STATISTICAL ANALYSIS PLAN

Study number: OG-6219-P001

Study protocol title: A Phase 2a/b, Randomized, Double-blind, Placebo-controlled, Parallel Group, Multicenter, Clinical Study to Evaluate the Efficacy and Safety of OG-6219 in 3 Dose Levels, in Women 18 to 49 Years of Age with Moderate to Severe Endometriosis-related Pain

Protocol version and date: V6, 16Feb2024

Study Phase: 2a/b

AUTHOR: PPD

REVIEWERS:
PPD

STATISTICAL ANALYSIS PLAN VERSION AND DATE: V2.0, 09JUN2025

STATISTICAL ANALYSIS PLAN SIGNATURE PAGE

Statistical Analysis Plan V2.0 (Dated 05MAY2025) for Protocol OG-6219-P001

	Name	Signature	Date (DDMmmYYYY)
Author:	PPD	Refer to eSignature	
Position:	PPD		

Upon review of this document, the undersigned approves this version of the Statistical Analysis Plan, authorizing that the content is acceptable for all reviewers of the document.

	Name	Signature	Date (DDMmmYYYY)
Approved By:	PPD	Refer to eSignature	06/09/2025
Position:	PPD		
PPD			
Approved By:	PPD	Refer to eSignature	
Position:	PPD		
PPD			

MODIFICATION HISTORY

Unique Identifier for this Version	Date of the Document Version	Author	Significant Changes from Previous Authorized Version
2.0	09JUN2025	PPD	This is the first version in the Organon SAP template. It is based on the IQVIA stable SAP 1.0, dated 15Nov2023, with some additions, modifications, and clarifications.

TABLE OF CONTENTS

1. List of Abbreviations

Abbreviation	Term
AE	Adverse event
AESI	Adverse event of special interest
ALT	Alanine aminotransferase
ANCOVA	analysis of covariance
AR (1)	first-order autoregressive 1
AST	Aspartate aminotransferase
BC	Baseline Cycle
BID	Twice a day
BMI	Body mass index
CI	Confidence interval
COVID-19	Coronavirus Disease 2019
CS	compound symmetry
C-SSRS	Columbia-Suicide Severity Rating Scale
CCI	[redacted]
[redacted]	[redacted]
DMC	Data Monitoring Committee
DYS	Dysmenorrhea
E1	Estrone
E2	Estradiol
ECG	Electrocardiogram
eCOA	Electronic clinical outcome assessment
eCRF	Electronic case report form
eDiary	Electronic diary
EHP-30	Endometriosis Health Profile-30
ENR	All Participants Enrolled Analysis Set
EOS	End of Study
EOT	End of Treatment
ePRO	Electronic Patient Reported Outcome
EQ VAS	EQ Visual Analog Scale
CCI	[redacted]
ERP	Endometriosis-related pain
FAS	Full Analysis Set
CCI	[redacted]
hCG	Human chorionic gonadotropin
HRQoL	Health-related quality of life
ICH	International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use
C	[redacted]
MFAS	Modified Full Analysis Set
NCA	Noncompartmental analysis
NMPP	Non-menstrual pelvic pain
NRS	Numeric rating scale

CCI [redacted]

OPP	Overall Pelvic Pain (endometriosis-related)
PD	Pharmacodynamic
PGI-C	Patient Global Impression of Change
PGI-S	Patient Global Impression of Severity
PK	Pharmacokinetic
PP	Per Protocol Analysis Set
PRO	Patient-reported outcome
PROMIS	Patient Reported Outcomes Measurement Information System
PT	Preferred term
QTcF	Corrected QT interval according to CCI formula
SAE	Serious Adverse Event
CCI	[redacted]
SC	Screening Cycle
SD	Standard deviation
SFAS	Safety Follow-up Analysis Set
SFC	Safety Follow-up Cycle
SoA	Schedule of Activities
TEAE	Treatment-emergent adverse event
T_{max}	Time taken to reach the maximum concentration
TRC1	Treatment Cycle 1
TRC2	Treatment Cycle 2
TRC3	Treatment Cycle 3
TSH	Thyroid stimulating hormone
TVUS	Transvaginal ultrasound
ULN	Upper limit of normal
VAS	Visual Analog Scale
WPAI-Endometriosis	Work Productivity and Activity Impairment Questionnaire- Endometriosis

2. INTRODUCTION

The purpose of this statistical analysis plan (SAP) is to describe the methods, specifications, and conventions to be used in the analysis of efficacy and safety data for Protocol OG-6219-P001. The SAP is based on version 6.0 of the protocol. Results of the analyses described in this SAP will be included in the study report. Any post-hoc or unplanned analyses that are not described in this SAP will be clearly identified in the study report.

3. OBJECTIVES, ENDPOINTS, AND ESTIMANDS

The primary and secondary objectives and the corresponding endpoints are described in Table 1 (copied from the study protocol). The estimand and the analysis for the primary efficacy endpoint are described in Table 2. Estimands for secondary efficacy endpoints are defined using the same principle as for the primary efficacy endpoint.

Table 1: Objectives and Endpoints (as described in the study protocol)

Objectives	Endpoints
Primary Endpoints	
The following primary efficacy endpoint will be assessed comparing Group A (CCI BID), B (CCI BID), C (CCI BID) with Group D (Placebo)	
• To evaluate the efficacy of 3 dose levels of OG-6219 (Group A, B, C) versus Placebo (Group D) in reducing OPP during TRC3, as measured by a NRS in the eDiary	• Change from BC to TRC3 in the mean OPP score[a]
The following primary safety endpoints will be assessed for Group A CCI BID), B (CCI BID), C (CCI BID) and Group D (Placebo)	
• To evaluate the safety and tolerability of OG-6219	• Proportion of participants who experienced any AEs/SAEs • Abnormalities in clinical laboratory assessments, vital signs, and physical examination • Proportion of participants who prematurely discontinued study treatment due to AEs/SAEs
Secondary Efficacy Endpoints	
The following secondary endpoints will be assessed comparing Group A (CCI BID), B (CCI BID), C (CCI BID) with Group D (Placebo)	
1. To evaluate the efficacy of 3 dose levels of OG-6219 (Groups A, B, C) versus Placebo (Group	• Change from BC to TRC3 in the mean DYS score

CCI

D) in reducing DYS during TRC3, as measured by a NRS in the eDiary	
2. To evaluate the efficacy of 3 dose levels of OG-6219 (Groups A, B, C) versus Placebo (Group D) in reducing NMPP during TRC3, as measured by a NRS in the eDiary	• Change from BC to TRC3 in the mean NMPP score
3. To evaluate the efficacy of 3 dose levels of OG-6219 (Group A, B, C) versus Placebo (Group D) reducing dyspareunia during TRC3, as measured by a NRS in the eDiary	• Change from BC to TRC3 in the mean dyspareunia score
4. To evaluate the daily use of Sponsor provided rescue medication taken for ERP at TRC1, TRC2, and TRC3	• Change from BC to TRC1, TRC2, and TRC3 in the mean number of tablets of rescue medication for ERP • Change from BC to TRC1, TRC2, and TRC3 in the proportion of days participant has used rescue medication for ERP
5. To evaluate the change in the PGI-S Score at different time points	• Change in PGI-S Score from C to Phone Contact CCI
6. To evaluate the change in the PGI-C Score in different time points	• Percentage of participants with any improvement on the PGI-C at TRC1, TRC2, and TRC3
7. To evaluate the change in the EHP-30 domains (CCI) from BC to TRC3	• Change from BC to TRC3 in the EHP-30 Domain Scores
Secondary Safety Endpoints	
8. To assess the incidence of clinically significant[b] changes from V1 to C in bone biomarkers: C C I	• Mean change from V1 to C in bone biomarker levels
9. To assess the incidence of any clinically significant[b] changes in laboratory parameters from V1 to CC I	• Proportion of participants with clinical parameters of significance from V1 to CCI
10. To assess vaginal bleeding pattern over the 3 menstrual cycles (TRC1, TRC2, and TRC3) as captured in the eDiary	• Mean change from BC to TRC1, TRC2, TRC3 in the number of days on period and number of days with irregular vaginal bleeding/spotting
11. To assess any potential change in the ECG parameters at CCI	• ECG parameter changes at CCI
Secondary PD / PK Endpoints	
12. To assess the change from CCI in serum hormone[c] concentrations	• Mean change from CCI in serum hormone levels
13. To assess any difference in serum hormone[c] concentrations at C between active treatment versus placebo	• Serum hormone levels at V5 comparing Group A (CCI BID), B (CCI BID), C (CCI BID) with Group D (Placebo)

CCI

14. To assess any difference in serum hormone[c] concentrations at C between active treatment versus placebo	• Serum hormone levels at V7 comparing Group A (CCI BID), B (CCI BID), C (CCI BID) with Group D (Placebo)
15. To assess the PK of orally administered OG-6219 and the active metabolite FOR-1011 using sparse and intensive[d] PK sampling.	• Plasma concentrations of OG-6219 and FOR-1011 at scheduled assessments using sparse PK sampling during the treatment period. • C_{max}, t_{max}, and AUC_{tau} for both OG-6219 and FOR-1011 on C and C for the population with intensive PK sampling, where applicable.
CCI	
[redacted]	• [redacted]
[redacted]	• [redacted]
[redacted]	• [redacted]
[redacted]	• [redacted]
[redacted]	• [redacted]
[redacted]	• [redacted]
[redacted]	• [redacted]
[redacted]	• [redacted]
[redacted]	• [redacted]
[redacted]	• [redacted]

Note: Refer below and to Figure 1.

- Screening Cycle is between CCI .
- Baseline Cycle is defined as CCI which starts at C (first day of the menses) and ends at one-day before next menses.

CCI [redacted]

- Treatment Cycle 1 is the period between the first day of menses associated with treatment start (C) to the day prior to the first day of next menses, regardless of number of days.
- Treatment Cycle 2 is the period between the first day of second menses after end of BC to the day prior to first day of next menses, regardless of number of days.
- Treatment Cycle 3 or EOT is the period between the first day of third menses after end of BC to either end CCI on study treatment, whichever occurs first.
- Safety Follow-up Cycle is the period time between the first day after stopping study treatment during TRC3 and ends after 30 consecutive days.

a. OPP score is defined as the combination of the DYS and NMPP scores in a cycle.
b. Clinical significance is defined as any variation in physical findings or laboratory assessments that has medical relevance and may result in an alteration in medical care.
c. CCI
d. Intensive PK sampling is optional and applicable for subset of PK analysis set, where a separate consent would be provided.

Abbreviations:

AE=Adverse event;
BC=Baseline Cycle;
CCI ;
C-SSRS=Columbia-Suicide Severity Rating Scale;
CCI
DYS=Dysmenorrhea;
CCI
eDiary=Electronic diary;
EOT=End of Treatment;
ERP=Endometriosis-related pain;
CCI
NRS=Numeric rating scale;
OPP=Overall Pelvic Pain (endometriosis-related);
PD=Pharmacodynamic;
PGI-S=Patient Global Impression of Severity;
PROMIS= Patient Reported Outcomes Measurement Information System;
CCI ;
t_{max}=Time taken to reach the maximum concentration
TRC1=Treatment Cycle 1;
TRC3=Treatment Cycle 3;
V=Visit;

AUC_{tau}=Area under the plasma-concentration-time curve during a dosing interval, calculated by linear up/log down trapezoidal summation;
BID=Twice a day;
C_{max}=Maximum concentration, obtained directly from the observed concentration versus time data;
CCI ;
CCI
CCI
ECG=Electrocardiogram;
EHP-30=Endometriosis Health Profile-30;
CCI
CCI
NMPP=Non-menstrual pelvic pain;
CCI
CCI
PGI-C=Patient Global Impression of Change;
PK=Pharmacokinetic;
SC=Screening Cycle;
CCI
CCI ;
TRC2=Treatment Cycle 2;
CCI
CCI

CCI

Table 2 Estimand and analysis for primary efficacy endpoint

Primary Efficacy Objective: To evaluate the efficacy of 3 dose levels of OG-6219 (Group A, B, C) versus Placebo (Group D) in reducing OPP during TRC3, as measured by a NRS in the eDiary	
Estimand: The mean difference between each OG-6219 dose arm and placebo in change from BC to TRC3 of the mean OPP score, assuming that all participants completed the treatment period.	
Treatment: OG-6219 CCI BID; OG-6219 CCI BID; OG-6219 CCI BID; placebo BID	
ESTIMAND	**ANALYSIS**
Target Population	**Analysis Set**
Pre-menopausal women between 18 and 49 years of age (inclusive) who have moderate to severe ERP and meet all other study inclusion/exclusion criteria	Participants allocated to treatment who took at least one dose of study medication and who completed at least one subsequent diary entry. Corresponds to the MFAS.
Variable	**Outcome Measure**
Change from BC to TRC3 of the mean OPP score	Change from BC to TRC3 of the mean OPP score Data after the last day of treatment will not be included in TRC1-TRC3, but only in SFC.
Handling of Intercurrent Events	**Handling of Missing Data**
Premature treatment discontinuation for any reason, including adverse event or lack of efficacy: a hypothetical strategy will be used, i.e. it will be assumed that treatment discontinuation did not occur and that participants completed the treatment period. Use of study supplied rescue medication for ERP per the protocol is not considered to be an intercurrent event. Use of prohibited pain medication is not considered to be an intercurrent event.	MAR (Missing at Random) will be assumed for the primary analysis.
Population-level summary measure	**Analysis approach**

Mean difference between each OG-6219 dose arm and placebo in change from BC to TRC3 of the mean OPP score.	The mean OPP score will be analyzed using a longitudinal model, i.e. a Mixed Models for Repeated Measures (MMRM), fitted to data from all participants with a post-treatment OPP measurement. The response vector for the model will be change from BC in the mean OPP score during TRC1, TRC2, and TRC3. Model terms will include the Baseline Cycle value (mean OPP score for BC), treatment, time (i.e., cycle), and treatment-by-time interaction. An unstructured error structure will be assumed. If nonconvergence under the UN assumption is encountered a compound symmetry (CS) and autoregressive AR(1) error structures will be assumed in sequence. Estimated treatment differences (each OG-6219 treatment group versus Placebo) with corresponding 95% CI and 2-sided p-values will be derived from the model for each treatment cycle.

4. STUDY DESIGN

4.1. Overall Study Design

This is a global multicenter, Phase 2a/b, randomized, double-blind, Placebo-controlled, 4-arm proof of concept, dose-range study to assess the efficacy, safety, and tolerability of 3 dose levels of OG-6219, in pre-menopausal women 18 to 49 years of age (inclusive), who have been surgically diagnosed with endometriosis within the last 4 months to 10 years prior to Visit 1 (V1) and with moderate to severe ERP.

The study duration of approximately 28 weeks includes: an initial period of approximately 2 to 4 weeks between V1 and CCI to collect safety laboratory assessments. Each cycle refers to one menstrual cycle. A Screening Cycle (SC), between CC and CCI A Baseline Cycle (BC) or Single-blind CCI, between CC and CCI A Randomization Visit CCI, with treatment allocation CCI to three OG-6219 dosing arms and one Placebo arm. The randomized treatment period consists of three menstrual cycles, Treatment Cycle 1 (TRC1), Treatment Cycle 2 (TRC2), and Treatment Cycle 3 (TRC3), followed by a Safety Follow-up Cycle (SFC). The study includes C C on-site visits and C C phone contacts, as detailed below and illustrated in Figure 1:

CCI

Screening Period:

- SC: Starts at CC (first day of the menses) and ends one-day before next menses.
- BC during the single-blind, CCI period: Starts at CC (first day of the menses) and ends at one-day before next menses.

Randomization: CC, this is on first day of the next menses after the end of BC.

Treatment Period: Three treatment cycles (TRC1, TRC2, and TRC3) after CC:

- TRC1 is the period between the first day of menses associated with treatment start CCI to the day prior to the first day of next menses, regardless of number of days. CCI
- TRC2 is the period between the first day of second menses after end of BC to the day prior to first day of next menses, regardless of number of days. Phone Contact 1 is done during the menses.
- TRC3 is the period between the first day of third menses after end of BC to either end after C C completed cycles or CCI on study treatment, whichever occurs first.

Safety Follow-up:

- SFC, is the period of time between the first day after stopping study treatment ending after 30 consecutive days (regardless of whether it aligns with first day of menses or not). Phone contact C C is done CCI after last intake of study treatment. Visit C C is considered End of Study (EOS).

At CC, participants will receive the eDiary (eCOA) which will be used CCI until the EOS (CC) to CCI collect the (worst) CCI pelvic pain score. The eDiary should be completed at about the same time (before going to bed), within a 24-hour recollection period. The eDiary will also collect CCI as well

CCI

as mental health assessment via C-SSRS will be collected at visits according to the Schedule of Activities (SoA) included in Section 4.4.

Figure 1: Study Schema (copied from the study protocol)

4.2. Statistical Hypotheses

The following three hypotheses will be tested for the primary endpoint:

1) **H0**: The change in mean OPP score from BC to TRC3 is the same for Group C (CCI BID) and Group D (Placebo) vs.
   H1: The change in mean OPP score from BC to TRC3 is different for Group C (CCI BID) and Group D (Placebo).
2) **H0**: The change in mean OPP score from BC to TRC3 is the same for Group B (CCI BID) and Group D (Placebo) vs.
   H1: The change in mean OPP score from BC to TRC3 is different for Group B (CCI BID) and Group D (Placebo).

3) **H0:** The change in mean OPP score from BC to TRC3 is the same for Group A (CCI BID) and Group D (Placebo) vs.
H1: The change in mean OPP score from BC to TRC3 is different for Group A (CCI BID) and Group D (Placebo).

The three hypotheses will be tested separately at the 2-sided α level of 0.05 without any adjustment for multiplicity.

4.3. Sample Size

Approximately 760 pre-menopausal females aged 18 to 49 years old (inclusive), who have been surgically diagnosed with endometriosis within the last 4 months to 10 years prior to V1, will be screened to achieve 380 randomly assigned to the study treatment (95 participants in each group). A total of 380 participants will provide 84% power to detect a difference of 1 point between an OG-6219 dose group and Placebo in change from baseline to TRC3 of mean OPP, assuming a common SD of 2.2. This power calculation is based on a 2-sided test with Type I error of 0.05 and a dropout rate of 10%.
A minimum subset of 10 participants per group (~10%) will be enrolled for optional intensive PK sampling for the entire duration of the study.

4.4. Schedule of Activities

The schedule of activities is shown in the Study Protocol, section 1.3.

5. COVARIATES AND SUBGROUPS

5.1. Planned Covariates and Factors

The statistical analyses will include treatment, time, and treatment-by-time interaction as factors (if only one timepoint is included in the analysis, treatment will be the only factor). The baseline value of the endpoint will be included as a covariate (for endpoints where a baseline value exists).

5.2. Subgroups

Demographic data, baseline characteristics, concomitant medications, and selected safety data will be summarized by region (US vs. Non-US). No specific subgroup analyses are planned, but such may be performed if deemed to be necessary.

6. ANALYSIS SETS

6.1. Definitions of Analysis Sets and Subgroups

Agreement of participants inclusion/exclusion from each analysis set will be conducted prior to database lock.

Analyses will be performed in the respective analysis sets (efficacy analyses in the MFAS for on-treatment efficacy, SFAS for post-treatment efficacy, selected efficacy analyses in the PP analysis set, safety analyses in the SAF, PK analysis in the PK analysis set and PK analysis Set for NCA, and PD analysis in the PD analysis set). The randomized treatment will be used for the ENR, FAS, MFAS and PP analysis sets. The actual treatment received will be used for the SAF, SFAS, PD, PK, and PKNCA analysis sets. If for any participant the randomized treatment differs from the actual treatment received, the primary efficacy analysis will be repeated using the actual treatment received. If a participant would receive a different treatment than the one randomized to during parts of the treatment period, the actual treatment received for this participant will be defined as the treatment with the longest

exposure time.

6.2. All Participants Enrolled Analysis Set [ENR]

The all participants enrolled analysis set will contain all participants who provided informed consent for the study.

6.3. Full Analysis Set [FAS]

The Full Analysis Set will include all randomized participants.

6.4. Modified Full Analysis Set [MFAS]

The Modified Full Analysis Set (MFAS) is defined as all randomized participants who received at least one dose of study medication and who completed at least one subsequent diary entry for DYS or NMPP.

6.5. Per-Protocol Analysis Set (PP)

The Per-Protocol analysis set is defined as a subset of all participants in the MFAS with the following exclusions.

The PP analysis set will exclude participants that either:

- Had critical or major protocol deviation with major impact on the primary endpoint.
- Were non-compliant with randomized treatment, i.e. used <75% of the protocol dose of randomized study treatment (only participants for which compliance could be calculated from randomization to last dose of treatment, and for which this compliance was <75%, will be excluded).

6.6. Safety Analysis Set [SAF]

The Safety Analysis Set will consist of all randomized participants who received at least one dose of randomized study medication, as determined on the drug accountability eCRF page (i.e., participants are expected to either have a date of first dose or indicate that no study medication has been taken).

6.7. Safety Follow-up Analysis Set [SFAS]

The Safety Follow-up Analysis Set will contain all randomized participants who receive at least one dose of study medication and had at least one diary entry for DYS or NMPP after the date of last treatment.

6.8. Pharmacokinetic Analysis Set

The PK analysis set is a subset of the safety analysis set (SAF) and will include all participants who received at least one dose of OG-6219 and have at least 1 quantifiable OG-6219 or FOR-1011 concentration, which is not impacted by protocol violations or events with potential to affect the PK.

6.9. Pharmacokinetic Analysis Set for Noncompartmental Analysis

The PK analysis set for noncompartmental analysis (NCA) will include all participants in the PK analysis set who chose intensive PK sampling in Day 1 (V4) and V5 and have sufficient samples to determine at least one PK parameter for OG-6219 or FOR-1011. Intensive PK sampling is optional where a separate consent would be provided.

6.10. Pharmacodynamic Analysis Set

The Pharmacodynamic Analysis Set is a subset of the SAF and will include all participants who took at least one dose of OG-6219 or Placebo and have at least one quantifiable PD endpoint without protocol deviations or events affecting the PD results.

7. PLANNED ANALYSES AND DATA MONITORING

7.1. Interim Analysis and Early Stopping Guidelines

A Data Monitoring Committee (DMC) will review unblinded safety data during the study. The planned safety analyses for DMC are described in a separate DMC SAP, and are not in scope for this SAP.

An interim psychometric analysis using blinded data will be performed to affirm methods for

establishing pain responder cutoffs. The interim psychometric analysis will be conducted by Adelphi Values under a separate SAP.

An interim PK and exposure-response modelling analysis will be performed by Certara using unblinded data. This analysis is described in a separate PK modelling plan.

7.2. Final Analysis

All planned analyses identified in this SAP will be performed by the ORD Biostatistics and Programming team following authorization of this SAP, database lock, and unblinding.

Final psychometric analyses, PK analyses, and exposure-response modelling analyses are not in scope for this SAP.

8. Statistical Methods

8.1. General Methodology

8.1.1. Summary tables, listings, and figures

All data will be listed and summarized with descriptive statistics in summary tables.

Continuous variables will be summarized for both the measured values and the change from baseline value, using the number of observations, mean, standard deviation (SD), median, minimum, and maximum. Categorical and ordinal variables will be summarized with counts and percentage for each category.

Box-and-whisker plots will be used to visualize efficacy endpoints by timepoint and treatment. Box-and-whisker plots will also be used to visualize laboratory data, vital signs, and ECG measurements. The results of the main efficacy analyses will be illustrated by plotting differences vs. placebo with 95% CIs for each OG-6219 dose group vs. cycle (or visit).

Mean values for each day relative randomization will be plotted by treatment group for OPP, number of rescue medication tablets used, and interference with daily activities.

8.1.2. Reference Start Date and Study Day

Study Day will be calculated from the reference start date and will be used to show start/stop day of assessments and events. Reference start date (Day 1) is defined as the day CCI, which is also the first dose of randomized study medication.

8.1.3. Baseline and Change from Baseline

Baseline is defined as the last non-missing measurement (including unscheduled assessments) taken prior to first dose of randomized study medication (reference start date). In the case where the last non-missing measurement and the reference start date coincide, that measurement will be considered pre-baseline with the following exceptions:
Adverse Events (AEs) and medications commencing on the reference start date will be considered as either pre-baseline or post-baseline based on the collected eCRF assessment of whether the item started prior to the start of study medication.
Visit 4 measurements scheduled to be made after first dose of randomized study medication (e.g. post-dose ECG measurements) will be regarded as post-baseline. Baseline ECG is the average of the triplicates taken prior to first dose of randomized study medication.
For related assessments (e.g., systolic and diastolic blood pressure), it is desirable that all baseline values come from the same measurement date and not from different dates/visits in case one value is missing. In the event where one value is missing, values from the same measurement dates will be used.

For HRQoL, Fatigue, and Productivity Questionnaires (EHP-30, EQ-5D-5L, WPAI-Endometriosis, PROMIS Fatigue Short Form 6a, and PGI-S) assessments collected on CCI will be considered Baseline. In case data were not entered CCI but within a few days from visit 4, these data will still be considered as baseline.

8.1.4. Retests, Unscheduled Visits and Early Termination Data

For by-visit summaries, only scheduled time points will be included. In the case of a retest (same visit number assigned and same date), the last available measurement for that visit will be used for by-visit summaries.

Listings will include scheduled, unscheduled, retest and early discontinuation data.

8.1.5. Imputation of Missing Data

All data will be evaluated as observed and no imputation methods will be used.

8.1.6. Intercurrent Events

Intercurrent events for the include premature treatment discontinuation for any reason, including adverse event or lack of efficacy. Data after the last day of treatment will not be included in TRC1-TRC3, but only in SFC. A hypothetical strategy will be used in the statistical analyses, assuming that premature treatment discontinuation did not occur and therefore all participants completed the treatment period. Consequently, if the mean score for TRC2 and/or TRC3 cannot be calculated, data will be assumed to be missing at random (MAR).

Use of study supplied rescue medication for ERP per the protocol is not considered to be an intercurrent event. Also, use of prohibited pain medication is not considered to be an intercurrent event. In clinical practice, use of rescue medication and pain medication cannot be restricted for subjects with endometriosis. Therefore, treatment effects must be evaluated in the presence of these additional medications.

8.1.7. Multiple Comparisons/Multiplicity

There will be no adjustments for multiple comparisons. The three doses of OG-6219 will be compared to placebo pairwise without any adjustments for multiplicity. Nominal p-values will be reported.

8.2. Study Participants

8.2.1. Disposition of Participants

Participant disposition and withdrawals, and reasons for exclusion from each analysis set will be presented for the ENR set.

The number and percentage of participants who were screened, screen failed (including reason screen failed), allocated to treatment, completed, or discontinued the randomized treatment period (including reasons for treatment discontinuation), completed or withdrew from the study, (including reasons for study withdrawal) will be presented.
In addition, the primary reason for discontinuation for participants who were allocated to study treatment but discontinued prior to study treatment (participants did not receive study treatment) will be presented. Participants who did not receive study treatment will be identified by "no study drug taken" indicated on the Drug Accountability CRF. Completion of treatment and completion of study are recorded on the End of Study Treatment and Disposition eCRF pages, respectively.

8.2.2. Protocol Deviations

The IQVIA Clinical Trial Management System (CTMS) will collect protocol deviations and as per the Protocol Deviation Management Plan.
The number of participants with at least one protocol deviation, and the number of participants with at least one critical or major protocol deviation, will be summarized by deviation category as described in the PDMP for the FAS. Protocol deviations which lead to exclusion from the PP will be summarized.
All critical and major protocol deviations will be listed.

8.2.3. Demographic and Baseline Characteristics

Demographic data and baseline characteristics will be presented by treatment group as well as by region (US vs. Non-US). All variables will be summarized by descriptive statistics for the SAF.

The following demographic and baseline characteristics will be summarized:

Data collected at CC (screening):

- Age (years): 18-34, 35-49 - calculated relative to date of consent
- Sex: Female
- Race (where participants can contribute to multiple race categories, not collected for

France)

- Ethnicity (not collected for France)
- Weight
- Height
- Body Mass Index (BMI) [kg/m^2]

Data collected during the full Baseline Cycle / Run-in:

- Mean DYS pelvic pain score
- Mean NMPP score
- Mean OPP score
- OPP score severity (4≤ mean <8 moderate; mean >=8 severe)
- Daily rescue medication use (% of days in cycle and mean number of tablets per day)
- Mean Dyspareunia score

8.2.4. Medical, Endometriosis Surgical, Gynecological and Obstetric History

General medical history, endometriosis surgical history and gynaecological history at CC will be presented for the Safety Analysis Set. The data will be summarized by treatment group as detailed in sections 8.2.4.1, 8.2.4.2, and will be listed by participant.

8.2.4.1. General Medical History

Medical History will be coded using Medical Dictionary for Regulatory Activities (MedDRA) version 25.1 or higher.

- Medical History conditions are defined as relative medical conditions which are reported at screening prior to start of treatment and may or may not be ongoing as of the screening visit.
- Medical history conditions will be presented by System Organ Class (SOC) and Preferred Term (PT), with SOC in alphabetical order and PT in order of decreasing frequency.

8.2.4.2. Endometriosis Surgical history

Summaries of endometriosis surgical history data will include the number and percentage of study participants by treatment group for the following:

- Surgical Diagnosis Method (Laparoscopy/Laparotomy)
- Years of Surgical Diagnosis of Endometriosis (<5 years, >=5 years) - calculated relative to date of consent

8.2.4.3. GYNECOLOGICAL AND OBSTETRIC HISTORY

Gynecological history data will be summarized by treatment group for the following:

- Age of Menarche (years)
- Typical Cycle Length (days)
- Usual Duration of Flow (days)
- Usual Volume of Flow (Low/Moderate/Heavy)
- Presence of Clots during Flow (Yes/No)

Summaries of pregnancy history will include the number and percentage of study participants with previous pregnancies as well as descriptive statistics for the overall number of previous pregnancies (gravidity number), number of live births and the number of times the participant has given birth to a fetus with gestational age of 24 weeks or more (parity number).

- Number of Previous Pregnancies
- Gravidity
- Number of Live Births
- Parity

8.2.5. Prior, Concomitant, and Post-Treatment Medications

Medications entered in the Medical History form of the eCRF will be presented for the Safety Analysis Set and coded using WHO-Drug dictionary, September 2022 or later.
In the case where it is not possible to define a medication as prior, concomitant, or post-treatment, the medication will be classified by the worst case, i.e., concomitant.

- 'Prior' medications are defined as any medications started prior to the day of the first dose of randomized study treatment.
- 'Concomitant' medications are defined as any medications taken on the day of or after the first dose of randomized study treatment and before or on the day of last

dose of randomized study treatment.

- 'Post-Treatment' medications are defined as any medications started after the day of last dose of randomized study treatment

Prior, concomitant, and post-treatment medication use will be summarized separately by preferred name using frequencies and percentages. Medications will be sorted alphabetically by Anatomical Therapeutic Chemical (ATC) Level 3 and preferred name in summaries. Participants with multiple occurrences of a medication in the same ATC Level 3 or preferred name will only be counted once within the PT for each summary. If a medication is unable to be categorized to ATC Level 3, the most detailed level available will be used. The type of medication (Prior/Concomitant) will be indicated in the listing of all medications.

The use of prohibited pain medication registered in the eDiary will be summarized by treatment and treatment cycle for the MFAS. All use of prohibited pain medication will be listed in detail.
Use of supplied rescue medication for management of endometriosis related pain, that was reported in the eDiary as an efficacy endpoint, will not be included in the summaries and listings described above.

8.2.6. Study Treatment Exposure and Compliance

8.2.6.1. Study Treatment Exposure

Exposure to randomized study treatment in days will be presented for the Safety Analysis Set.

The date of first randomized study treatment administration will be defined as the date of the randomization visit. The date of last randomized study treatment will be taken from the eCRF End of Study Treatment form.
Interruptions and lack of compliance will not be taken into account for duration of exposure.

The number and percentage of participants with exposure to randomized treatment will be summarized by treatment group for the following categories:

- <4 weeks

- 4-<8 weeks
- 8-<12 weeks
- >=12 weeks

Duration of exposure (days) = (date of last randomized study treatment administration – date of first randomized study treatment administration) + 1.

8.2.6.2. STUDY TREATMENT COMPLIANCE

Compliance with the use of study treatment overall will be presented for the SAF. Number of tablets dispensed and number of tablets returned will be taken from the eCRF Drug Accountability form. Non-compliance is defined as compliance to study treatment <75%. Non-compliance to study treatment will be summarized overall using descriptive statistics by treatment group.

Compliance (%) with randomized study treatment will be calculated based on drug accountability data as the total number of tablets taken (total dispensed – total returned) divided by the total expected number of tablets (number of tablets should be taken per day times duration of exposure) expressed as a percentage, see calculation below.

$$100 \times \frac{\sum(\#\text{tablets dispensed} - \#\text{tablets returned})}{\#\text{tablets that should be taken per day} \times \text{duration of exposure}}$$

Where # of tablets that should be taken per day is 6 tablets (1 tablet from Bottle A BID and 2 tablets from Bottle B BID), and duration of exposure (days) = date of last randomized study treatment administration – date of first randomized study treatment administration + 1.

For participants where bottles were not returned or lost, compliance will not be calculated but set to missing.

8.3. Analyses Supporting Primary Efficacy Objective

8.3.1. Efficacy assessments

Participants will complete the efficacy assessments in the CCI eDiary in the order presented

below:

- Vaginal bleeding pattern
- Endometriosis-related pain NRS for OPP (including vaginal bleeding)
- Dyspareunia
- Interference with daily activities
- Rescue medication use and number of tablets
- Prohibited pain medication use and reason

HRQoL and Productivity Questionnaires are conducted at Study visits per SoA:

- PGI-S
- PGI-C
- EHP-30
- WPAI-Endometriosis
- PROMIS Fatigue Short Form 6a
- EQ-5D-5L

Participants will complete eDiary items CCI asking about the severity of their worst pain over the last 24 hours on 11-point NRS for endometriosis-related pain. Endometriosis-related pain severity will be provided based upon an NRS from 0 (no pain) to 10 (worst pain imaginable) where lower value represents a better outcome.
CCI DYS is defined as the NRS endometriosis-related pain severity entered on a menstrual bleeding day, and CCI NMPP is defined as the NRS endometriosis-related pain severity entered on a non-menstrual bleeding day. CCI OPP is defined as the NRS endometriosis-related pain severity entered on any day regardless of menstrual bleeding (i.e. one cycle of CCI OPP scores consists of CCI DYS scores and CCI NMPP scores for the cycle).

The definition for a menstrual bleeding day used in the determinations for CCI DYS and CCI NMPP will be based upon bleeding information collected from the CCI diary for the two questions below:

- (1a) During the past 24 h did you have any vaginal bleeding or spotting?

- (1b) During the past 24 h have you been on your period?

If both questions are answered "Yes", then that day qualifies as a menstrual bleeding day and the pain score is DYS. If either question is answered "No", then that day is a non-menstrual bleeding day and the pain score is NMPP. OPP score is defined as DYS for menstrual bleeding days and as NMPP for non-menstrual bleeding days.

8.3.2. Primary Efficacy Endpoint and Estimand

The primary efficacy endpoint is change from BC to TRC3 of the mean OPP score. The estimand is the mean difference between each OG-6219 dose arm and placebo in change from BC to TRC3 of the mean OPP score, assuming that all participants completed the treatment period, i.e. a hypothetical estimand. The estimand framework is described in Table 2.

8.3.3. Derivation of Primary Efficacy Endpoint

The mean OPP will be calculated for each cycle as the total of CCI pelvic pain scores reported during the cycle divided by the number of days during the cycle when a CCI score was reported. The mean OPP score will be calculated if there is at least one CCI pain score reported within that cycle.

8.3.3.1. Definition of cycles

The definition of cycles for the calculation of averages over cycles are as follows:

- Screening cycle (SC) – starts at the start of menses associated with CCI and ends the day before the start of menses associated with CCI
- Baseline cycle (BC) – starts at the start of menses associated with CCI and ends at the day before the start of menses associated with CCI or the day of CCI whichever comes first.
- Treatment cycle 1 (TRC1) – starts at the day of CCI (randomization visit=day of first dose of randomized treatment) and ends the day before the start of menses associated with phone contact C (PC1).

CCI

- Treatment cycle 2 (TRC2) – starts at the start of menses associated with phone contact CCI and ends the day before the start of menses associated with CCI
- Treatment cycle 3 (TRC3) – starts at the start of menses associated with CCI and ends the day before the start of menses associated with CCI or, at the latest after CCI of randomized treatment.
- Safety follow-up cycle (SFC) – starts the day after the last dose of randomized treatment and ends 30 days after the last dose of randomized treatment.

Days after last dose of randomized treatment will only be included in SFC, and not in TRC1-TRC3.

8.3.3.2. START OF MENSES

The primary source for determining start of menses dates is the eCRF. At each visit the first date of current menses/menstrual cycle was recorded in the eCRF. Since visits could be performed before the start of menses, the same first date of current menses could be recorded at two subsequent visits (if the next menses had not yet started when the visit was performed, the start date of the current menses was the same as at the previous visit). Therefore, the eCRF did not capture all start of menses days necessary to define the cycles.
To account for start of menses days not captured by the eCRF and correctly define the cycles, additional start of menses days could be identified using the information from the eCOA. The eCOA included the following two daily questions:

- 1a. During the past 24 h, did you have any vaginal bleeding or spotting?
- 1b. During the past 24h, have you been on your period?

If both questions were answered "yes", the participant was regarded to have menses according to the eCOA.
Additional start of menses days were identified from the eCOA using the following algorithm:

- After randomization: If the participant had menses according to the eCOA in the time interval starting 15 days after the previous start of menses day recorded in the eCRF, and ending 4 days before the next start of menses day recorded in the eCRF

(the latter criteria only applies if there is a next start of menses recorded in the eCRF), the first day of menses according to the eCOA in that time interval was defined as an additional start of menses day for that participant.

- Before randomization: If the participant had menses according to the eCOA in the time interval starting 36 days before CCI and ending 4 days before CCI the first day of menses according to the eCOA in that time interval was defined as an additional start of menses day for that participant.

Additional start of menses days identified from the eCOA were added to the start of menses days in the eCRF **only** in cases where either a start of menses day in the eCRF was missing, or when a start of menses day in the eCRF was the same as the start of menses day recorded at the previous visit (for the start of menses associated with CCI also when it was the day before the start of menses day associated with CCI If the start of menses day associated with CCI was missing, or recorded as the same as the start of menses associated with CCI (or before), the start of menses date associated with CCI was defined as the date of CCI .

8.3.4. Intercurrent Event Handling and Data Imputation for Primary Efficacy Endpoint

See section 8.1.6, a hypothetical strategy will be used in the statistical analyses. If the mean score for TRC2 and/or TRC3 cannot be calculated, data will be assumed to be missing at random (MAR). No missing data will be imputed for the primary efficacy endpoint.

8.3.5. Analysis of Primary Efficacy Endpoint

The primary endpoint will be analysed using a longitudinal model, i.e. a Mixed Models for Repeated Measures (MMRM), fitted to data from all participants with a post-treatment OPP measurement. The response vector for the model will be the change from BC in the mean OPP during TRC1, TRC2, and TRC3. Model terms will include the baseline cycle value (mean OPP for BC), treatment, time (i.e., cycle), and treatment-by-time interaction. An unstructured error structure will be assumed. If nonconvergence under the UN assumption is encountered a compound symmetry (CS) and autoregressive AR(1) error structures will be assumed in sequence. Estimated treatment differences (each OG-6219 treatment group versus

Placebo) with corresponding 95% CI and 2-sided p-values will be derived from the model for each treatment cycle.

Analysis using MMRM method will be implemented using the SAS® procedure MIXED as follows (example code only):

```
PROC MIXED DATA = …;
    CLASS CYCLE TREATMENT SUBJID;
        MODEL CHG = BASELINE TREATMENT CYCLE TREATMENT*CYCLE /
        DDFM=KR;
    REPEATED CYCLE / SUBJECT = SUBJID TYPE=UN;
    LSMEANS TREATMENT*CYCLE / PDIFF CL;
RUN;
```

where CHG = change from baseline at each treatment cycle
TREATMENT = randomized treatment group (CCI BID, CCI BID, CCI BID, and placebo)
BASELINE = baseline cycle value
CYCLE = treatment cycle
SUBJID = subject ID

In addition, the mean OPP will be summarized by cycle and treatment group using descriptive statistics of observed data and change from baseline cycle values. Mean OPP by cycle will also be presented in listings.

8.3.6. Sensitivity Analyses for the Primary Efficacy Endpoint

The following sensitivity analyses will be performed for the primary efficacy variable:

- The primary analysis will be repeated for the PP analysis set (sensitivity to analysis set).
- The primary analysis will be repeated, excluding cycles with <10 CCI NMPP scores or <2 CCI DYS scores (sensitivity to missing eCOA data).
- The primary analysis will be repeated, excluding the OPP score on days when prohibited pain medication was used (as reported in the eDiary) (sensitivity to effect of concomitant medication).

8.3.7. Supplementary Analyses for Primary Efficacy Endpoint

Two supplementary analyses will be performed for the MFAS, both using analysis of covariance (ANCOVA) models on the change from baseline to TRC3 with treatment as a fixed factor and baseline OPP as a covariate. In the first analysis, only subjects with data from TRC3 will be included. In the second analysis, subjects without data from TRC3 will have their TRC3 value imputed using last observation carried forward (i.e. the last available of the TRC2 and TRC1 values will be used.)
Estimated treatment differences (each OG-6219 treatment group versus Placebo) with corresponding 95% CI and 2-sided p-values will be derived from the model.

Analysis using an ANCOVA model will be implemented using the SAS® procedure MIXED as follows (example code only):

```
PROC MIXED DATA = …;
    CLASS TREATMENT;
    MODEL CHG = BASELINE TREATMENT;
    LSMEANS TREATMENT / PDIFF CL;
RUN;
```

where CHG = change from baseline to TRC3
TREATMENT = randomized treatment group (CCI BID, CCI BID, CCI BID, and placebo)
BASELINE = baseline cycle value

8.4. Analyses Supporting Secondary Efficacy Objectives

8.4.1. Secondary Efficacy Endpoints

ENDOMETRIOSIS-RELATED PAIN (DYS, NMPP, AND DYSPAREUNIA) SCORES

The secondary efficacy variables of endometriosis-related pain scores are:

- Change from BC to TRC3 in the mean DYS score
- Change from BC to TRC3 in the mean NMPP score
- Change from BC to TRC3 in the mean dyspareunia score

The mean DYS and mean NMPP scores will be calculated similar to the mean OPP score. The mean DYS score will be calculated for each cycle as the total of CCI DYS scores reported during the cycle divided by the number of days during the cycle when a DYS score was reported. The mean NMPP score will be calculated for each cycle as the total of CCI NMPP scores reported during the cycle divided by the number of days during the cycle when a NMPP score was reported.

The mean dyspareunia score will be calculated for each cycle as the total of dyspareunia scores reported during each cycle divided by the number of days when participants engaged in any sexual activity that involved full vaginal penetration during each cycle (i.e., the number of days during the baseline cycle when a dyspareunia score was reported).

Rescue Medication Use for ERP

Participants will be asked daily whether they used the rescue medication in the past 24 hours to treat their ERP. If yes, then number of tablets will be documented.
The secondary efficacy variables of rescue medication use for ERP are:

- Change from BC to TRC1, TRC2, and TRC3 in the mean number of tablets of rescue medication taken for ERP
- Change from BC to TRC1, TRC2, and TRC3 in the proportion of days participant has taken rescue medication for ERP

The mean number of tablets of rescue medication for ERP will be calculated for each cycle as the total number of tablets of rescue medication for ERP reported during the cycle, divided by the number of days during the cycle when a value was reported.
The proportion of days participant has used rescue medication for ERP will be calculated for each cycle as the total number of days participant has used ANY rescue medication for ERP, divided by the number of days during the cycle when a value was reported.

PGI-S and PGI-C Score

The PGI-S is a single item measuring the overall severity of pelvic pain over the past 7 days on a 4-point Likert-type scale (0=None, 1=Mild, 2=Moderate, 3=Severe) where lower values represent a better outcome.

The PGI-C is a single item measuring the change in pelvic pain since the start of the study treatment on a 5-point scale: 0=much better, 1=a little better, 2=no change, 3=a little worse, and 4=much worse.

The secondary efficacy variables of PGI-S and PGI-C are:

- Change in PGI-S Score from V4 to Phone Contact CCI
- Percentage of participants with any improvement on the PGI-C at TRC1, TRC2, and TRC3
- Percentage of participants with any improvement on the PGI-S at TRC1, TRC2, and TRC3

Participants with any improvement on the PGI-C is defined as a PGI-C score of 0 or 1 (0=much better, 1=a little better).

Participants with any improvement on the PGI-S is defined as an improvement from baseline of at least one step in the score.

ENDOMETRIOSIS HEALTH PROFILE-30 (EHP-30)

The EHP-30 is a validated disease specific patient-reported outcome (PRO) instrument measuring the HRQoL of women with endometriosis. The EHP-30 consists of 30 items covering the following domains: pain (11 items, Q1-Q11), controls and powerlessness (6 items, Q12-Q17), emotional well-being (6 items, Q18-Q23), social support (4 items, Q24-Q27), and self-image (3 items, Q28-Q30). Each item is answered on a 5-point Likert-type scale (0=never, 1=rarely, 2=sometimes, 3=often, 4=always).

The secondary efficacy variables of EHP-30 are:

- Change from CCI (BC) to CCI (TRC3) in the EHP-30 domain pain (11 items, Q1-Q11)
- Change from CCI (BC) to CCI (TRC3) in the EHP-30 domain control and powerlessness (6 items, Q12-Q17)
- Change from CCI (BC) to CCI (TRC3) in the EHP-30 domain emotional well-being (6 items, Q18-Q23)
- Change from CCI (BC) to CCI (TRC3) in the EHP-30 domain social support

CCI

(4 items, Q24-Q27)

- Change from CCI (BC) to CCI (TRC3) in the EHP-30 domain self-image (3 items, Q28-Q30)

Scores for each domain are transformed from raw scores to range from 0 (indicating the best health status) to 100 (worst health status). The domain score is calculated as the total of raw scores for each item in the domain divided by the total possible raw score times 100. Lower scores indicate better QoL.

8.4.2. Intercurrent Event Handling and Data Imputation for Secondary Efficacy Endpoints

See section 8.1.6, no missing data will be imputed for the secondary efficacy endpoints.

8.4.3. Analysis of Secondary Efficacy Endpoint(s)

ANALYSIS OF ENDOMETRIOSIS-RELATED PAIN (DYS, NMPP, AND DYSPAREUNIA) SCORES

Similar analysis for the primary efficacy variable described in Section 8.3.5 using longitudinal model will be performed for change from BC to TRC3 in the mean DYS score, the mean NMPP score, and the mean dyspareunia score.

Similar sensitivity/supplementary analysis as for the primary efficacy variable described in Section 8.3.6 and 8.3.7 will be performed for change from BC to TRC3 in the mean CCI DYS score and the mean CCI NMPP score.

For the analysis of change from BC to mean CCI DYS score, the sensitivity analysis will exclude cycles with <2 CCI DYS scores.

For the analysis of change from BC to mean CCI NMPP score, the sensitivity analysis will exclude cycles with <10 CCI NMPP scores.

Similar supplementary analyses as for the primary efficacy variable described in Section 8.3.7 will be performed for change from BC to TRC3 in mean dyspareunia score. The sensitivity analysis described in Section 8.3.6 based on the PP analysis set will also be performed.

ANALYSIS OF RESCUE MEDICATION USE FOR ERP

Similar analysis for the primary efficacy variable described in Section 8.3.5 using a

longitudinal model will be performed for change from BC to TRC3 in the mean number of tablets of rescue medication for ERP and the proportion of days participant has used rescue medication for ERP. Similar supplementary analyses as for the primary efficacy variable described in Section 8.3.7 will be performed for change from BC to TRC3 in the mean number of tablets of rescue medication for ERP and the proportion of days participant has used rescue medication for ERP.

ANALYSIS OF PGI-S

Change from V4 to Phone Contact CCI [redacted] in the PGI-S Score will be summarized in a shift table by treatment group and visit.

ANALYSIS OF PGI-C

The number and percentage of participants with any improvement on the PGI-C at phone contact CCI [redacted] will be summarized by treatment group and visit.

Results will be expressed as the number and percentage of subjects with any improvement on the PGI-C.

ANALYSIS OF EHP-30

The five EHP-30 Domain scores will be summarized by descriptive statistics of observed data and change from baseline values.

A similar analysis for the primary efficacy variable described in Section 8.3.7 using ANCOVA will be performed for change from baseline in each EHP-30 Domain score.

8.5. Analyses Supporting Exploratory Efficacy Objectives

8.5.1. Exploratory Efficacy Endpoints

ENDOMETRIOSIS-RELATED PAIN (DYS, NMPP, OPP) SCORES

The exploratory efficacy endpoints for endometriosis-related pain scores are:

- Change from BC to TRC1 of the mean OPP score
- Change from BC to TRC1 of the mean DYS score

- Change from BC to TRC1 of the mean NMPP score
- Change from BC to TRC2 of the mean OPP score
- Change from BC to TRC2 of the mean DYS score
- Change from BC to TRC2 of the mean NMPP score
- Change from SC to SFC of the mean OPP score
- Change from SC to SFC of the mean DYS score
- Change from SC to SFC of the mean NMPP score

Change from SC to SFC of the score where means for a cycle will be calculated if there is at least one CCI score is reported within each of these cycles.

EUROQOL-5 DIMENSIONS 5-LEVEL VERSION (EQ-5D-5L)

The EQ-5D-5L consists of the EQ-5D descriptive system and the EQ Visual Analog Scale (EQ-VAS). The descriptive system consists of five dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each dimension has five levels: no problems, slight problems, moderate problems, severe problems, unable to/extreme problems. The participant is asked to indicate her health state in the eDiary by ticking the box next to the most appropriate statement in each of the five dimensions. This decision results in a 1-digit number that expresses the level selected for that dimension. The digits for the five dimensions are combined into a digit number that describes the participant's health state. The EQ-VAS records the participant's self-rated health on a vertical VAS where the anchors are labeled 'Best health you can imagine' and 'Worst health you can imagine'.

The following EQ-5D-5L exploratory efficacy endpoints will be evaluated:

- Change from CCI (BC) to CCI (TRC3) in EQ-VAS
- Change from CCI (BC) to CCI (TRC3) in each of the five dimensions

INTERFERENCE WITH DAILY ACTIVITIES

Participants will complete eDiary items asking about any interference with daily activities over the last 24 hours on 11-point NRS. The NRS measures the level of interference 0 (not difficult) to 10 (extremely difficult) where lower value represents a better outcome.

The following exploratory efficacy endpoint will be evaluated:

CCI

- Mean change from BC to TRC3 in level of difficulty to do daily activities

The mean level of difficulty to do daily activities will be calculated for each cycle as the total of level of difficulty to do daily activities scores reported during the cycle divided by the number of days during the cycle in which the score was reported. The mean level of difficulty to do daily activities score for a cycle will be calculated if there is at least one level of difficulty to do daily activities score reported within that cycle.

WPAI-ENDOMETRIOSIS

The WPAI-Endometriosis is a PRO instrument that measures the effect of endometriosis on work productivity and daily tasks. The WPAI-Endometriosis questionnaire contains six items (Q1-Q6) that produce four different scale scores. WPAI outcomes are expressed as impairment percentages, with higher numbers indicating greater impairment and less productivity. The scoring algorithms for the four WPAI: Endometriosis scales are as follows:

1. Absenteeism (% work time missed due to endometriosis) = $\frac{Q2}{Q2+Q4} \times 100$;
2. Presenteeism (% impairment while working due to endometriosis) = $\frac{Q5}{10} \times 100$;
3. Work productivity impairment (% work impairment due to endometriosis) = $[\frac{Q2}{Q2+Q4} + ((1 - \frac{Q2}{Q2+Q4}) \times \frac{Q5}{10})] \times 100$; and
4. Daily activity impairment (% activity impairment due to endometriosis) = $\frac{Q6}{10} \times 100$.

Note. If Q4=0, Q5 will be missing from the questionnaire and should be replaced by 0 in the calculation; If Q2=0 and Q4=0, Absenteeism will be defined as 0.

The WPAI-Endometriosis diary is completed during CCI and CCI EOT.

The following WPAI-Endometriosis exploratory efficacy endpoints will be evaluated:

- Mean change from CCI (baseline) to CCI EOT in absenteeism
- Mean change from CCI (baseline) to CCI EOT in presenteeism
- Mean change from CCI (baseline) to CCI EOT in work productivity impairment
- Mean change from CCI (baseline) to CCI EOT in daily activity impairment

PROMIS FATIGUE SHORT FORM 6A

The PROMIS Fatigue Short Form 6a is a validated PRO instrument measuring the severity of

CCI

fatigue in participants. The PROMIS Fatigue Short Form 6a consists of 6 items in which each item is answered using a 5-point Likert-type scale (1=not at all, 2=a little bit, 3=somewhat, 4=quite a bit, 5=very much), and the total score is the sum of the six item scores. Total scores range from 6 (indicating the best health status) to 30 (worst health status). Raw total scores are converted to T-scores using conversion tables (provided in the PROMIS Fatigue scoring manual and shown in the table below), which standardizes the score with a mean of 50 and standard deviation of 10 (a higher T-score indicates worse fatigue).

Fatigue 6a - Adult v1.0		
Short Form Conversion Table		
Raw Score	T-score	SE*
6	33.4	4.9
7	39.1	2.9
8	42.0	2.4
9	44.2	2.2
10	46.1	2.1
11	47.8	2.1
12	49.4	2.1
13	50.9	2.0
14	52.4	2.0
15	53.7	2.0
16	55.1	2.0
17	56.3	1.9
18	57.5	1.9
19	58.8	1.9
20	60.0	1.9
21	61.2	1.9
22	62.4	1.9
23	63.7	2.0
24	65.0	2.0
25	66.4	2.0
26	67.8	2.0
27	69.3	2.0
28	71.0	2.1
29	73.0	2.5
30	76.8	3.8

*SE = Standard Error on T-score metric

The PROMIS Fatigue Short Form 6a is completed during Visit 4 and Visit 7/EOT.

The following exploratory efficacy endpoint will be assessed:

- Change from Visit 4 to Visit 7/EOT in the PROMIS Fatigue Short Form 6a T-score

MIRNA ANALYSIS

A panel of 6 circulating microRNAs (miR-125b-5p, miR-451a, miR-3613-5p, miR-150-5p, miR-342-3p, and let-7b) will be collected at CC/Screening and CCI EOT.

The following exploratory efficacy endpoints will be assessed:

- Change from CC/Screening to CC/EOT in miRNA expression levels (miR-125b-5p, miR-451a, miR-3613-5p, miR-150-5p, miR-342-3p, and let-7b)

8.5.2. Intercurrent Event Handling and Data Imputation for Exploratory Efficacy Endpoints

See section 8.1.6, no missing data will be imputed for the exploratory efficacy endpoints.

8.5.3. Analysis of Exploratory Efficacy Endpoint(s)

ANALYSIS OF ENDOMETRIOSIS-RELATED PAIN (DYS, NMPP, OPP) SCORES

Results for changes from BC to TRC1 and TRC2 will be extracted from the longitudinal analyses used to analyse changes from BC to TRC1, TRC2, and TRC3 (see section 8.3.5).

A similar analysis for the primary efficacy variable described in Section 8.3.7 using ANCOVA will be performed for the change from SC to SFC in the mean OPP score, the mean DYS score, and the mean NMPP score for SFAS.

ANALYSIS OF EUROQOL-5 DIMENSIONS 5-LEVEL VERSION (EQ-5D-5L)

A similar analysis for the primary efficacy variable described in Section 8.3.5 using ANCOVA will be performed for the change from BC to TRC3 in EQ-VAS.

The change from BC to TRC3 in each of the five dimensions (Mobility, Self-care, Usual Activities, Pain/Discomfort, and Anxiety/Depression) scores will be summarized by treatment group.

ANALYSIS OF INTERFERENCE WITH DAILY ACTIVITIES

A similar analysis for the primary efficacy variable described in Section 8.3.5 using longitudinal model will be performed for the mean change from BC to TRC3 in level of difficulty to do daily activities.

ANALYSIS OF WPAI-ENDOMETRIOSIS

Similar analyses as for the primary efficacy variable described in Section 8.3.7 using ANCOVA will be performed for the WPAI-Endometriosis exploratory efficacy endpoints.

ANALYSIS OF PROMIS FATIGUE SHORT FORM 6A

Similar analyses as for the primary efficacy variable described in Section 8.3.7 using ANCOVA will be performed for the change from CCI to CCI EOT in the PROMIS Fatigue Short Form 6a T-score.

ANALYSIS OF MIRNA ANALYSIS

Similar analyses as for the primary efficacy variable described in Section 8.3.7 using ANCOVA will be performed for the change from baseline to CCI EOT in miRNA expression levels (miR-125b-5p, miR-451a, miR-3613-5p, miR-150-5p, miR-342-3p, and let-7b).

8.6. Analyses Supporting Secondary PD Objectives

8.6.1. Secondary PD Endpoints

Serum hormone levels (LH, FSH, E2, E1, Progesterone) and androgens (testosterone, free testosterone, DHEA, and DHEAS) will be assessed at CCI /EOT.

8.6.2. Intercurrent Event Handling and Data Imputation for Secondary PD Endpoints

See section 8.1.6, no missing data will be imputed for secondary PD/PK endpoints.

CCI

8.6.3. Analyses of Secondary PD Endpoints

Secondary PD endpoints will be analysed using the PD Analysis Set. Treatment comparisons between the OG-6219 dose groups and placebo will be conducted using ANCOVA on the change from baseline to CC and CC separately, with treatment as a fixed factor and the baseline value as a covariate. Estimated treatment differences with corresponding 95% CI and 2-sided p-values will be derived from the ANCOVA model. A multiplicative model will be used, i.e. the dependent variable and the covariate will be log-transformed before analysis and estimated treatment differences and 95% CI will be obtained by back-transformation.

8.7. Safety Analyses

8.7.1. Adverse Events

AEs will be coded using MedDRA coding dictionary, Version 25.1 or higher.

Treatment emergent adverse events (TEAEs) are defined as events that first occur or worsen (increase in severity) after the first dose of the study medication, during the Treatment Period, and up to 14 days (inclusive) after the last dose of study medication. Events starting on the day of first dose will be considered treatment-emergent unless information is available to indicate the event started prior to the time of first dose. In the case where it is not possible to define an AE as treatment emergent or not due to a partial start date, the AE will be classified by the worst case, i.e., treatment emergent. Pre-treatment AEs/SAEs starting before the on-treatment period, including those starting on the day of first dose of study medication yet prior to the study medication intake, will be excluded from the TEAE analyses (unless there is an increase in investigator assessed intensity or causality, or the AE meets SAE criteria post-treatment allocation).

An overview table of number of participants within each of the categories described in the sub-section below, will be provided. The broad AE categories consisting of the percentage of participants with any TEAE, a TEAE related to study treatment, a serious TEAE, a TEAE leading to treatment discontinuation, a TEAE leading to study discontinuation, an AE with an outcome of death, and an AESI will be summarized.

Listings will include all AEs, including those that were not treatment emergent.

8.7.1.1. ALL TEAEs

Incidence of TEAEs will be presented by SOC and PT, with SOC in alphabetical order and PT in order of decreasing frequency. The incidence of TEAEs will be broken down further by maximum severity and relationship to study treatment.
The TEAE summary by SOC and PT will also be presented for the region subgroup.

Severity

Severity is classed as mild/ moderate/ severe (increasing severity). TEAEs with a missing severity will be classified as severe. If a participant reports a TEAE more than once within that SOC/ PT, the AE with the worst-case severity will be used in the corresponding severity summaries.

Relationship to Study Treatment

Relationship, as indicated by the Investigator, is classed as 'unrelated', 'unlikely to be related', 'possibly related', and 'probably related' (increasing severity of relationship). A 'related' TEAE is defined as a TEAE with a relationship to study treatment as 'possibly related' or 'probably related' to study treatment. TEAEs with a missing or 'unknown' relationship to study treatment will be regarded as 'related' to study treatment. If a participant reports the same AE as both related and unrelated within that SOC/ PT, the related AE will be used in the corresponding relationship summaries.

8.7.1.2. TEAEs LEADING TO DISCONTINUATION OF STUDY TREATMENT

TEAEs leading to permanent discontinuation of study treatment will be identified by using the 'Action taken with study treatment' variable equal to 'Drug withdrawal' from the AE page of the eCRF.
A summary of TEAEs leading to discontinuation of study treatment by SOC and PT will be prepared. A listing of all TEAEs leading to discontinuation of study treatment will be provided, along with the details as to whether the event was related or serious.

8.7.1.3. TEAEs LEADING TO DISCONTINUATION FROM THE STUDY

TEAEs leading to discontinuation from the study will be identified by using the 'Did the AE

cause the participant to discontinue from the study?' variable equal to 'Yes' from the AE page of the eCRF.

A summary of TEAEs leading to discontinuation from the study by SOC and PT will be prepared. A listing of all TEAEs leading to discontinuation from the study will be provided, along with the details as to whether the event was related or serious.

8.7.1.4. Serious Adverse Events

Serious adverse events (SAEs) are those events recorded as 'Serious' on the Adverse Events page of the eCRF. A summary of serious TEAEs by SOC and PT will be prepared.

8.7.1.5. Adverse Events Leading to Death

TEAEs leading to Death are those events which are recorded as 'Fatal' on the Adverse Events page of the eCRF. A summary and listing of TEAEs leading to death by SOC and PT will be prepared.

8.7.1.6. Adverse Events Of Special Interest

Adverse Event of Special Interest (AESI), as defined in section 8.4.6 of the protocol, are those events recorded as 'AESI' on the Adverse Events page of the eCRF.

AESIs include haemolysis and methemoglobinemia.

Haemolysis is defined as at least three of the following four laboratory parameter criteria fulfilled (high and low refers to the upper and lower reference limits from the central lab):

a) Low haemoglobin OR haematocrit
b) Low haptoglobin
c) High reticulocyte count
d) High indirect (unconjugated) bilirubin

Methemoglobinemia is defined as both the following two criteria fulfilled:

a) Methemoglobin reading by pulse oximetry >2% and
b) Methemoglobin by blood test > ULN

Note: If pulse oximeter is not available or can't be used, the blood test value will be used alone to meet AESI reporting criteria.

A summary of AESIs by SOC and PT will be prepared. The proportion of participants with prespecified events listed as AESIs during the on-treatment period will be provided. All AESIs will be included in the summary table, even if the incidence is zero.

8.7.2. Laboratory Evaluations

Quantitative laboratory measurements reported as '< X', i.e., below the lower limit of normal (LLN), or '> X', i.e., above the upper limit of normal (ULN), will be converted to X for the purpose of quantitative summaries, but will be presented as recorded, i.e., as '< X' or '> X' in the listings. If local and central laboratory assessments are available from the same day, central laboratory assessments will be used in the summary.

The following summaries will be provided for laboratory data:

- Observed values and change from baseline by visit for Chemistry, Hematology and Urinalysis (quantitative measurements)
- Observed values for categorical measurements
- Shift from baseline value to each post-baseline visit value according to normal range criteria (quantitative measurements and categorical measurements)

8.7.2.1. Laboratory Reference Ranges

Quantitative laboratory measurements will be compared with the relevant laboratory reference ranges in SI units and categorized as:

- Low: Below the lower limit of the laboratory reference range.
- Normal: Within the laboratory reference range (upper and lower limit included).
- High: Above the upper limit of the laboratory reference range.

8.7.2.2. Liver Function Related Criteria

Summary details for liver enzyme abnormalities will be presented by treatment group and assessed by using post-baseline results for Alanine Aminotransferase (ALT), Aspartate Aminotransferase (AST) and Total Bilirubin (TBL) given an upper limit of normal (ULN):

- ALT or AST >8 × ULN
- ALT or AST >5 × ULN
- ALT or AST >3 × ULN and TBL >2 × ULN or international normalized ratio >1.5
- ALT or AST >3 × ULN

8.7.3. ECG Evaluations

The following ECGs measurements will be reported for this study:

- Axis (degree)
- Heart Rate [beats per minute (bpm)]
- PR Interval (msec)
- QT Interval (msec)
- QTcB Interval (msec)
- QRS Interval (msec)
- RR interval (msec)
- QTcF Interval (msec)
- Overall interpretation of ECG:
  - Normal
  - Abnormal, Not Clinically Significant (Abnormal NCS)
  - Abnormal, Clinically Significant (Abnormal CS)

Triplicate standard 12-lead ECGs will be obtained after the participant has been in a supine position for at least 5 minutes. ECG measurements will be summarized by taking the average of the available assessments.

The following summaries will be provided for ECGs data:

- Observed values and change from baseline by visit and timepoint
- Shift in normal/abnormal NCS/abnormal CS in the overall interpretation from baseline to each timepoint

The number of participants with absolute QTcF values CCI [redacted], and the number of participants with change from baseline in QTcF, CCI [redacted] during

CCI [redacted]

treatment will be summarized by treatment group. QTcF prolongation is defined as CCI or a change CCI.

A scatter plot will be presented, showing baseline QTcF on the x-axis, post-baseline QTcF on the y-axis (one panel for each treatment), and diagonal lines indicating changes of -60, -30, 0, +30, and +60. Different symbols will be used to identify the different post-treatment timepoints.

8.7.4. Vital Signs

The following Vital Signs measurements will be reported for this study:

- Systolic Blood Pressure (mmHg)
- Diastolic Blood Pressure (mmHg)
- Heart Rate (bpm)
- Respiratory rate (breaths/min)
- Body temperature (ºC)
- Weight (kg)
- BMI (kg/m^2)
    - Observed values and changes from baseline will be presented by visit and treatment group.

8.7.5. COVID-19 Impact

To assess the impact that COVID-19 had on the trial, a listing of comments related to COVID-19 will be presented. This listing will also include information about visits and by-visit assessments that were impacted by COVID-19.
Additionally, a listing of participants who discontinued the study due to COVID-19, including participants who contracted the disease or discontinued for any other reason related to COVID-19, will be presented along with the specified text associated with discontinuation.

8.7.6. Other Safety Assessments

8.7.6.1. VAGINAL BLEEDING PATTERN

Participants records the presence of bleeding or spotting CCI in the eDiary. The CCI questions to assess bleeding pattern in the eDiary are "1a: During the past 24 hours, did you have any vaginal bleeding or spotting? If the response is "Yes", then the next question is, "1b: During the past 24 hours, have you been on your period?".

The following endpoints will be calculated:

- Mean change from BC to TRC1, TRC2, TRC3 in the number of days on period
- Mean change from BC to TRC1, TRC2, TRC3 in the number of days with irregular vaginal bleeding/spotting (i.e. days when question 1a was answered yes and question 1b was answered no)

The definition of cycles for the above endpoints will be different from the definition used for efficacy endpoints. For the vaginal bleeding endpoints, cycles will be defined based on start of menses dates only as below:

- Screening cycle (SC) – starts at the start of menses associated with CCI and ends the day before the start of menses associated with CCI
- Baseline cycle (BC) – starts at the start of menses associated with CCI and ends at the day before the start of menses associated with CCI
- Treatment cycle 1 (TRC1) – starts at start of menses associated with visit 4 and ends the day before the start of menses associated with phone contact C (PC1).
- Treatment cycle 2 (TRC2) – starts at the start of menses associated with phone contact C and ends the day before the start of menses associated with CCI
- Treatment cycle 3 (TRC3) – starts at the start of menses associated with CCI and ends the day before the start of menses associated with CCI
- Safety follow-up cycle (SFC) – starts at the start of menses associated with CCI and ends the day before the start of menses associated with CCI

The following summaries will be provided for vaginal bleeding pattern:

- Observed values and change from BC in the number of days on period by treatment cycle

CCI

and by treatment group

- Observed values and change from BC in the number of days with irregular vaginal bleeding/spotting by treatment cycle and by treatment group

8.7.6.2. Endometrial Assessment: Tissue collection and ultrasound

A Transvaginal Ultrasound or Transabdominal (Pelvic) Ultrasound (TVUS) will be performed at CC/Screening and CCI / End of TRC3 to assess endometrial thickness and if there is a clinically significant changes in endometrial thickness and histology. Endometrial biopsies are optional for all participants except in the case of abnormal vaginal bleeding (at discretion of the Investigator).

- The following variables will be assessed.
    - Endometrial thickness (mm)
    - Proportion of participants with clinically significant changes from CCI
        - i) In endometrial thickness as assessed by transvaginal ultrasound and
        - ii) Histological findings as assessed by optional endometriosis biopsy (using an endometrial suction curette Pipelle®

The endometrial thickness will be summarized by treatment group and visit using descriptive statistics of observed data and change from baseline values. Similar analysis for the primary efficacy variable described in Section 8.3.7 using ANCOVA will be performed for the change from baseline to CCI /EOT in endometrial thickness.

The number and percentage of participants with clinically significant changes in endometrial thickness and histology from CCI will be summarized.

8.7.6.3. Pregnancy Test and Lactation

Urine pregnancy test (human chorionic gonadotropin) and/or serum β-hCG pregnancy tests are performed throughout the study in female subjects of childbearing potential.
Pregnancy outcomes and lactation information collected on eCRF Pregnancy/Lactation page will be presented in a listing.

8.7.6.4. Columbia Suicide Severity Rating Scale (C-SSRS)

A positive C-SSRS response is defined when there is a positive answer ("Yes"), at any time during treatment, to any one of the questions (suicidal ideation / behavior) on the C-SSRS questionnaire.

Positive C-SSRS responses observed at any post-baseline assessment will be summarized by treatment group for the SAF using descriptive statistics for categories of Suicidal Ideation, Suicidal Behavior and Self-Injurious Behavior without Suicidal Intent, as well as for the 11 individual questions associated with each category.

8.7.6.5. Physical, Gynecological and Breast Examination

Physical, gynecological and breast examination findings after ICF are recorded in the eCRF AE page and will be presented in the AE listing. A listing of gynecological examination assessments in participants who had abnormal PAP smear results will be provided.

8.7.6.6. Bone Biomarker

The following bone biomarker measurements will be reported for this study:

CCI [redacted]

The following summaries will be provided for bone biomarker data:

- Observed and change from baseline (CCI) to end of treatment (CCI)

9. Pharmacokinetic Analysis

Plasma concentrations of OG-6219 and FOR-1011 will be summarized for the PK analysis set and for the PK analysis set for NCA using descriptive statistics.

Calculation of pharmacokinetic parameters will be performed by IQVIA and the methods of calculation are described in a separate document.

Calculated pharmacokinetic parameters for OG-6219 and FOR-1011 will be summarized for the PK analysis set for NCA using descriptive statistics. In addition to the standard descriptive statistics described in section 8.1.1, geometric mean and coefficient of variation

(CV) will be included for all PK parameters except t_{max} and t_{last}.

Individual and mean plasma concentration data will be illustrated using both a linear and a logarithmic concentration scale.

10. CHANGES TO PROTOCOL-PLANNED ANALYSES

There were no changes to the analyses planned in the study protocol.

The study protocol did not include a description of the estimand. This has now been included in the SAP.

11. REFERENCES

PROMIS_Fatigue_Scoring_Manual.pdf (healthmeasures.net)

OG-6219-P001_CSR_SAP_V2_09JUN2025

Final Audit Report 2025-06-10

Created:	2025-06-09
By:	PPD
Status:	Signed
Transaction ID:	PPD

"OG-6219-P001_CSR_SAP_V2_09JUN2025" History

Document created by PPD
2025-06-09 - 12:49:40 PM GMT- IP address: PPD

Document emailed to PPD for signature
2025-06-09 - 12:52:49 PM GMT

Email viewed by PPD
2025-06-09 - 1:02:28 PM GMT- IP address: PPD

PPD authenticated with Adobe Acrobat Sign.
Challenge: The user opened the agreement.
2025-06-09 - 1:03:02 PM GMT

PPD authenticated with Adobe Acrobat Sign.
Challenge: The user completed the signing ceremony.
2025-06-09 - 1:04:14 PM GMT

Document e-signed by PPD
Signing reason: Reviewed
Signature Date: 2025-06-09 - 1:04:15 PM GMT - Time Source: server- IP address: PPD

Document emailed to PPD for signature
2025-06-09 - 1:04:17 PM GMT

Email viewed by PPD
2025-06-09 - 3:43:38 PM GMT- IP address: PPD

PPD authenticated with Adobe Acrobat Sign.
Challenge: The user opened the agreement.
2025-06-09 - 3:44:46 PM GMT

PPD authenticated with Adobe Acrobat Sign.

Challenge: The user completed the signing ceremony.

2025-06-09 - 5:06:55 PM GMT

Document e-signed by PPD

Signing reason: Approved

Signature Date: 2025-06-09 - 5:06:56 PM GMT - Time Source: server- IP address: PPD

Document emailed to PPD for signature

2025-06-09 - 5:06:58 PM GMT

Email viewed by PPD

2025-06-10 - 6:32:43 AM GMT- IP address: PPD

PPD authenticated with Adobe Acrobat Sign.

Challenge: The user opened the agreement.

2025-06-10 - 6:33:24 AM GMT

PPD authenticated with Adobe Acrobat Sign.

Challenge: The user completed the signing ceremony.

2025-06-10 - 6:34:17 AM GMT

Document e-signed by PPD

Signing reason: Author

Signature Date: 2025-06-10 - 6:34:19 AM GMT - Time Source: server- IP address: PPD

Agreement completed.

2025-06-10 - 6:34:19 AM GMT