 (1.1%)	13.6% (5.6%, 21.7%)
MVT-601-3002	8/93 (8.6%)	6/95 (6.3%)	2.3% (-5.2%, 9.8%)

Source: Reviewer's analysis.

Note that, for this endpoint, a responder is defined as achieving a maximum NRS score ≤ 1 in the last 35 days of treatment having a maximum NRS score ≥4 at baseline. However, based on the definition of the NRS scale, absence of pain is represented by an NRS score=0. This reviewer summarized the proportion of subjects with an NRS score=0 in the last 35 days of treatment among all subjects with a baseline maximum NRS score ≥4 and based on the pain evaluable population. Both analyses provided numerically favorable results to the R+E2/NETA arm (Table 25 and Table 26).

Table 25: Proportion of patients with a maximum NRS score = 0 during the last 35 days of treatment (Pain evaluable population)

Study	Treatments		% Difference (95% Cl)
	R+E2/NETA	Placebo	
MVT-601-3001	15/55 (27.3%)	6/61 (9.8%)	17.4% (3.5%, 31.4%)
MVT-601-3002	21/68 (30.9%)	13/82 (15.8%)	15.3% (1.5%, 28.6%)

Source: Reviewer's analysis.

Table 26: Proportion of patients with a maximum NRS score =0 during the last 35 days of treatment (All subjects with maximum NRS≥4 at baseline)

Study	Treatments		% Difference (95% CI)
	R+E2/NETA	Placebo	
MVT-601-3001	22/81 (27.2%)	10/85 (11.8%)	15.4% (3.5%, 27.3%)
MVT-601-3002	22/93 (23.7%)	14/95 (14.7%)	8.9% (-2.3%, 20.1%)

Source: Reviewer's analysis.

Reviewer's remark: As discussed above, over 70% of all subjects who had a baseline NRS score≥4 have received at least one concomitant medications for uterine fibroid pain. Consequently, it might be difficult to solely attribute the observed improvement in NRS pain score to the investigative drug alone. (b) (4) [redacted]

E. Hemoglobin

The proportion of women with a hemoglobin ≤10.5 g/dL at baseline (hemoglobin evaluable) who achieved an increase of ≥2 g/dL from baseline at Week 24 was one of the key secondary efficacy endpoints. The results for this endpoint are summarized in Table 27.

Table 27: Proportion of women who achieved an increase of > 2 g/dL from baseline at Week 24

Study	R+E2/NETA	Placebo	% Difference (95% CI)
MVT-601-3001	15/30 (50.0%)	5/22 (22.7%)	27.3% (2.24%, 52.3%)
MVT-601-3002	19/31 (61.3%)	2/37 (5.4%)	55.9% (37.2%, 74.5%)

Source: Source: Reviewer's analysis (MVT-601-3001); Table 18 of the Applicant's Study reports (MVT-601-3002).

Reviewer's remark: This reviewer disagreed with the Applicant's original analysis of this endpoint which excluded subjects who did not have hemoglobin data at Week 24 (see reviewer's additional analysis below). After discussions, the Applicant and the Agency have agreed to use the Applicant's analysis results conducted based on a mixed-effects model for labeling purposes. The results based on the mixed effects model are presented in Table 28.

Table 28: Proportion of women who achieved an increase of > 2 g/dL from baseline at Week 24

Study	Treatments		% Difference (95% CI)
	R+E2/NETA	Placebo	
MVT-601-3001	19/43 (44.2%)	5/29 (17.2%)	27.0% (6.7%, 47.2%)
MVT-601-3002	22/40 (55.0%)	3/53 (5.7%)	49.3% (32.7%, 66.0%)

Source: Table 6 of the Sponsor's response to information request sent on 02/08/2021

Additional analysis conducted by the statistical reviewer

The Applicant's analysis for this endpoint excluded subjects who did not have hemoglobin data at Week 24. As can be seen in Table 29, subjects did not have data because they discontinued the study early for a variety of reasons. Because missing data and noncompliance are post-randomization intercurrent events, analysis with only observed data may introduce selection bias. The reviewer conducted the analyses of this endpoint with subjects who did not have data

at Week 24 considered as treatment failures. The treatment difference is still numerically favorable to R+E2/NETA. However, the treatment difference is not statistically significant in Study MVT-601-3001.

Table 29: Patient disposition for the evaluation of hemoglobin

	MVT-601-3001		MVT-601-3002	
	R+E2/NETA	**Placebo**	**R+E2/NETA**	**Placebo**
Randomized	122	113	126	129
Evaluable [1]	43/122 (35.2%)	29/113 (25.7%)	40/126 (32%)	53/129 (41.1%)
Completed the Week 24	30/43 (69.7%)	22/29 (75.9%)	31/40 (77.5%)	37/53 (69.8%)
Discontinued the Study[2]	13/43 (30.2%)	7/30 (23.3%)	9/40 (22.5%)	16/53 (30.2%)
Reasons				
Adverse Event(s)	0/43 (0.0%)	1/29 (3.4%)	1/40 (2.5%)	2/53 (3.8%)
Lack of Efficacy	3/43 (6.9%)	0/29 (0.0%)	2/40 (5.0%)	0/53 (0.0%)
Lost to Follow-Up	0/43 (0.0%)	1/29 (3.4%)	0/40 (0.0%)	1/53 (1.9%)
Withdrawal by Patient	5/43 (11.6%)	2/29 (6.9%)	3/40 (7.5%)	3/53 (5.7%)
Protocol Deviation	1/43 (2.3%)	0/29 (0.0%)	0/40 (0.0%)	0/53 (0.0%)
Other	2/43 (4.6%)	1/29 (3.4%)	1/40 (2.5%)	4/53 (7.5%)
No reason	2/43 (4.6%)	2/29 (6.9%)	2/40 (5.0%)	6/53 (11.3%)

Source: Reviewer's analysis.1 Subjects with baseline hemoglobin levels of <10.5g/dL; 2 Subjects with no Week 24 hemoglobin data either because they discontinued the study early or did not provided data and hence were excluded from the analysis.

Table 30: Percentage of hemoglobin (All missing data set as failure)

Study	**Treatments**		**Difference (95% CI)**
	R+E2/NETA	**Placebo**	
MVT-601-3001	15/43 (34.9%)	5/29 (17.2%)	17.6% (-2.2%, 37.4%)
MVT-601-3002	19/40 (47.5%)	2/53 (3.8%)	43.7% (27.4%, 60.0%)

Source: Reviewer's Analysis.

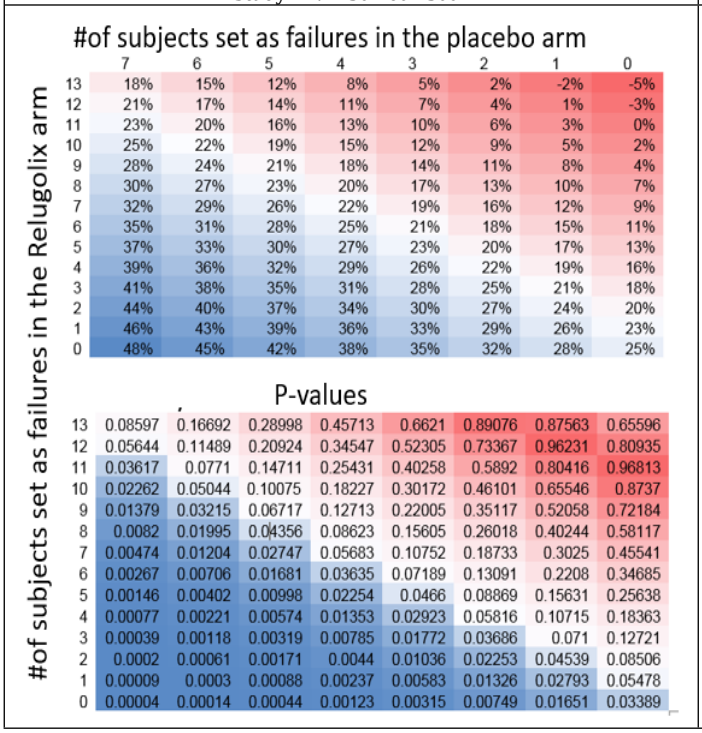
To further evaluate the impact of missing data, this reviewer conducted a tipping point analysis. The heatmaps provided in Figure 3 depict the treatment differences and p-values under different combination of number of subjects who did not have data at Week 24 treated as treatment failures. For example, 18% (13,7) for Study MVT-601-3001 corresponds to the treatment difference based on the analysis in which all 13 subjects in the R+E2/NETA arm and all 7 subjects in the placebo arm who did not have Week 24 data set as treatment failures. The p-value corresponding to this analysis is 0.0859.

As can be seen, in Study MVT-601-3001, except for the extreme cases in which almost all 13 subjects in the R+E2/NETA arm and none or only one of the 7 subjects in the placebo group who did not have Week 24 data were set as failures, the treatment differences are numerically favorable to the R+E2/NETA arm. However, the differences are not statistically significant.

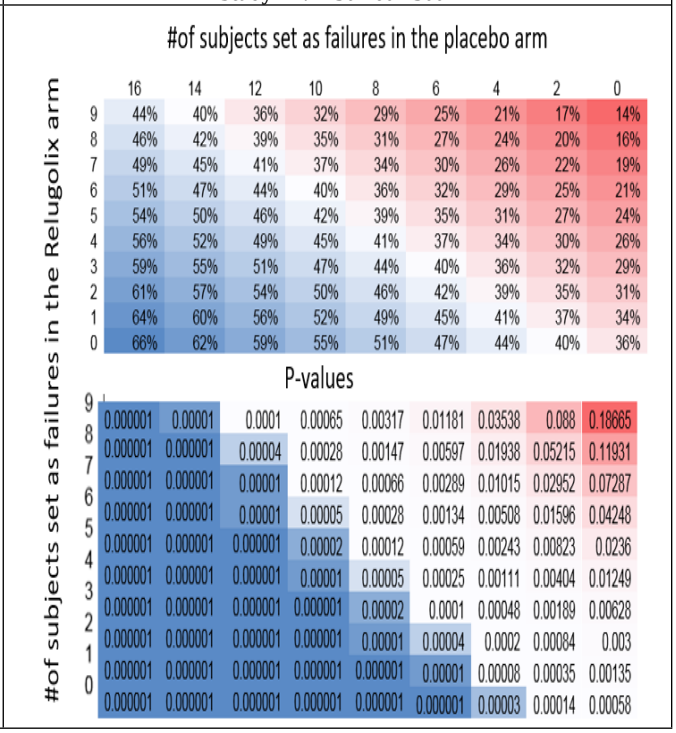
In Study MVT-601-3002, the results are consistently favorable to the R+E2/NETA arm, and for most cases, statistically significant differences are observed. Based on these observations, it appears that the results in Study MVT-601-3002 are only minimally affected by the missing data while in Study MVT-601-3001, the effect of missing data on the observed treatment difference is slightly higher.

Figure 3: Heat map for tipping analysis summary

Study MVT-301-601-3001



Study MVT-301-601-3002



Source: Reviewer's Analysis.

F. Uterine fibroid volume

The percent change from baseline to Week 24 in primary uterine fibroid volume was the sixth key secondary efficacy endpoint. In both studies, the treatment comparisons for this endpoint showed numerically favorable results for the R+E2/NETA group but the treatment differences were not statistically significant in either study (Table 31).

The mean percentage change from baseline to Week 24 in uterine fibroid volume was -12.6 in the R+E2/NETA arm compared with -0.01 in the placebo arm [Diff (95% CI): -12.6 (-27.7, 2.52)]. Similarly, In MVT-601-3002, the mean percent change from baseline to Week 24 in uterine fibroid volume was -17.4 in the R+E2/NETA arm compared with -7.4 in the placebo arm [Diff (95% CI): -10.0 (-25.8, 5.8)].

Table 31: Percent change from baseline to Week 24 in primary uterine fibroid volume

Study	Treatments		Difference (95% CI)
	R+E2/NETA	Placebo	
MVT-601-3001	-12.6 (5.95)	-0.01 (6.19)	-12.6 (-27.7, 2.52)
MVT-601-3002	-17.4 (5.9)	-7.4 (5.92)	-10.0 (-25.8, 5.8)

Source: Table 6 of the Sponsor's response to information request sent on 02/08/2021

G. Uterine Volume

The percent change from baseline to Week 24 in uterine volume was the final key secondary efficacy endpoint. The treatment comparisons for this endpoint in both trials showed statistically significant treatment difference in favor of the R+E2/NETA arm (Table 32).

In MVT-601-3001, the mean percentage change from baseline to Week 24 in uterine volume was -12.9 in the R+E2/NETA arm compared with 2.2 in the placebo arm [diff (95% CI): -15.1 (-23.3, -6.9)]. Similarly, In MVT-601-3002, the mean percent change from baseline to Week 24 in uterine volume was -13.8 in the R+E2/NETA arm compared with -1.5 in the placebo arm [Diff (95% CI): -12.2 (-21.3, -3.2)].

Table 32: Percent change from baseline to Week 24 in uterine fibroid volume

Study	Treatments		Difference (95% CI)
	R+E2/NETA	Placebo	
MVT-601-3001	-12.9 (3.21)	2.2 (3.37)	-15.1 (-23.3, -6.9)
MVT-601-3002	-13.8 (3.4)	-1.5 (3.37)	-12.2 (-21.3, -3.2)

Source: Table 6 of the Sponsor's response to information request sent on 02/08/2021.

Reviewer's remark: (b) (4)

3.3 Evaluation of Safety

This section presents descriptive summaries of the percentages of treatment-emergent adverse events from Study MVT-601-3001 and Study MVT-601-3002. These summaries are provided for the safety analysis population, which is defined in the SAP as all randomized patients who receive at least 1 dose of study medication. Unlike for the efficacy summary, the Agency has decided to include all data including data from Site 1152 in the safety summaries.

3.3.1 Adverse Events

The percentage of subjects who reported at least one adverse event of any type was 61.0% in the R+E2/NETA group, 72.1% in the R+delE2/NETA group, compared to 62.5% in the placebo group (Table 33). The percentage of subjects who reported at least one serious adverse event was 3.1% in the R+E2/NETA group, 1.9% in the R+delE2/NETA group, and 2.3% in the placebo group. No death was reported in either study. Adverse events leading to study discontinuation were reported for 3.9% (10/254) subjects in the R+E2/NETA group, 11.6% (30/258) subjects in the R+delE2/NETA group, and 4.3% (11/256) subjects in the placebo group (Table 33).

The most frequently reported adverse events in the R+E2/NETA group were headache (9.8%), hot flash (8.3%), and hypertension (4.7%). The corresponding figures in the placebo group were 13.3% [headache], 5.9% [hot flash] and 1.6% [hypertension]. The incidence rate of these events in the R+delE2/NETA group were (16.3% [headache], 35.3% [hot flash] and 3.9% [hypertension]; Table 34).

Table 33: Overall summary of adverse events (MVT-601-3001, MVT-601-3002)

	R+E2NETA N=254	R+delE2NETA N=258	Placebo N=256
Any AE	155 (61.0%)	186 (72.1%)	160 (62.5%)
AE leading to treatment discontinuation	10 (3.9%)	30 (11.6%)	11 (4.3%)
AE leading to treatment interruption	3 (1.2%)	3 (1.2%)	4 (1.6%)
Grade 2 or Above	92 (36.2%)	144 (55.8%)	66 (25.8%)
Serious AE	8 (3.1%)	5 (1.9%)	6 (2.3%)

Source: Adapted from Table 4.1.1 of the integrated summary of safety (ISS).

Table 34: Summary of adverse events reported in at least 2% of subjects (MVT-601-3001, MVT-601-3002)

Preferred Terms	R+E2NETA N=254	R+delE2NETA N=258	Placebo N=256
Headache	25 (9.8%)	42 (16.3%)	34 (13.3%)
Hot flush	21 (8.3%)	91 (35.3%)	15 (5.9%)
Hypertension	12 (4.7%)	10 (3.9%)	4 (1.6%)
Nausea	10 (3.9%)	9 (3.5%)	16 (6.3%)
Alopecia	9 (3.5%)	9 (3.5%)	2 (0.8%)
Abdominal pain	9 (3.5%)	5 (1.9%)	4 (1.6%)
Upper respiratory tract infection	7 (2.8%)	10 (3.9%)	10 (3.9%)
Back pain	7 (2.8%)	10 (3.9%)	9 (3.5%)
Menorrhagia	7 (2.8%)	6 (2.3%)	1 (0.4%)
Libido decreased	7 (2.8%)	5 (1.9%)	1 (0.4%)
Nasopharyngitis	6 (2.4%)	5 (1.9%)	11 (4.3%)
Anaemia	6 (2.4%)	2 (0.8%)	14 (5.5%)
Irritability	6 (2.4%)	1 (0.4%)	0 (0.0%)
Bronchitis	6 (2.4%)	0 (0.0%)	4 (1.6%)
Arthralgia	5 (2.0%)	15 (5.8%)	8 (3.1%)
Fatigue	5 (2.0%)	13 (5.0%)	7 (2.7%)
Metrorrhagia	5 (2.0%)	6 (2.3%)	1 (0.4%)
Hyperhidrosis	5 (2.0%)	5 (1.9%)	2 (0.8%)
Pelvic pain	5 (2.0%)	3 (1.2%)	6 (2.3%)
Insomnia	5 (2.0%)	3 (1.2%)	5 (2.0%)
Dyspepsia	5 (2.0%)	2 (0.8%)	1 (0.4%)
Breast cyst	5 (2.0%)	0 (0.0%)	0 (0.0%)

Source: Table 4.3.1 of the ISS.

3.3.1.1 Adverse events of special interest

Bone Mineral Density (BMD)

Per the Applicant, low-estrogen levels can weaken bone health and cause unwanted side effects such as hot flashes. Low doses of estradiol and progestins help offset these effects. To this effect, the Applicant is investigating whether their combination product maintains estrogen in

the low normal range to balance the potential benefits of Relugolix, while also maintaining bone health and mitigating the side effects that can result from a low-estrogen state.

Per Section 9.3.5 of the SAP for the two pivotal studies and Section 7.4.3.4 of the SAP for integrated summary of safety (ISS), to support the inclusion of E2/NETA in the treatment regimen, the safety endpoint of mean percent change from baseline in BMD at the lumbar spine at Week 12 will be analyzed using pooled data from the two replicate studies (MVT-601-3001 and MVT-601-3002). A formal superiority test comparing R+E2/NETA with R+delE2/NETA will be performed using a mixed effects model. The model includes treatment group, age at baseline, visit, baseline BMD value, stratification factors (geographic region and menstrual blood loss volume), race (African American versus Other), and BMI at baseline as fixed effects. An unstructured variance covariance matrix was used to account for within subject correlation. From this model, least square means with associated 95% CIs will be provided. The summaries of the percent changes from baseline to Weeks 12 and 24 in BMD at the lumbar spine for each study separately and the two studies combined are summarized in Table 35.

Table 35: Mean percentage change from baseline in BMD at lumbar spine (Safety population)

	Treatments			**Difference (95% Confidence Interval)**		
Visit	**R +E2/NETA**	**R + del E2/NETA**	**Placebo**	**R+E2/NETA Vs Placebo**	**R+ delE2/NETA Vs Placebo**	**R+E2/NETA Vs R+ delE2/NETA**
	Study MVT-301-601-3001					
Baseline	N=128 1.186 (0.016)	N=132 1.22 (0.016)	N=127 1.242 (0.016)			
12	N=101 -0.47 (0.291)	N=103 -1.99 (0.285)	N=103 0.198 (0.286)	-0.667 (-1.40, 0.067)	-2.193 (-2.91, -1.468)	1.52 (0.79, 2.26)
24	N=100 -0.356 (0.293)	N=100 -1.82 (0.289)	N=102 0.052 (0.289)	-0.407 (-1.15, 0.341)	-1.869 (-2.61, -1.127)	1.46 (0.711, 2.21)
	Study MVT-301-601-3002					
Baseline	N=126 1.227 (0.016)	N=126 1.217 (0.016)	N=129 1.235 (0.015)			
12	N=103 -0.819 (0.269)	N=95 -1.919 (0.277)	N=104 0.508 (0.265)	-1.326 (-2.03, -0.6284)	-2.243 (-3.142, -1.711)	1.10 (0.381, 1.819)
24	N=95 -0.126 (0.297)	N=94 -2.124 (0.299)	N=95 0.315 (0.291)	-0.441 (-1.225, 0.342)	-2.439 (-3.226, -1.1.652)	1.998 (1.206, 2.789)
	Studies MVT-301-601-3001 and MVT-301-601-3002 Combined					
12	N=204 -0.626 (0.198)	N=198 -1.961 (0.198)	N=207 0.342 (0.194)	-0.968 (-1.476, -0.460)	-2.304 (-2.813, -1.794)	1.335 (0.820, 1.850)
24	N=195 -0.233 (0.207)	N=194 -1.972 (0.207)	N=197 0.184 (0.203)	-0.417 (-0.955, 0.121)	-2.156 (-2. 694, -1.619)	1.739 (1.197, 2.281)

Source: Table 32 (MVT-601-3001) and Table 29 (MVT-601-3002) of the Applicant's Study Reports and Table 8.5.11of ISS

Categorical summaries of percent change from baseline at Weeks 12 and 24 for each study separately is provided in Table 37. Per this summary, the proportion of subjects who showed no change or had an increase in BMD were comparable between the R+E2/NETA arm and the placebo arm. The R+delE2/NETA arm had the lowest percentage of subjects with no change or increase in BMD. In both studies, the percentage of missing BMD data was comparable across the three treatment groups (around 18-26%).

Reviewer's remark: Based on the results summarized above, (b) (4)

(b) (4)

Reviewer's remark: Note, per the SAP, the difference of percent change from baseline between treatment groups (b) (4) *will be summarized* (b) (4) *with the corresponding 95% CIs. However, there will be no p-values provided for these comparisons.* (b) (4)

4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

The subgroup analyses results presented in Figure 8 and Figure 9 provided treatment differences in favor of R+E2/NETA across all categories including race, age, and BMI. In both trials, the proportion of responders was numerically higher in subjects with baseline NRS pain score <4. Subjects with baseline MBL volume <225 mL, subjects with uterine volume <300 cm^3 (MVT-600-3001) and subjects enrolled in sites located outside the North American region (MVT-600-3002) also had numerically higher responders compared to the other subgroups.

Reviewer's remark: For each subgroup, the Applicant presented the proportion of responders for each treatment arm. However, the treatment comparisons were done based on odds ratios as opposed to the difference in the proportion of responders. The reviewer provided the difference in proportions and the associated 95% CI for each subgroup treatment comparison.

5 SUMMARY AND CONCLUSIONS

5.1 Statistical Issues

No major statistical issues related to the primary efficacy analysis were encountered in this review. One issue related to the analysis population used for the hemoglobin endpoint was addressed. The Applicant analyzed this endpoint by excluding eligible subjects based on post-randomization intercurrent events. Per this reviewer, doing so might introduce selection bias in the treatment comparison. Therefore, this reviewer recommended the analyses results based on all eligible subjects to be used in the drug labeling. The clinical team found the Applicant's analysis in which missing data were handled via a mixed effects model appropriate for this endpoint. Note, this analysis assumes missing data to follow the missing at random mechanism and measures the hypothetical estimand (treatment difference in the hypothetical scenario where treatment discontinuations did not occur).

5.2 Collective Evidence

The primary efficacy analysis results in both trials support a conclusion of superiority of R+E2/NETA over placebo. The secondary and sensitivity analyses all show similar robust findings and demonstrated the efficacy of R+E2/NETA in patients with heavy menstrual bleeding associated with uterine fibroids. Regarding safety, the most frequently reported adverse events in the R+E2/NETA group were headache, hot flash, and hypertension.

Reviewer's remarks: Due to the observed bone mineral density loss, the duration of use of the study might be limited to two years.

5.3 Labeling Recommendations

This reviewer provided the following recommendations and comments for the drug labeling. The comments were accepted by the team.

- We recommend confidence intervals to be provided for each endpoint (b) (4)
- (b) (4)
- (b) (4) bone mineral density is a safety endpoint, (b) (4)

Section 14 of the current label is included below

The efficacy and safety of MYFEMBREE were evaluated in two replicate, 24-week, multinational, randomized, double-blind, placebo-controlled studies in a total of 768 premenopausal women with heavy menstrual bleeding associated with uterine fibroids in Study L1 (NCT03049735) and Study L2 (NCT03103087). For study inclusion, women had to have uterine fibroids confirmed by ultrasound examination in which at least one fibroid met at least one of the following criteria:

a. Subserosal, intramural, or < 50% intracavitary submucosal fibroid with a diameter ≥ 2 cm, or
b. Multiple small fibroids with a total uterine volume of ≥ 130cm3.

Women also had to have menstrual blood loss (MBL) volume of ≥ 80 mL per cycle for two menstrual cycles or ≥ 160 mL during one cycle quantified by the alkaline hematin method from menstrual products collected during baseline menstrual cycles to be included in the studies. Women with hemoglobin < 8.0 g/dL were excluded from the study. Iron therapy was required

for women with hemoglobin ≥ 8 g/dL and ≤ 10 g/dL. Women were allowed, but not required, to take calcium and vitamin D during the study.

In Studies L1 and L2, women were randomized 1:1:1 to receive a once daily relugolix 40 mg tablet plus an over encapsulated tablet of E2 1 mg and NETA 0.5 mg (relugolix+E2/NETA), which is equivalent to 1 tablet of MYFEMBREE, for 24 weeks, placebo for 24 weeks, or relugolix 40 mg monotherapy for 12 weeks followed by MYFEMBREE for 12 weeks. Treatment was initiated within the first seven days after the onset of menses.

The primary endpoint was the proportion of women in the MYFEMBREE group compared with women in the placebo group, who achieved menstrual blood loss volume of < 80 mL and at least a 50% reduction from baseline MBL volume over the last 35 days of treatment, as measured by the alkaline hematin method. Key secondary endpoints were related to amenorrhea, MBL volume, and change in hemoglobin.

A total of 768 women were randomized and treated in Studies L1 and L2 (741 women in the efficacy population used for these studies). Of the 741 women, 247 received treatment with MYFEMBREE (122 and 125 in Studies L1 and L2, respectively), 252 received treatment with relugolix followed by MYFEMBREE (125 and 127 in Studies L1 and L2, respectively), and 242 received placebo (113 and 129 in Studies L1 and L2, respectively). The median age of women included in the efficacy analysis was 42 years (19 - 51 years), and mean body mass index was 31.6 (b) (4) kg/m2. Approximately 53% of study participants were Black, 41% were White, and 6% were of other races. Across studies at baseline, mean (± standard deviation) MBL volume at baseline was 231 mL (± 156). Baseline uterine size in Studies L1 and L2 ranged from normal to greater than 28 weeks gestation size (47 - 2625 cm3).

Menstrual Blood Loss

In both Study L1 and Study L2, a statistically higher proportion of women treated with MYFEMBREE achieved the primary endpoint of both an MBL volume of less than 80 mL and at least a 50% reduction from baseline in MBL volume over the last 35 days of treatment compared with placebo (Table A1).

Table A1: Proportion of Responders for Reduction in MBL Volume Over Last 35 days of Treatment in Women with Uterine Fibroids (Studies L1 and L2)

	Study L1		**Study L2**	
	MYFEMBREE (N = 122)	**Placebo (N = 113)**	**MYFEMBREE (N = 125)**	**Placebo (N = 129)**
Women with MBL Volume < 80 mL and ≥50% Reduction in MBL Volume from Baseline to the Last 35 Days of Treatment	**72.1%**	**16.8%**	**71.2%**	**14.7%**
%Difference from placebo 95% CI* p-value	**55.3% (44.2%, 65.6%) < 0.0001**		**56.5% (46.6%, 66.5%) < 0.0001**	

***CI = confidence interval.**

Amenorrhea

In Studies L1 and L2, 50.0% and 50.4% of women treated with MYFEMBREE, respectively, achieved amenorrhea compared to 6.2% and 3.1% treated with placebo, respectively, over the last 35 days of treatment.

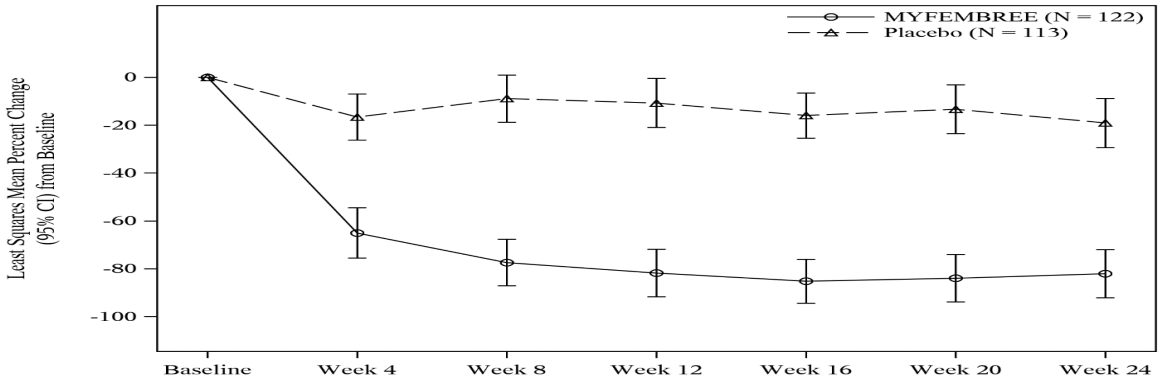
Percent Change in MBL Volume

The mean MBL volumes in Studies L1 and L2 at baseline were 243.8 mL and 246.7 mL in the MYFEMBREE group and 223.2 mL and 211.8 mL in the placebo group, respectively. The mean reduction in MBL volume from baseline to Week 24 in the MYFEMBREE group was 82.0% in Study L1 and 84.3% in Study L2, compared with placebo which was 19.1% and 15.1%, respectively.

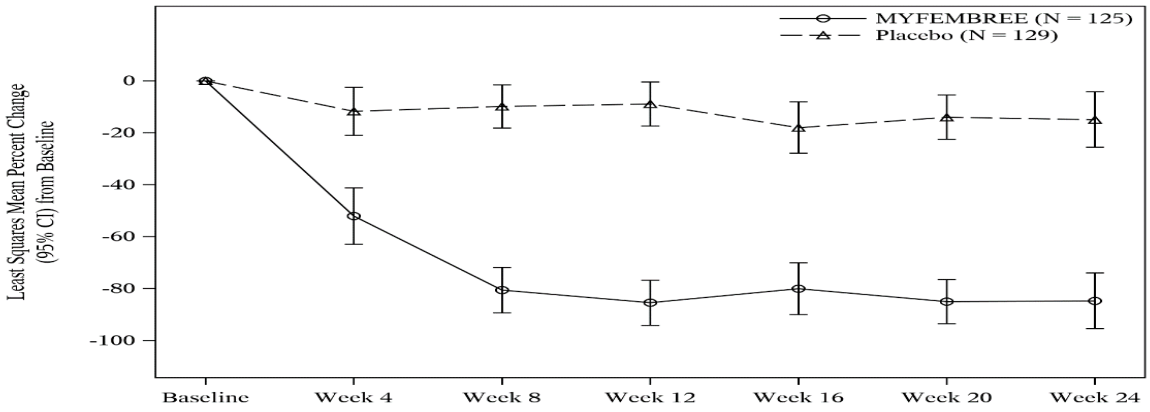
Reductions in MBL volume for MYFEMBREE and placebo groups are depicted in Figure A1.

Figure A1: Percent Change from Baseline in Menstrual Blood Loss over Time

Study L1



Study L2



Hemoglobin Levels

For efficacy, a hemoglobin response was defined as a hemoglobin increase > 2 g/dL from baseline to Week 24 in the subgroup of women with anemia at baseline (hemoglobin ≤ 10.5 g/dL). A statistically higher proportion treated with MYFEMBREE compared with placebo had > 2 g/dL improvement in hemoglobin levels, see Table A2.

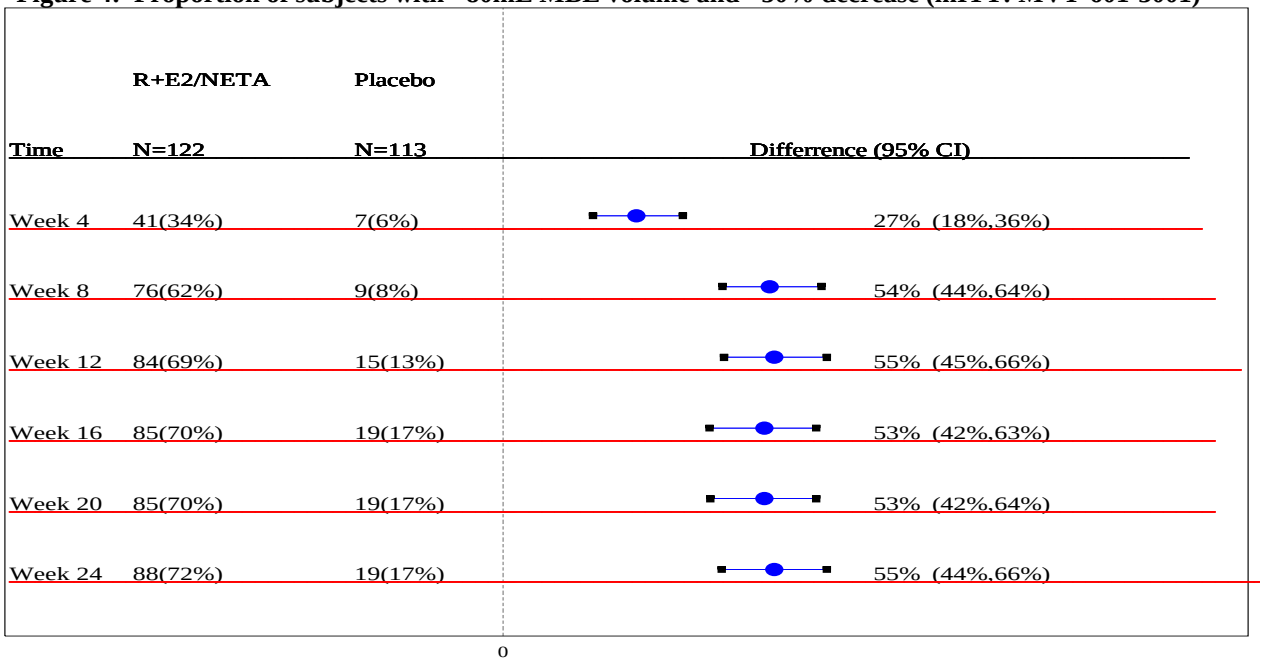
Table A2: Proportion of Women with Baseline Hgb ≤ 10.5 and > 2 g/dL Improvement in Hemoglobin Levels from Baseline at Week 24

	Study L1		**Study L2**	
	MYFEMBREE n=43 (N = 122)	**Placebo n=29 (N = 113)**	**MYFEMBREE n=40 (N = 125)**	**Placebo n=53 (N = 129)**
% at Week 24	**44.2%**	**17.2%**	**55.0%**	**5.7%**
%Difference from placebo 95% CI p-value	**26.9% (6.7%, 47.2%) 0.0177**		**49.3% (32.7%, 66.0%) <0.0001**	

CI = confidence interval
n = number of patients with Hgb ≤10.5 g/dL at baseline (b) (4)
N = number of patients in each treatment group

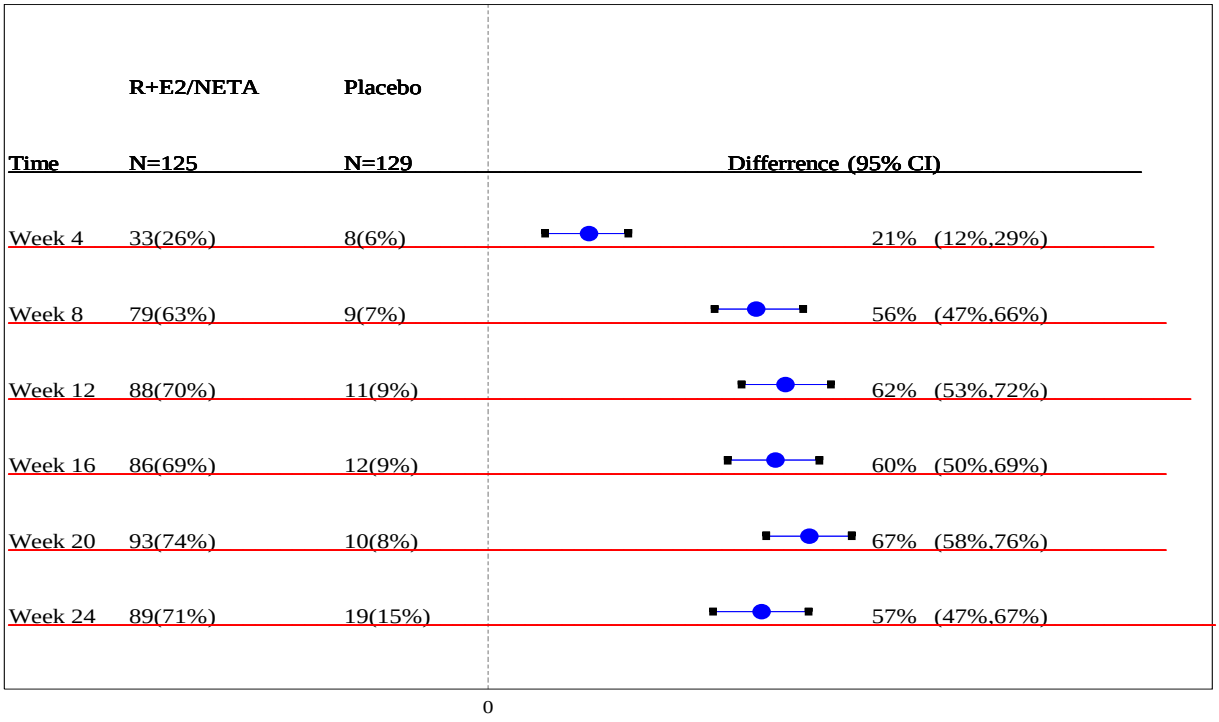
6 Appendix A: Selected Efficacy and Safety Summaries

Figure 4: Proportion of subjects with <80mL MBL volume and >50% decrease (mITT: MVT-601-3001)



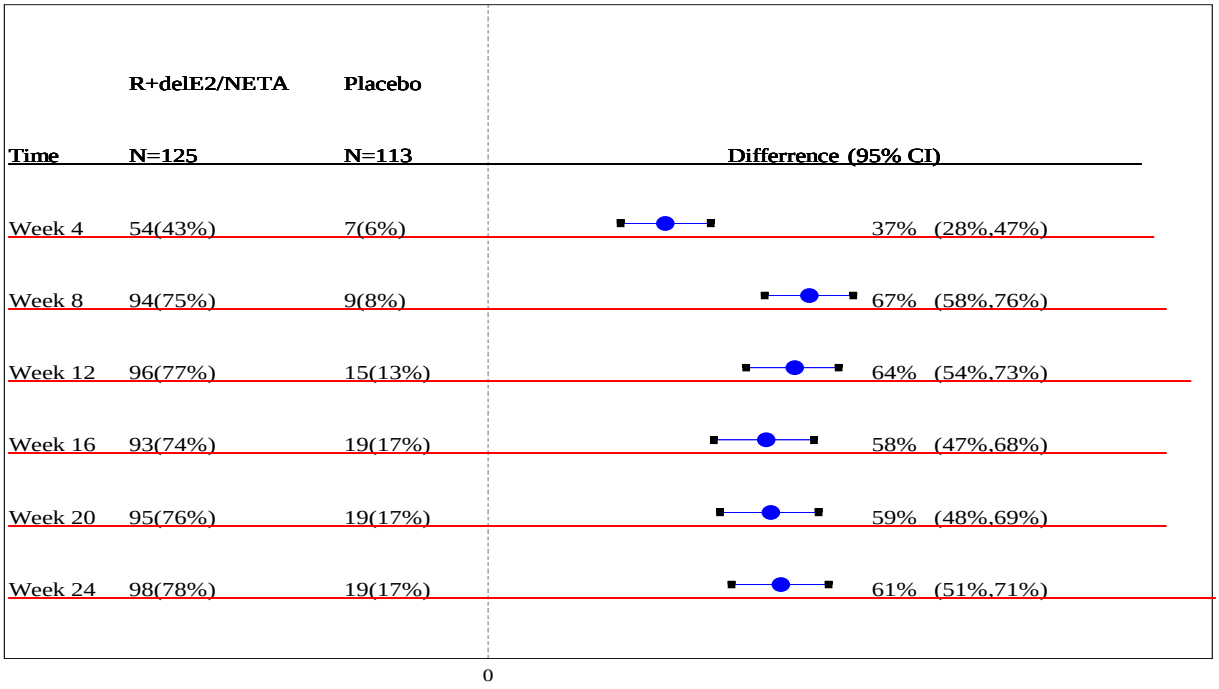
Source: Reviewer's analysis.

Figure 5: Proportion of subjects with <80mL MBL volume and >50% decrease (mITT: MVT-601-3002)



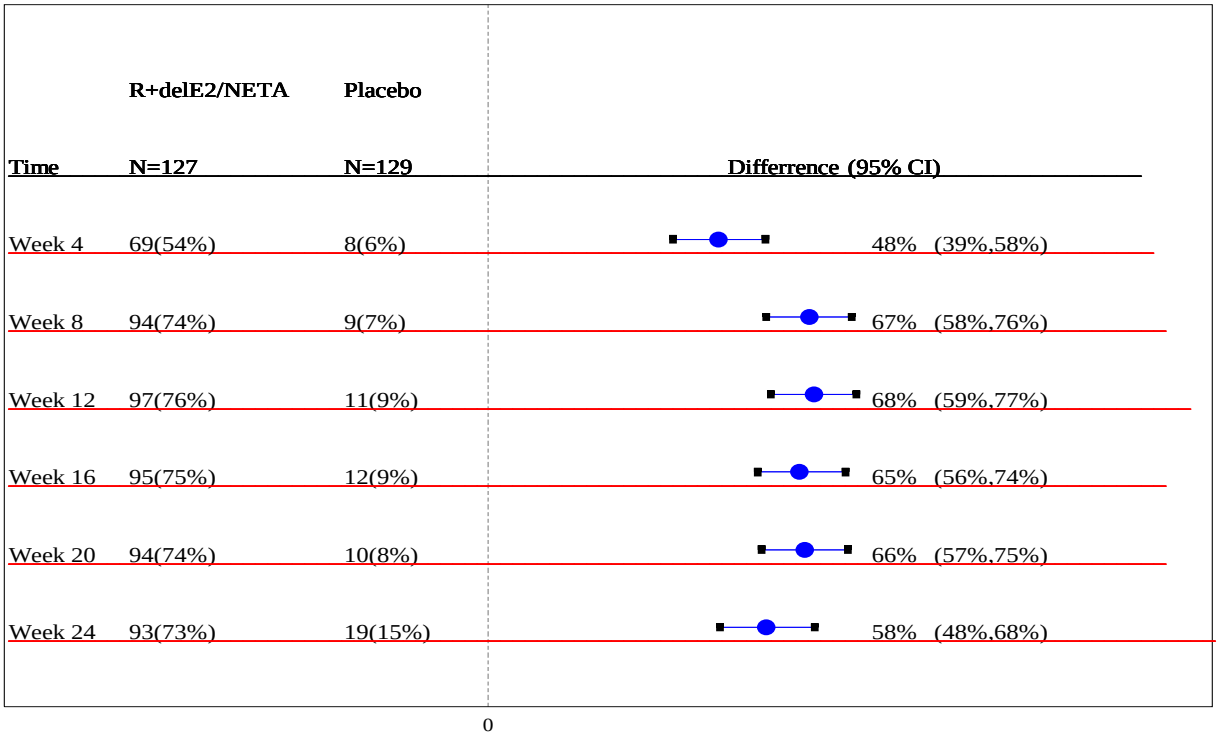
Source: Reviewer's analysis.

Figure 6: Proportion of subjects with <80mL MBL volume and >50% decrease: Delayed arm (mITT: MVT-601-3001)



Source: Reviewer's analysis.

Figure 7: Proportion of subjects with <80mL MBL volume and >50% decrease: Delayed arm (mITT: MVT-601-3002)



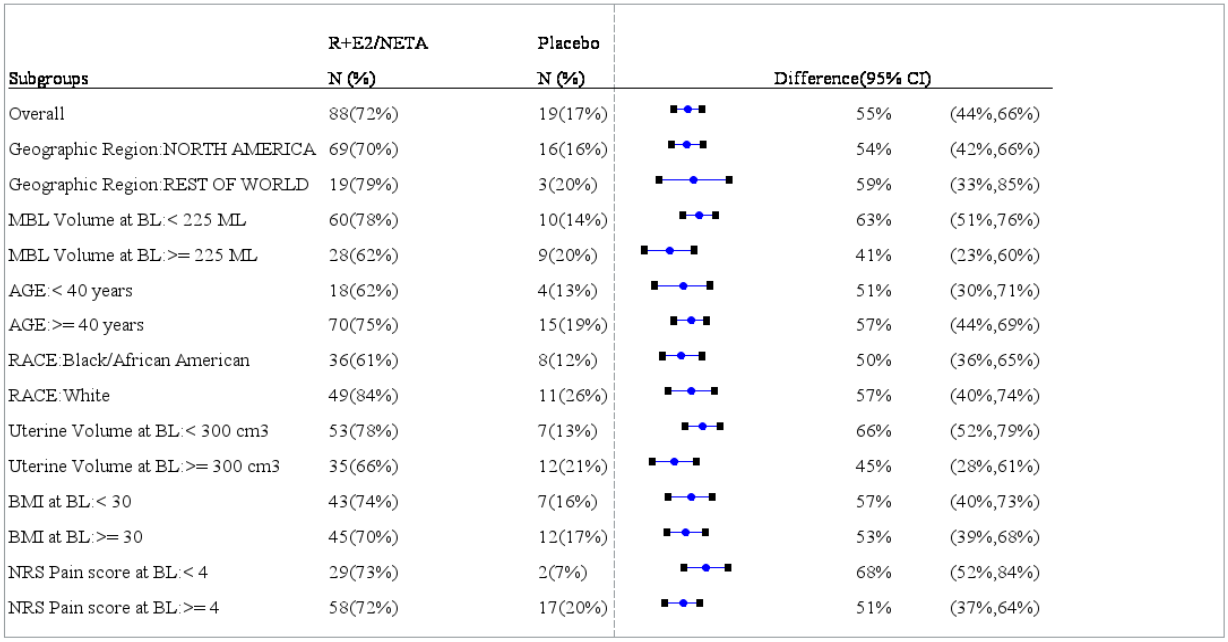
Source: Reviewer's analysis.

Table 36: Summary of key secondary efficacy endpoints (Delayed Arm)

Alpha -adjusted Endpoints	MVT-601-3001			MVT-601-3002		
	R + del E2/NETA	Placebo	Difference (95% Cl)	R + del E2/NETA	Placebo	Difference (95% CI)
Proportion of women who achieved amenorrhea over the last 35 days of treatment	70/125 (56.0%)	7/113 (6.2%)	49.8% (40.0%, 59.6%)	63/127 (49.6%)	4/129 (3.1%)	46.5% (37.3%, 55.7%)
Percent change from baseline to Week 24 in MBL volume [mean (SD)]	-86.4 (5.01)	-19.1 (5.2)	-67.3 (-80.7, -53.9)	-89.4 (5.74)	-15.1 (5.47)	-74.3 (-89.6, -59.0)
Change from baseline to Week 24 in UFS-QoL BPD scale score as measured by UFS-QoL (Q1, Q2, Q5) [mean (SD)]	-51.8 (3.11)	-16.4 (3.24)	-35.3 (-43.3, -27.3)	-48.9 (3.02)	-18.3 (2.93)	-30.6 (-38.6, -22.6)
Proportion of women with hemoglobin level ≤10.5g/dL at baseline who achieved an increase of >2 g/dL from baseline at Week 24	18/32 (56.3%)	5/22 (22.7%)	33.5% (9.0%, 58.1%)	18/31 (58.1%)	2/37 (5.4%)	52.7% (33.8%, 71.5%)
Proportion of women who achieved a maximum NRS score ≤1 for uterine fibroid-associated pain over the 35 days of treatment in the subset of women with a maximum pain score ≥4 during the 35 days prior to randomization	26/62 (41.9%)	7/61 (11.5%)	30.5% (15.8%, 45.1%)	24/58 (41.4%)	14/82 (17.1%)	24.3% (9.2%, 39.4%)
Percent change from baseline to Week 24 in primary uterine fibroid volume [mean (SD)]	-22.3 (5.98)	-0.01 (6.19)	-22.2 (-37.5, -7.0)	-30.2 (6.25)	-7.4 (5.92)	-22.8 (-38.9, -6.7)
Percent change from baseline to Week 24 in uterine volume [mean (SD)]	-18.1 (3.21)	2.2 (3.01)	-20.3 (-28.5, -12.1)	-17.7 (3.54)	-1.5 (3.37)	-16.2 (-25.4, -7.1)

Source: Reviewer's analysis (MVT-601-3001); Adapted from Table 7.2.3.2 (amenorrhea) 7.2.2.1 (MBL), 7.2.4.1 (hemoglobin), 7.2.5.9 (UFS-QoL), 7.2.8.5 (NRS), 7.2.9.1 (Uterine volume), 7.2.9.3 (uterine fibroid volume; MVT-601-3002).

Figure 8: Proportion of responders at Week 24: Subgroup analysis (mITT: MVT-601-3001)



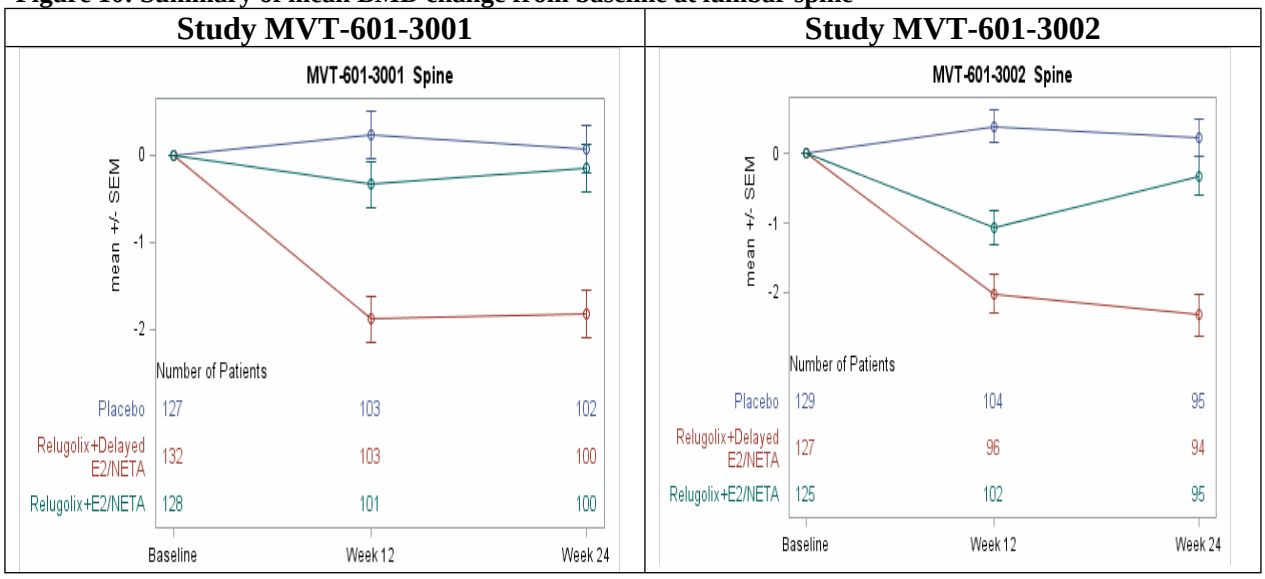
Source: Reviewer's analysis. BL: Baseline.

Figure 9: Proportion of responders at Week 24: Subgroup analysis (mITT: MVT-601-3002)

Subgroups	R+E2/NETA N (%)	Placebo N (%)	Difference(95% CI)		
Overall	89(71%)	19(15%)		57%	(47%,67%)
Geographic Region:NORTH AMERICA	63(68%)	16(17%)		51%	(39%,63%)
Geographic Region:REST OF WORLD	26(81%)	3(9%)		74%	(59%,89%)
MBL Volume at BL:< 225 ML	54(71%)	12(15%)		56%	(43%,69%)
MBL Volume at BL:>= 225 ML	35(71%)	7(14%)		58%	(42%,74%)
AGE:< 40 years	22(69%)	5(12%)		57%	(39%,74%)
AGE:>= 40 years	67(72%)	14(16%)		57%	(45%,69%)
RACE:Black/African American	41(65%)	11(15%)		50%	(36%,65%)
RACE:White	44(76%)	7(14%)		62%	(47%,77%)
Uterine Volume at BL:< 300 cm3	54(73%)	10(15%)		57%	(43%,71%)
Uterine Volume at BL:>= 300 cm3	35(69%)	9(14%)		57%	(42%,72%)
BMI at BL:< 30	40(69%)	7(11%)		59%	(45%,73%)
BMI at BL:>= 30	48(73%)	12(18%)		55%	(41%,69%)
NRS Pain score at BL:< 4	23(77%)	3(10%)		74%	(56%,92%)
NRS Pain score at BL:>= 4	65(70%)	16(17%)		53%	(41%,65%)

Source: Reviewer's analysis. BL: Baseline.

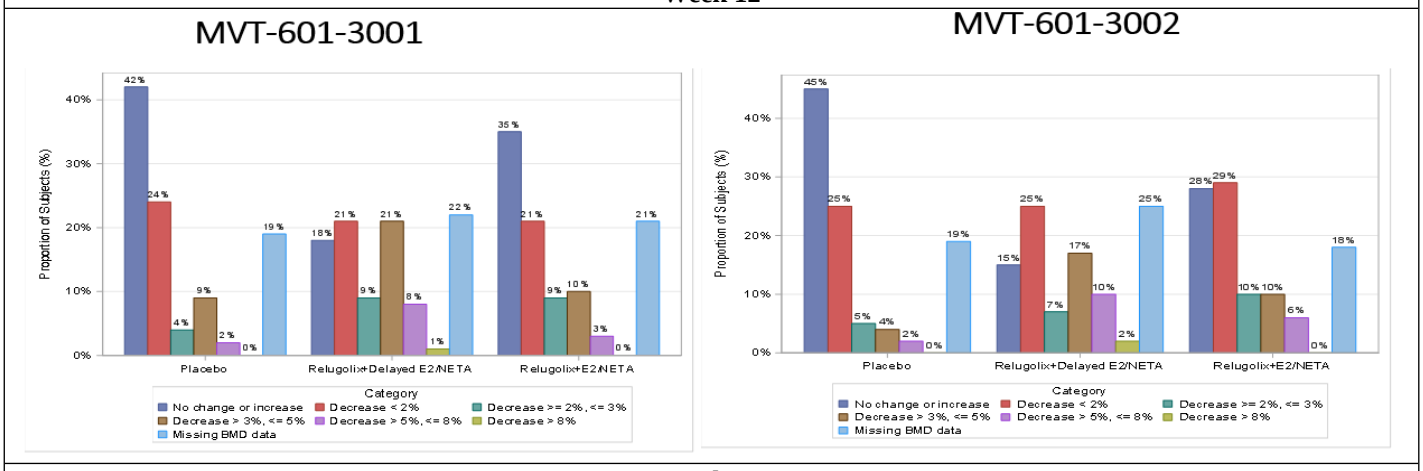
Figure 10: Summary of mean BMD change from baseline at lumbar spine



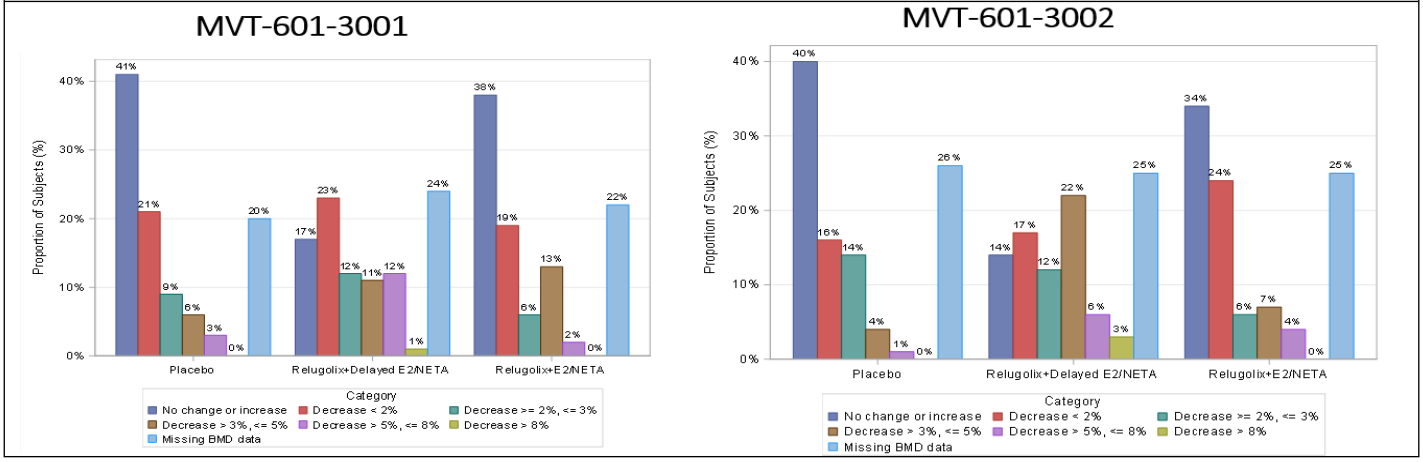
Source: Reviewer's analysis.

Table 37: Summary of mean percentage change from baseline in BMD at lumbar spine categories

Week 12



Week 24



Source: Reviewer's analysis.

7 Appendix B: Description of the Applicant's Sensitivity analysis

Sensitivity Analysis 1: This analysis is conducted to assess the potential impact of unvalidated feminine product use, the primary endpoint will be analyzed as sensitivity analysis in a similar fashion to the primary analysis using the Week 24/EOT validated MBL volume (obtained from the validated or validated-but unauthorized feminine products only and excluding unvalidated products).

Sensitivity Analysis 2: The aim of this analysis is to assess the potential impact of missing data due to inadequate collection of feminine products, the primary endpoint will be analyzed with a sensitivity analysis using the missing data handling rules as described in the table below where the observed MBL volume will be used to assess the responder status at Week 24/EOT when feminine product collection was incomplete. These rules differ from those used in the primary analysis in that no imputation will be implemented for patients with < 100% feminine product

compliance and the reported MBL volume both < 80 mL and a ≥ 50% reduction from Baseline as highlighted in the table below.

Treatment Exposure	FP Collection (FPRR)	Observed MBL Volume	Reason for No FP Collection	Responder Status
< 4 weeks	N/A	N/A	N/A	Imputed as non-responder
≥ 4 weeks	100% FP Compliance	N/A	N/A	Based on the observed MBL volume
	< 100% FP Compliance	MBL volume ≥ 80mL or < 50% reduction from Baseline	N/A	Imputed as non-responder based on the observed MBL volume
		MBL volume < 80mL and ≥ 50% reduction from Baseline	N/A	**Based on the observed MBL volume**
	No FP Collection	N/A	Reported "Amenorrhea"	Imputed as responder
			Reported "Spotting or negligible bleeding" and confirmed by eDiary[a]	Imputed as responder
			Reported "Spotting or negligible bleeding" although not confirmed by eDiary or any other reason, had at least 8 weeks of MBL volume data	Based on the imputed MBL volume
			The entries in the eDiary did not verify "Spotting or negligible bleeding" or any other reason and if had < 8 weeks of MBL volume data	Imputed as non-responders

Abbreviations: eDiary, electronic diary; FP, feminine product; EOT, end of treatment; MBL, menstrual blood loss; N/A, not available.

[a] Defined as those patients who meet the following criteria: eDiary entry rate >70% and no more than 3 consecutive days and no more than 5 days of bleeding/spotting and use of feminine product reported on the eDiary over the Week 24/EOT visit window (see Table 5).

Sensitivity Analysis 3: This analysis is conducted to assess the potential impact of early discontinuation on the primary endpoint, the primary endpoint will be analyzed with a sensitivity analysis defining the patients' responder status as follows:

- Patients who discontinued study drug during the first 4 weeks for any reason or who discontinued study drug between Week 4 and Week 12 due to an adverse event, surgery or other intervention for heavy menstrual bleeding, reported lack of efficacy, or bleeding complaints will be considered as non-responders;

- All other patients will have their responder status defined using data from the Week 24/EOT assessment period using the last observation carried forward method.

Sensitivity Analysis 4: This analysis is aimed at assessing the potential impact of the length and full exposure of the treatment. In this analysis, the primary endpoint will be analyzed for the Completers population as a sensitivity analysis. The Completers population is defined as patients in the mITT population who completed 24 weeks of study treatment.

Sensitivity Analysis 5: The primary endpoint will be analyzed on the Per-Protocol population as a sensitivity analysis, using the methods specified for the primary analysis.

Sensitivity Analysis 6: As a sensitivity analysis to the primary analysis using the mixed-effects model for imputing missing MBL volumes at Week 24/EOT, multiple imputation approach will be implemented. This approach will be used to impute missing or partially missing MBL volume. In this approach, if the patient has amenorrhea, her MBL volume will be set to zero.

Table 38: Summary of study discontinuation reported as "OTHER" and "WITHDRAWAL BY PATIENT"

STUDYID	SUBJID	Treatment	Reported Reason	Description
MVT-601-3001	(b) (6)		OTHER	PER SPONSOR REQUEST. SUBJECT NOT COMPLIANT WITH STUDY VISITS AND PROCEDURES.
			WITHDRAWAL BY SUBJECT	PERSON REASONS - REFUSES ANY ADDITIONAL TESTING INCLUDING EARLY TERMINATION VISIT - STOPPED PILLS 10 DAYS AGO
				SUBJECT WITHDREW CONSENT DUE TO A NEW JOB THAT SHE DID NOT WANT TO JEOPARDIZE WITH THE TIME OFF FOR STUDY VISITS.
				SUBJECT STATES SHE IS MOVING OUT OF STATE AND WILL NO LONGER WILL BE AVAILABLE TO PARTICIPATE.
				PERSONAL DECISION DUE TO PERSONAL AND WORK SCHEDULE
		Placebo		PATIENT HAD A LOT OF FAMILY STUFF GOING ON AND SHE COULD NO LONGER CONTINUE.
				SUBJECT HAS DECIDED NOT TO CONTINUE IN THE TRIAL, DUE TO A RECENT DIAGNOSIS THAT SHE DID NOT WISH TO DISCLOSE AT THIS TIME.
				PATIENT DECIDED NOT TO RETURN TO CLINIC FOR ANY OTHER STUDY VISITS AS SHE IS "DONE MISSING WORK FOR THIS" (RESEARCH VISITS)
			OTHER	EARLY TERMED DUE TO NEEDED PROHIBITED MEDICATION TO TREAT LATENT TB THAT WAS PRESENT AT BASELINE
				NON-COMPLAINT TO IP
		R+delE2/NETA		PROLONGED QTCF WAS AT THE BASELINE VISIT.
			WITHDRAWAL BY SUBJECT	SUBJECT WITHDREW DUE TO WORK SCHEDULE
				WITDREW CONSENT BY SUBJECT DUE TO OF HER OPINION MYOMAS NOT DISAPEARAING
				PT STARTED NEW JOB AND IS UNABLE TO MAKE STUDY VISITS OR COLLECT FEMININE PRODUCTS WHILE WORKING
			OTHER	GENERAL SUBJECT NON-COMPLIANCE (INCLUDING IP AND STUDY PROCEDURES)
				POOR COMPLIANCE WITH STUDY: NOT ATTENDING STUDY VISIT, NOT RETURNING CALLS/EMAILS FROM SITE/ NOT TAKING RECOMMENDED IRON AND VITAMIN D SUPPLEMENTS,
		R+E2/NETA		NON-COMPLIANCE DUE TO TRAVEL OUT THE COUNTRY
				PATIENT WAS NON-COMPLIANT WITH IP
				SUBJECT IS DISCONTINUED BASED OFF OF PI DISCRETION FOR SAFETY, DUE TO HER ELEVATE PROLACTIN LEVELS AT BASELINE.
				SUBJECT MOVED OUT OF STATED AND WITHDREW FROM STUDY.
				SUBJECT STATED NOT HAVING TIME AND SHE IS GOING TO HAVE SURGERY FOR HER FIBROIDS
				SUBJECT NO LONGER WISHED TO PARTICIPATE

STUDYID	SUBJID	Treatment	Reported Reason	Description
	(b) (6)		WITHDRAWAL BY SUBJECT	SUBJECT STARTED A NEW JOB & REFUSES TO ASK OFF FROM HER NEW JOB FOR ANY VISITS.
				SUBJECT DID'T SPECIFY IN DETAIL-ONLY THAT SHE NO LONGER WANTS TO PARTICIPATE & THAT MEDS NOR ANY OF HER EXPERIENCES HAD ANYTHING TO DO WITH DECISION.
				SUBJECT WALKED IN TO THE OFFICE AND RETURNED IP AND E-DIARY AND SAID SHE NO LONGER WANTED TO BE A PART OF THE RESEARCH STUDY AND LEFT.
				SUBJECT WAS ET D DUE TO HER SON S ILLNESS.
				SUBJECT HAS BEEN OUT OF THE COUNTRY DUE TO A FAMILY EMERGENCY. SHE WILL NOT BE ABLE TO CONTINUE SINCE SHE HAS TO GO OUT OF THE COUNTRY AGAIN.
				MOVED OUT OF STATE SUDDENLY
				WITDRAW CONSENT-PATIENT DID NOT GIVE REASON FOR STUDY DISCONTINUATION
MVT-601-3002			OTHER	LOW DIARY COMPLIANCE
				EARLY TERMINATED AS PER INDICATED BY MEDICAL MONITOR DUE TO IP NONCOMPLIANCE. THE SUBJECT MISSED 2 CONSECUTIVE VISITS.
				PI DECISION (NON-COMPLIANCE) WITH STUDY VISIT AND IP
				SUBJECT WAS NON-COMPLIANT IN DOSING
				NON-COMPLIANCE WITH STUDY VISITS
		Placebo	WITHDRAWAL BY SUBJECT	PATIENT WANTS TO GET PREGNANT.
				PATIENT IS RELOCATING TO EAST LONDON
				SUBJECT DECIDED TO TERMINATE THE STUDY; FUP VISIT ON SITE WAS REFUSED BY THE SUBJECT; ONLY MENSTRUATION RECOVERY STATUS WAS OBTAINED FROM THE SUBJECT
				DUE TO SUBJECT'S CHANGE IN HER SHIFT AT WORK AND CAN NOT COMMIT TO THE STUDY.
				WITHDRAWAL BY SUBJECT DUE TO NOT WANTING TO REPEAT THE DEXA SCAN
				SUBJECT WITHDREW CONSENT NO REASON GIVEN
			OTHER	NON-COMPLIANCE WITH EDIARY, IP DOSES, STUDY VISITS, STUDY PROCEDURES.
				NON-COMPLIANCE WITH IP AND STUDY PROCEDURES.
				CONSISTENT NON-COMPLIANCE WITH PRODUCT COLLECTIONS AND FOLLOW UP CARE.
		R+delE2/NETA	WITHDRAWAL BY SUBJECT	SUBJECT DID NOT WANT TO CONTINUE WITH STUDY AND DISCONTINUED IP ON (b) (6)
				SHE WITHDRAWS FROM THE STUDY DUE TO FAMILY PROBLEMS
				PATIENT WANTED SURGERY
				PATIENT NO LONGER WANTED TO PARTICIPATE IN STUDY. RETURNED TO SITE FOR ET VISIT ON (b) (6)
				SUBJECT WITHDREW CONSENT FROM THE STUDY DUE TO WORK CONSTRAINTS
				PATIENT WISHED TO WITHDRAW FROM STUDY ON HER OWN TERMS.
			OTHER	NON-COMPLIANCE WITH E-DIARY ENTRIES
				SUBJECT STATED SHE NO LONGER WISHED TO CONTINUE STUDY.

STUDYID	SUBJID	Treatment	Reported Reason	Description
	(b) (6)	R+E2/NETA	WITHDRAWAL BY SUBJECT	SUBJECT NO LONGER WANTED TO PARTICIPATE IN STUDY DUE TO PERSONAL REASONS
				SHE HAD NO DESIRE TO KEEP GOING, SUBJECT WITHDREW CONSENT.
				PT NO LONGER WISHES TO CONTINUE WITH IP DOSING OR PARTICIPATION IN STUDY
				SUBJECT WITHDREW ICF DUE TO PERSONAL/SOCIAL ISSUES
				PATIENT DECIDED TO WITHDRAW BECAUSE SHE HAS A VERY BUSY WORK SCHEDULE AND SHE IS NO ABLE TO COME FOR REGULAR VISITS.
				SUBJECT STATED SHE CANNOT PARTICIPATE IN THE TRIAL BECAUSE OF HER FAMILY ISSUES
				PATIENT CHOSE TO DISCONTINUE FROM THE STUDY DUE TO THINKING THE MEDICATION WAS AFFECTING HER BONES.
				FAMILY PROBLEMS
				SUBJECT BEGAN A NEW JOB AND WAS NOT BALE TO COMPLETE THE REMAINING VISIT
				SUBJECTS JOB WILL NOT ALLOW HER TO CONTINUE PARTICIPATION
				SUBJECT WITHDREW SECONDARY TO JOB
				SUBJECT DECLINED PROCEDURES AND FURTHER PARTICIPATION

Source: Reviewer's analysis. ET: End of treatment.

This is a representation of an electronic record that was signed electronically. Following this are manifestations of any and all electronic signatures for this electronic record.

/s/

ABEL T ESHETE
05/10/2021 04:08:30 PM

GUOXING SOON
05/10/2021 05:12:00 PM

TSAE YUN D LIN
05/10/2021 08:38:58 PM