Official Title: **Efficacy and Safety of Oral OPS-2071 in Subjects With Crohn's Disease Showing Symptoms of Active Inflammation Despite Ongoing Treatment**

NCT Number: NCT03850509

Document Date: SAP Version 1.0: 09 June 2020

Otsuka Pharmaceutical Development & Commercialization, Inc.

Investigational New Drug OPS-2071

Protocol No. 341-201-00004
IND No. 137377
EudraCT No. 2019-000176-41

A Phase 2, Multicenter, Randomized, Double-blind, Placebo-controlled, Proof of Concept Trial to Assess the Efficacy and Safety of Orally Administered OPS 2071 for 12 Weeks in Subjects With Crohn's Disease Showing Symptoms of Active Inflammation Despite Ongoing Treatment

Statistical Analysis Plan

Version: 1.0
Date: 09 Jun 2020
Protocol Version 1 Date: 05 Feb 2019
Protocol Amendment 1: 12 Apr 2019
Protocol Amendment 2: 09 Dec 2019
Protocol Amendment 3: 22 Jan 2020

Table of Contents

List of Appendices

List of Abbreviations and Definition of Terms

Abbreviation	Definition
5-ASA	5-aminosalicylic acid
AE	Adverse event
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
AUC_{24h}	Area under the concentration-time curve from time zero to 24 hours
AUC_{∞}	Area under the concentration-time curve from time zero extrapolated to infinity
AUC_t	Area under the concentration-time curve from time zero to the last quantifiable concentration
BCRP	Breast cancer resistance protein
BID	Twice a day
CDAI	Crohn's Disease Activity Index
C_{max}	Maximum (peak) plasma concentration of the drug
CMV	Cytomegalovirus
CRO	Clinical Research Organization
CRP	C-reactive protein
DNA	Deoxyribonucleic acid
ECG	Electrocardiogram
FBR	Future biospecimen research
FCP	Fecal calprotectin
FDA	Food and Drug Administration
FSH	Follicle-stimulating hormone
GCP	Good Clinical Practice
HIV	Human immunodeficiency virus
IB	Investigator's brochure
IBD	Inflammatory bowel disease
ICF	Informed consent form
ICH	International Council on Harmonisation
ICMJE	International Committee of Medical Journal Editors
IDMC	Independent Data Monitoring Committee
IEC	Independent ethics committee
IMP	Investigational medicinal product
IND	Investigational new drug
IRB	Institutional review board
IRE	Immediately reportable event
ITT	Intent-to-treat
LPS	Lipopolysaccharide
NOAEL	No observed adverse effect level
O&P	Ova and parasite
PD	Pharmacodynamic

1 Introduction

This statistical analysis plan (SAP) documents the statistical methodology and data analysis algorithms and conventions to be applied for statistical analysis and reporting of efficacy and safety data of Trial 341-201-00004. All amendments to the protocol are taken into consideration in developing this SAP.

2 Trial Objectives

The primary objective of the trial is to investigate the therapeutic effect of OPS-2071 (150, 300, or 600 mg BID) add-on therapy administered orally for 12 weeks in subjects with Crohn's disease showing symptoms of active inflammation despite ongoing treatment.

The secondary objective of the trial is to determine the safety and tolerability of OPS-2071 (150, 300, or 600 mg BID) add-on therapy administered orally for 12 weeks in subjects with Crohn's disease showing symptoms of active inflammation despite ongoing treatment.

3 Trial Design

3.1 Design of Trial

This is a phase 2, multicenter, randomized, double-blind, placebo-controlled, POC trial to assess the efficacy and safety of 3 doses of OPS-2071 (150, 300, or 600 mg BID) in subjects with Crohn's disease showing symptoms of active inflammation despite ongoing treatment. OPS-2071 will be added to ongoing treatment for Crohn's disease, with some restrictions.

Screening will occur up to 28 days prior to randomization. At the start of the screening period, a fecal sample will be collected for stool microbiology testing (eg, for culture, C difficile, ova and parasites [O&P]) and to exclude any type of infectious colitis, and to assess the fecal microbiota. During the screening period, subjects will be dispensed an electronic diary to complete to collect data for the baseline CDAI assessment. After completion of the diary for 7 days, subjects will undergo bowel preparation for 3 days (consisting of 2 days of dietary modifications and 1 day of cleansing bowel preparation) and then an ileocolonoscopy with biopsy will occur (within 7 to 10 days prior to Day 1). The ileocolonoscopy will be centrally read to assess inflammation in order to confirm trial eligibility. Optional cytomegalovirus (CMV) testing can be done on the biopsy specimen, if clinically indicated.

Eligible subjects will be randomized on Day 1 in a 1:1:1:1 ratio to 1 of 4 treatments: placebo BID, OPS-2071 150 mg BID, OPS-2071 300 mg BID, or OPS-2071 600 mg BID. The randomization will be stratified by CDAI score (CDAI 180 - 300 and CDAI > 300 450) and concomitant Crohn's disease medication status (subjects who are treated with anti TNF-α monoclonal antibody, subjects who are treated with low-dose steroids, and subjects who are treated with other Crohn's disease medications). The IMP will be administered twice a day, 8 to 12 hours apart, starting on Day 1 for a total of 12 weeks. Blood samples for PK analysis will be collected throughout the 12 week treatment period according to the schedule of assessments (Table 3.7-1). Efficacy will be assessed using the CDAI, Simple Endoscopic Score for Crohn's Disease (SES-CD), and the two-item patient reported outcome (PRO-2). Definitions of remission and nonremission subjects (ie, treatment failures) can be found in Section 7.4.1. A 30 day (post treatment) follow-up assessment will be performed for all subjects; the CDAI score will be calculated, inflammatory biomarkers and microbiota flora will be checked, and safety measures will be performed. Subjects who increase dosages of baseline Crohn's disease medications or take a new therapy for Crohn's disease during the trial are considered as treatment failures (nonresponders). These subjects are encouraged to stay in the trial for continuous data collections.

A blood sample for future biospecimen research (FBR) will be taken from subjects who consent to the collection of this sample.

3.2 Trial Treatments

Eligible subjects will be randomized on Day 1 to 1 of 4 treatments: placebo BID, OPS-2071 150 mg BID, OPS-2071 300 mg BID, or OPS-2071 600 mg BID. The IMP will be supplied as 150-mg and 300-mg OPS-2071 tablets, and matching placebo tablets. The IMP will be administered as oral tablets, BID (8 to 12 hours apart), starting on Day 1 for a total of 12 weeks. The timing of the first dose of IMP will be at the subject's or trial site's discretion. Each subject will take 6 tablets daily (to maintain blinding); 3 tablets for the morning dose and 3 tablets 8 to 12 hours later. The IMP will be administered with up to 240 mL (8 fluid ounces) of still water, at least 1 hour before or after a meal. Water intake will be restricted for 1 hour before and after dosing, except as part of the dosing procedure.

3.3 Trial Population

Approximately 240 male and female subjects between the ages of 18 and 70 (inclusive), with Crohn's disease and showing symptoms of active inflammation despite ongoing

treatment, will be randomized. Active mucosal inflammation will be confirmed by centrally read ileocolonoscopy. Discontinued or withdrawn subjects will not be replaced.

4 Sample Size

For the primary endpoint of clinical remission (CDAI score < 150), assuming a 40% response rate in the OPS-2071 600 mg BID and 300 mg BID dose groups and a 20% response rate in the placebo group, 60 randomized subjects per treatment group would provide 80% power to detect a treatment difference between the OPS-2071 pooled dose groups (600 mg BID and 300 mg BID) and the placebo group in the primary endpoint, using a 2-sided alpha of 0.05. The total sample size of the trial is 240 randomized subjects with 1:1:1:1 randomization to OPS-2071 600 mg BID, 300 mg BID, or 150 mg BID, or placebo BID. The power of individual dose comparison with placebo under the same assumption is 67%.

For the secondary endpoint of endoscopic response, assuming a 26.5% response rate in the OPS-2071 600 mg BID and 300 mg BID dose groups and a 10% response rate in the placebo group, 60 subjects per treatment group would provide 80% power for the comparison between the OPS-2071 pooled dose groups (600 mg BID and 300 mg BID) and the placebo group, and 65% power for the individual dose comparison with placebo, using a 2-sided alpha of 0.05.

5 Statistical Analysis Datasets

5.1 Analysis Datasets

The following samples are defined for this trial:

- Randomized: consists of all subjects who are randomized into this trial.
- Safety Analysis Set (SAS): consists of all subjects who are administered at least one dose of IMP.
- Full Analysis Set (FSA): consists of all subjects who are in the Randomized Sample, take at least one dose of IMP, and have a baseline and at least one postbaseline evaluation of the CDAI score.

In general, baseline of an efficacy endpoint is defined as the last observation of the endpoint before the subject takes the first dose of IMP.

The core dataset for all efficacy analyses is based on the intent-to-treat (ITT) population, which will consist of all randomized subjects who take at least one dose of IMP. The Full Analysis Set, which is identical to the Efficacy Sample, will also be used in the efficacy analyses. As described below, in order to handle missing data and restrictions imposed by

different types of analyses (eg, change from baseline analysis), other datasets derived from the ITT population will be used for the efficacy analyses.

5.2 Handling of Missing Data

The estimand defining the treatment effect of interest in the protocol uses a mixture of composite strategy and treatment policy strategy specified in the ICH E9 addendum. Subjects who increase dosages of baseline Crohn's disease medications or take a new therapy for Crohn's disease during the trial are considered as treatment failures (non-responders) in the ITT analysis of the primary and secondary endpoints, while these subjects are encouraged to stay in the trial for continuous data collection. These data collected after subjects change their baseline Crohn's disease medication will only be provided in data listings, but not used in the primary and secondary efficacy analyses.In addition, subjects who withdraw from the trial early are considered as treatment failures (non-responders) in the ITT analysis of all of the primary and secondary endpoints. Subjects who discontinue IMP during the trial with no baseline Crohn's disease medication dosage increases and/or new therapy for Crohn's disease through Week 12 are encouraged to stay in the trial and take primary and secondary endpoint evaluations at Week 12 for efficacy analyses. Hence, missing data at Week 12 due to subjects' change of baseline Crohn's disease medication and early withdrawal is handled by an ITT analysis with all these subjects treated as treatment failures(nonresponders).

For CDAI-related endpoints, if a subject has a Week 12 evaluation but a component of the score is missing, the value of the component in the most recent previous visit will be used for the calculation of the score and the endpoint if no more than 2 components of the CDAI score are missing. For a CDAI assessment, if more than 2 components of the score are missing, the CDAI score will be set to missing. This approach is conservative since subjects randomized to OPS-2071 treatment groups are expected to have their Crohn's disease symptoms improved over time. No component is allowed to be missing in a PRO-related endpoint. The abdominal pain and stool frequency scores of PRO-2 are averaged over a 7-day period from the subject's diary.

6 Study Conduct

6.1 Subject Disposition

The number of subjects who have been randomized, the number of subjects who are treated, and the number of subjects who discontinue from the trial, together with the reasons for discontinuation taken from the eCRF status page will be provided.

6.2 Prior and Concomitant Medications

The proportion of subjects taking concomitant medications will be tabulated for all randomized sample for prior to, during and after the trial medication.

6.3 Protocol Deviations

Protocol deviations data will be summarized by type of deviations (eg, deviations in entry criteria, dosing, concomitant medication, procedural, etc) by trial center. In addition, a subject listing will be provided describing the deviations for each subject.

7 Demographic and Baseline Characteristics

Demographics, disease severity, baseline characteristics, and medical history at (predose) baseline will be summarized by descriptive statistics; eg, proportion, mean, median, standard deviation, and minimum and maximum values. These summary statistics will be reviewed to identify any potential lack of balance between the treatment groups.

8 Efficacy Analysis

8.1 Primary Efficacy Endpoint

The primary efficacy endpoint is percentage of subjects who, at Week 12, achieve clinical remission (defined as a CDAI score < 150).

8.1.1 Primary Estimand

The primary objective is to investigate the therapeutic effect of OPS-2071 add-on therapy administered orally for 12 weeks in subjects with Crohn's disease showing symptoms of active inflammation despite ongoing treatment. The primary estimand defining the treatment effect of interest in the protocol uses a mixture of composite strategy and treatment policy strategy specified in the ICH E9 addendum. [2] Two types of intercurrent events are accounted for by these two strategies:

- Increasing dosages of baseline Crohn's disease medications or taking a new therapy for Crohn's disease: The occurrence of this intercurrent event may provide meaningful information on the treatment effect of interest such as lack of efficacy, and will be taken to be a component of the variable for the primary estimand. Following the composite strategy, the primary estimand assess the treatment effect based on a composite variable which combines percentage of subjects achieve clinical remission at week 12 and the intercurrent event of the changes of the medications for Crohn's disease. The subjects with this type of intercurrent event will be considered as treatment failures.

- Early withdrawal from the trial: Interest lies in the treatment effect regardless of this intercurrent event. Following treatment-policy strategy, occurrence of the early withdraw intercurrent event is ignored, and the primary estimand is then the difference between treatment groups in percentage of subjects who achieve clinical remission at week 12, regardless of whether or not early withdraw had occurred.

The primary estimand for this trial is defined by the following components:

- Population: subjects with Crohn's disease and showing symptoms of active inflammation despite ongoing treatment who have met the protocol-defined inclusion/exclusion Criteria
- Endpoint: the percentage of subjects who, at Week 12, achieve clinical remission (defined as a CDAI score < 150)
- Intercurrent Events: increase dosages of baseline Crohn's disease medications or take a new therapy for Crohn's disease during the trial, and early withdraw from the trial
- Population-level summary: difference in clinical remission proportions between OPS2071 and placebo

Subjects who increase dosages of baseline Crohn's disease medications or take a new therapy for Crohn's disease during the trial and subjects who withdraw from the trial early are considered as treatment failures in the ITT analysis of all of the primary and secondary response/remission endpoints. Subjects who discontinue IMP during the trial with no baseline Crohn's disease medication dosage increases and/or new therapy for Crohn's disease through Week 12 are encouraged to stay in the trial and take primary and secondary endpoint evaluations at Week 12 for efficacy analyses. Hence, missing data at Week 12 due to subjects' early withdrawal is handled by an ITT analysis with all subjects who early withdraw from the trial treated as treatment failures.

8.1.2 Primary Efficacy Analysis

Logistic regression with terms of treatment group, the randomization stratification factors, location of disease, country, age, sex, disease duration, smoking habit, and baseline CDAI, and baseline CDAI with CDAI stratification factor interaction will be applied to the primary endpoint. The FAS will be used in the efficacy analysis.

The primary analysis of this trial is comparison of the pooled OPS-2071 600 mg BID and 300 mg BID treatment groups with placebo in the primary endpoint of clinical remission.

If this analysis is significant at a 2-sided alpha level of 0.05, comparisons to OPS-2071 600 mg BID versus placebo and OPS-2071 300 mg BID versus placebo in the primary endpoint will be conducted with a 2-sided alpha level of 0.05 in each comparison. If the comparison of the pooled OPS-2071 dose groups versus placebo in the primary endpoint is significant at a 2-sided alpha level of 0.05, and a comparison, say, OPS-2071 600 mg BID versus placebo, is also significant at a 2-sided alpha level of 0.05, the null hypothesis of equal response of OPS-2071 600 mg BID and placebo in the primary endpoint will be considered rejected with an experiment-wise Type I error of 0.05.

The comparison between OPS-2071 150 mg BID versus placebo in the primary endpoint is considered as a secondary analysis and the details are provided in section 8.4.

8.1.3 Technical Computational Details for Primary Efficacy Analysis

1) CDAI score is the sum of 8 components score multiplying by each component's weight factor. The components of CDAI include: number of liquid or soft stools, abdominal pain, general well-being, presence of complications, taking Lomotil or opiates for diarrhea, presence of an abdominal mass, hematocrit, percentage deviation from standard weight. The formula for CDAI score calculation is: sum of number of liquid or soft stools over a week x 2 + sum of abdominal pain over a week x 5+ sum of general well-being over a week x 7 + presence of complications x 20 + taking Lomotil or opiates for diarrhea x 30 + presence of an abdominal mass x 10 +6 point per percent deviation of hematocrit from 47% in male and 42% in female+ percentage below standard weight x 1. One point each is added for each set of complications which include the presence of joint pains (arthralgia) or frank arthritis; inflammation of the iris or uveitis; presence of erythema nodosum, pyoderma gangrenosum, or aphthous ulcers; anal fissures, fistulae or abscesses; other fistulae; fever during the previous week.
2) Remission and none remission subjects (ie, treatment failures) are defined as follows:
    - Subjects who complete the trial and meet the primary endpoint definition at Week 12 without increasing dosages of baseline Crohn's disease medications or taking a new medication for Crohn's disease during the double blind treatment period will be considered as remission subjects for the endpoint.
    - Subjects who increase dosages of baseline Crohn's disease medications or take a new medication for Crohn's disease during the double blind treatment period, or subjects who withdraw from the trial during the

double blind period, in addition to the subjects who fail to meet the primary endpoint at Week 12, will be considered as nonremission subjects for the endpoint.

3) If a subject has a Week 12 evaluation but a component of the score is missing, the value of the component in the most recent previous visit will be used for the calculation of the score and the endpoint if no more than 2 components of the CDAI score are missing. CDAI score will be set to be missing if more than 2 components' scores are missing.
4) The 3 components of CDAI including liquid or soft stools, abdominal pain and general well-being are collected from the subjects' eDiary. The sum of the scores of the last 7 day's data will be used for each component. If a subject has 1 or 2 days' missing data, the 7 day's score will be calculated by the average of 6 or 5 days' data multiplying 7. If more than 2 days' data are missing, the components' score will set to be missing.
5) Clinical remission is defined as CDAI score < 150 for remission subjects as described in session 8.1.1.

8.1.4 Sensitivity Analyses

8.1.4.1 Placebo-based Pattern Imputation

A sensitivity analysis of the primary endpoint of clinical remission of missing data will be provided for subjects who withdraw from the trial before the Week 12 visit , or subjects who increase dosages of baseline Crohn's disease medications or take a new medication for Crohn's disease during the double-blind treatment period. The missing data under this scenario is considered as Missing Not at Random (MNAR) and will be imputed. A pattern-mixture model approach will be utilized to model the distribution of the observed responses and the distribution of the missing responses using a placebo-based multiple imputation(MI). Imputation of missing CDAI scores will be produced using a regression approach[1] and the imputed CDAI remission status will be determined by checking if the imputed CDAI score is less than 150.

Specifically, the placebo-based pattern imputation procedure follows these steps:

1) Using Monte Carlo Markov Chain (MCMC) methodology from PROC MI by treatment group to impute the intermittent missing data to a monotone missing pattern;

2) Using pattern-mixture model approach to impute the monotone missing data. By applying MNAR statement of PROC MI in SAS, missing values under MNAR will be imputed. Only the observations in the placebo control group will be used to impute the missing values by using the MODEL option in the MNAR statement. The imputed values for the 3 treatment groups of OPS-2071will be adjusted using the shift parameter. When the shift parameter is 0, the MI procedure would produce an analysis which is essentially something called “back to placebo”;
3) For subjects with imputed CDAI scores, CDAI remission status will be determined by checking if CDAI score < 150;
4) Using logistic regression model in the primary analysis to analyze the completed data along with the imputed data;
5) Obtaining the overall results using PROC MIANALYZE.

8.1.5 Supplemental Analysis

As another sensitivity analysis, the Cochran-Mantel-Haenszel (CMH) test stratified by the randomization stratification factors will be applied to the primary endpoint, and treatment comparisons between individual OPS-2071 dose groups vs. placebo will be conducted.

8.1.6 Subgroup Analysis

Subgroup analysis of the percentage of subjects who achieve clinical remission at week 12 will be conducted by gender, race (Caucasian and Non-Caucasian), age (18 to < 40, 40 to < 50, and at least 50), and region (US and non-US) using logistic regression with terms of treatment group, the randomization stratification factors, location of disease, disease duration, smoking habit, and baseline CDAI, and baseline CDAI with CDAI stratification factor interaction.

8.2 Secondary Efficacy Endpoints

8.2.1 Secondary Efficacy Analyses

If all both null hypotheses of OPS-2071 600 mg BID, 300 mg BIDversus placebo in the primary endpoint of clinical remission are rejected with an experiment-wise Type I error of 0.05 by the test procedure provided in session 8.1.1, a formal test of the secondary efficacy endpoint of endoscopic response (defined as a reduction of the SES-CD by at least 50%) for comparisons of OPS-2071 600 mg BID versus placebo and 300 mg BID versus placebo will be conducted, using a similar procedure used for the primary

endpoint of clinical remission except replacing baseline CDAI with baseline SES-CD score.

The other secondary efficacy endpoints of this trial are listed below by the order of the gatekeeping test procedure:

2) Change from baseline in the SES-CD score.
3) The percentage of subjects with PRO-2 remission (defined as stool frequency ≤ 3 times per day and abdominal pain ≤ 1) at Week 12.
4) The percentage of subjects with clinical response, defined as at least a 25% decrease in the CDAI score, at Week 12.
5) The percentage of subjects who, at Week 12, achieve endoscopic remission (defined as a SES-CD score of 0 to 2; or a score of 0 to 4, with no individual subscore greater than 1).
6) The percentage of subjects with a decrease from baseline of ≥ 100 points in the CDAI score.

Analysis of secondary endpoint #2 will be conducted using a procedure similar to the analyses of the secondary endpoint of endoscopic response, except the logistic regression will be replaced by an analysis of covariance (ANCOVA). The model used in the analysis of the primary endpoint will be applied to secondary endpoints #3, #4, and #6, while the model used in the analysis of the secondary endpoint of endoscopic response will be applied to the analysis of secondary endpoint #5. The multiplicity handling along with the gatekeeping to protect the experiment wise Type I error in the primary and secondary analyses are detailed in section 8.4.

A sensitivity analysis of secondary endpoint #3 (PRO-2 remission) of missing data will also be provided similar to the sensitivity analysis of the primary endpoint of CDAI clinical remission. Specifically, missing data for PRO-2 remission will be imputed for the subjects who withdraw from the trial before the Week 12 visit without increasing dosages of baseline Crohn's disease medications or taking a new medication for Crohn's disease during the double-blind treatment period. The missing data under this scenario is considered as MNAR and will be imputed by applying a regression approach[1] to the abdominal pain and stool frequency scores, averaged over a 7-day period. In detail, the placebo-based pattern imputation procedure follows these steps:

1) Using Monte Carlo Markov Chain (MCMC) methodology from PROC MI by treatment group to impute the intermittent missing data to a monotone missing pattern;
2) Using pattern-mixture model approach to impute the monotone missing data. By applying MNAR statement of PROC MI in SAS, missing values under MNAR will be imputed. Only the observations in the placebo control group will be used

to impute the missing values by using the MODEL option in the MNAR statement. The imputed values for the 3 treatment groups of OPS-2071will be adjusted using the shift parameter; When the shift parameter is 0, the MI procedure would produce an analysis which is essentially something called “back to placebo”;

3) The PRO-2 remission status for the subjects with the imputed data will be determined by checking if the imputed stool frequency score is ≤ 3 and the imputed abdominal pain score is ≤ 1;
4) Using logistic regression model in the primary analysis to analyze the completed data along with the imputed data;
5) Obtaining the overall results using PROC MIANALYZE.

8.2.2 Technical Computational Details for Secondary Efficacy Analyses

1) Subjects who increase dosages of baseline Crohn’s disease medications or take a new therapy for Crohn’s disease during the trial should be discontinued from the trial, and subjects who withdraw from the trial early are considered as treatment failures in the ITT analysis of the secondary endpoints. These subjects will be considered as non-remission and non-responders for the remission and response endpoints including endoscopic response, PRO-2 remission, clinical response, endoscopic remission.
2) For the PRO-2 score used to derive PRO-2 remission, the abdominal pain score is reported as the daily worst abdominal pain (0 to 3) score and the stool frequency score is recorded as the number of loose or watery bowel movements (Bristol stool form scale 6 or 7) per day. The average score of the last 7 day’s data will be used for each component. If a subject has 1 or 2 days’ missing data, the 7 day’s score will be calculated by the average of 6 or 5 days’ data multiplying 7. If more than 2 days’ data are missing, the components’ score will set to be missing.
3) No component is allowed to be missing in a PRO related endpoint. PRO-2 remission will be set to missing if either the abdominal pain score or the stool frequency score is missing.

8.3 Exploratory Endpoint

8.3.1 Exploratory Endpoint Analyses

The CDAI score and PRO-2 score will be evaluated at baseline and at Weeks 2, 4, 6, 8, 10, and 12 in this trial, so that the by-visit analysis of CDAI remission and PRO-2 remission (exploratory endpoints #1 and #2) at Weeks 2, 4, 6, 8, and 10, in addition to Week 12, will be provided and analyzed using a similar method with Week 12 replaced by Weeks 2, 4, 6, 8 or 10. Since the by-visit analysis of an endpoint is considered as supplemental to the analysis of the endpoint at Week 12, no adjustment in alpha is necessary.

Additional exploratory endpoints are as follows:

3) The percentage of subjects with a biological response in subjects with elevated levels of CRP, defined as a reduction in the level of CRP to within the normal range or by at least 15%, at Week 6, Week 12, and at the 30-day follow-up.
4) The percentage of subjects with a biological response in subjects with elevated levels of FCP, defined as a reduction in the level of FCP to within the normal range or by at least 15%, at Week 6, Week 12, and at the 30-day follow-up.
5) Characterization of the fecal microbiota profile and abundance based on taxonomic and functional annotation from next-generation sequencing, change from baseline, and association with phenotype, at Week 2, Week 12, and at the 30-day follow-up.

CMH test stratified by the randomization stratification factors will be applied to the exploratory endpoint #3 and #4. The mean change from baseline will be provided by visit for the additional exploratory endpoints#5. An analysis of covariance (ANCOVA) with baseline value as covariate and treatment as main effects will be applied to this endpoint.

8.3.2 Technical Computational Details for Exploratory Endpoint Analyses

To do the by visit analysis of exploratory endpoints #1 and #2, CDAI remission and PRO-2 remission will be determined at Weeks 2, 4, 6, 8, and 10, in addition to Week 12, using the same definition in session 8.1.1and session 8.2.1. At each visit, subjects' status of dosage increase of baseline Crohn's disease medications, taking a new medication for Crohn's disease during the double blind treatment period and withdraw from the trial during the double blind period will be checked. If any of these scenarios is observed, the subject will be considered as treatment failure – non remission.

8.4 Handle of Multiplicity for Efficacy Analyses

To control the family-wise error rate (FWE) of multiple comparisons at alpha level of 0.05, the average effect method as a global test [2] and the hierarchical testing procedure as a gatekeeping strategy will be applied to the primary and secondary endpoints.

In order to handle multiplicity issue in the analysis of the primary efficacy endpoint, the average treatment difference of the pooled OPS-2071 600 mg BID and 300 mg BID treatment groups with placebo will be tested first at 2-sided alpha level of 0.05. If this global test is significant, then comparisons for each dose group of OPS 2071 600 mg BID and 300 mg BID versus placebo will be performed at 2-sided alpha level of 0.05. The comparison of a dose group is significant at 0.05 FWE significant level if both the global test and the test of the dose group's comparison with placebo are significant at alpha level of 0.05. If both null hypotheses of OPS-2071 600 mg BID, and 300 mg BID in the primary endpoint are rejected with an experiment wise Type I error of 0.05, a formal test of the secondary efficacy endpoint of endoscopic response for comparisons of OPS-2071 600 mg BID versus placebo and 300 mg BID versus placebo will be conducted, using a similar procedure used for the primary endpoint. If both null hypotheses of OPS-2071 600 mg BID versus placebo and 300 mg BID versus placebo in this secondary efficacy endpoint are rejected with an experiment wise Type I error of 0.05, the formal comparison of OPS-2071 150 mg BID versus placebo in the primary efficacy endpoint will be conducted, using the same model provided in section 8.1.2 for the primary analysis. If the null hypothesis of the primary efficacy endpoint of OPS-2071 150 mg BID is rejected with experiment wise Type I error of 0.05, the formal comparison of OPS-2071 150 mg BID versus placebo in the secondary efficacy endpoint of endoscopic response will be conducted, using the same model provided in the section 8.2.1 for the analysis of this secondary efficacy endpoint. Formal analyses of other secondary efficacy endpoints will start if the null hypothesis of OPS-2071 150 mg BID in the secondary efficacy endpoint of endoscopic response is rejected with experiment wise Type I error of 0.05.

The other secondary efficacy endpoints of this trial will be tested by the order of the gatekeeping test procedure. Formal test of the secondary efficacy endpoint #2 will start with a comparison of pooled OPS-2071 600 mg BID and 300 mg BID versus placebo. If it is significant, comparisons of 600 mg BID versus placebo and 300 mg BID versus placebo will be performed. If both these two comparisons are significant, formal comparison of OPS-2071 150 mg BID versus placebo will be conducted for the secondary efficacy endpoint #2. If this comparison is again significant, analysis of the secondary efficacy endpoint #3 will start using the same procedure, up to the secondary

efficacy endpoint #6.: Overall, this analytical procedure along with the gatekeeping protects the experiment wise Type I error in the primary and secondary analyses.

9 Safety Analyses

9.1 Extent of Exposure

Extent of exposure to the trial medication for safety sample will be summarized using descriptive statistics (n, percentage).

9.2 Adverse Events

All AEs will be coded by system organ class and the Medical Dictionary for Regulatory Activities preferred term. The incidence of the following events will be summarized by treatment group:

- TEAEs
- TEAEs by severity
- TEAEs potentially causally related to the IMP
- TEAEs with an outcome of death
- Serious TEAEs
- TEAEs leading to discontinuation of the IMP

The TEAEs from the start of IMP to the first day of bowel preparation for ileocolonoscopy will be included in the TEAE summaries for all subjects who have bowel preparation for ileocolonoscopy at Week 12. The TEAEs that occur during the ileocolonoscopy preparation and procedure will be summarized separately.

9.3 Clinical Laboratory Data

Summary statistics for changes from baseline in the central clinical laboratory measurements will be provided for the safety sample. Potentially clinically significant results in laboratory tests identified using prospectively defined criteria, including criteria for liver enzyme elevations, will also be summarized for the safety sample. By-subject listings will be provided for data of local laboratory tests.

In addition, According to FDA Guidance,[3], laboratory measurements that signal the potential for Hy's Law will be reported. An incidence table and a listing will be provided for subjects who meet one or combinations of the following criteria, without initial findings of cholestasis (alkaline phosphatase activity > 2 times the ULN):

- ALT or AST ≥ 3 times the ULN
- Bilirubin ≥ 2 times the ULN

9.4 Vital Sign Data

Summary statistics for changes from baseline in vital signs and potentially clinically significant results in vital signs will be summarized for the safety sample.

9.5 Physical Examination Data

By-subject listings will be provided for physical examinations.

9.6 Electrocardiogram Data

The mean change from baseline and the incidence of clinically significant changes will be calculated for ECG parameters.

For the analysis of QT and the corrected QT interval (QTc), data from 3 consecutive complexes (representing 3 consecutive heart beats) will be measured to determine average values. The following QT corrections will be used for reporting purposes in the clinical study report:

1) QTcF is the length of the QT interval corrected for heart rate by the Fridericia formula: $QTcF=QT/(RR)^{0.33}$

QTcN is the length of the QT interval corrected for heart rate by the FDA Neuropharm Division formula: $QTcN=QT/(RR)^{0.37}$

A regression analysis will be used to derive the slope and its 90% confidence interval of the change in QTc, adjusted from the placebo effect (ΔΔQTc), against the drug concentration by robust blood sampling visit.

10 Interim Analysis

No interim analysis is planned for this trial.

11 Changes in the Planned Analyses

The planned analyses will not be performed due to the early termination of the trial with only three randomized subjects.

12 References

1 Yuan Y. Sensitivity analysis in multiple imputation for missing data. Proceedings of the SAS Global Forum Conference; 2014 Mar 23-26; Washington, DC; 2014.

2 Guidance for Industry, Drug-Induced Liver Injury: Premarketing Clinical Evaluation. US Department of Health and Human Services, Food and Drug Administration, and Center for Drug Evaluation and Research (CDER), July 2009.

3 Kong L, Koch G, Liu T and Wang H. Performance of some multiple testing procedures to compare three doses of a test drug and placebo. Pharmaceutical Statistics 2005; 4: 25-35.

13 Appendices

Appendix 1 Criteria for Identifying Vital Signs and Weight of Potential Clinical Relevance

Variable	Criterion Value[a]	Change Relative to Baseline[a]
Heart Rate[b]	> 120 bpm < 50 bpm	≥ 15 bpm increase ≥ 15 bpm decrease
Systolic Blood Pressure[b]	> 180 mmHg < 90 mmHg	≥ 20 mmHg increase ≥ 20 mmHg decrease
Diastolic Blood Pressure[b]	> 105 mmHg < 50 mmHg	≥ 15 mmHg increase ≥ 15 mmHg decrease
Orthostatic Hypotension	≥ 20 mmHg decrease in systolic blood pressure and a ≥ 25 bpm increase in heart rate from supine to sitting/standing	Not Applicable
Weight	-	≥ 7% increase ≥ 7% decrease

[a]In order to be identified as potentially clinically relevant, an on-treatment value must meet the "Criterion Value" and also represent a change from the subject's baseline value of at least the magnitude shown in the "Change Relative to Baseline" column.

[b]As defined in "Supplementary Suggestions for Preparing an Integrated Summary of Safety Information in an Original NDA Submission and for Organizing Information in Periodic Safety Updates," FDA Division of Neuropharmacological Drug Products draft (2/27/87).

Appendix 2 Criteria for Identifying Laboratory Values of Potential Clinical Relevance

Laboratory Tests	Criteria
Chemistry[a]	
AST (SGOT)	≥ 3 x upper limit of normal (ULN)
ALT (SGPT)	≥ 3 x ULN
Alkaline phosphatase	≥ 3 x ULN
LDH	≥ 3 x ULN
BUN	≥ 30 mg/dL
Creatinine	≥ 2.0 mg/dL
Uric Acid	
Men	≥ 10.5 mg/dL
Women	≥ 8.5 mg/dL
Bilirubin (total)	≥ 2.0 mg/dL
CPK	≥ 3 x ULN
Prolactin	> ULN
Hematology[a]	
Hematocrit	
Men	≤ 37 % and decrease of ≥ 3 percentage points from Baseline
Women	≤ 32 % and decrease of ≥ 3 percentage points from Baseline
Hemoglobin	
Men	≤ 11.5 g/dL
Women	≤ 9.5 g/dL
White blood count	≤ 2,800/ mm^3 or ≥ 16,000/ mm^3
Eosinophils	≥ 10%
Neutrophils	≤ 15%
Absolute neutrophil count	≤ 1,000/ mm^3
Platelet count	≤ 75,000/ mm^3 or ≥ 700,000/ mm^3
Urinalysis[a]	
Protein	Increase of ≥ 2 units
Glucose	Increase of ≥ 2 units
Casts	Increase of ≥ 2 units
Additional Criteria	
Chloride	≤ 90 mEq/L or ≥ 118 mEq/L
Potassium	≤ 2.5 mEq/L or ≥ 6.5 mEq/L
Sodium	≤ 126 mEq/L or ≥ 156 mEq/L
Calcium	≤ 8.2 mg/dL or ≥ 12 mg/dL
Glucose	
Fasting	≥ 115 mg/dL
Non-Fasting	≥ 200 mg/dL
Total Cholesterol, Fasting	≥ 240 mg/dL
LDL Cholesterol, Fasting	≥ 160 mg/dL
HDL Cholesterol, Fasting	≤ 30 mg/dL
Triglycerides, Fasting	
Men	≥ 160 mg/dL
Women	≥ 120 mg/dL

[a]As defined in "Supplementary Suggestions for Preparing an Integrated Summary of Safety Information in an Original NDA Submission and for Organizing Information in Periodic Safety Updates," FDA Division of Neuropharmacological Drug Products draft (2/27/87).

Appendix 3 Criteria for Identifying ECG Measurements of Potential Clinical Relevance

Variable	Criterion Value[a]	Change Relative to Baseline[a]
Rate		
Tachycardia	≥ 120 bpm	increase of ≥ 15 bpm
Bradycardia	≤ 50 bpm	decrease of ≥ 15 bpm
Rhythm		
Sinus tachycardia[b]	≥ 120 bpm	increase of ≥ 15 bpm
Sinus bradycardia[c]	≤ 50 bpm	decrease of ≥ 15 bpm
Supraventricular premature beat	all	not present → present
Ventricular premature beat	all	not present → present
Supraventricular tachycardia	all	not present → present
Ventricular tachycardia	all	not present → present
Atrial fibrillation	all	not present → present
Atrial flutter	all	not present → present
Conduction		
1° atrioventricular block	PR ≥ 0.20 second	increase of ≥ 0.05 second
2° atrioventricular block	all	not present → present
3° atrioventricular block	all	not present → present
Left bundle-branch block	all	not present → present
Right bundle-branch block	all	not present → present
Pre-excitation syndrome	all	not present → present
Other intraventricular conduction block[d]	QRS ≥ 0.12 second	increase of ≥ 0.02 second
Infarction		
Acute or subacute	all	not present → present
Old	all	not present → present ≥ 12 weeks post trial entry
ST/T Morphological		
Myocardial Ischemia	all	not present → present
Symmetrical T-wave inversion	all	not present → present
Increase in QTc	QTc > 450 msec (males and females)	

[a] In order to be identified as potentially clinically relevant, an on-treatment value must meet the "Criterion Value" and also represent a change from the subject's baseline value of at least the magnitude shown in the "Change Relative to Baseline" column.

[b] No current diagnosis of supraventricular tachycardia, ventricular tachycardia, atrial fibrillation, atrial flutter, or other rhythm abnormality.

[c] No current diagnosis of atrial fibrillation, atrial flutter, or other rhythm abnormality.

[d] No current diagnosis of left bundle branch block or right bundle branch block.

Appendix 4 List of Summary Tables

CT-1.1 Subject Disposition

CT-1.2 Subject Disposition by Country and Center

CT-1.3 Subject Completion Rates

CT-2 Reasons For Discontinuation

CT-3.1 Demographic Characteristics: Age, Weight, Height, BMI, Sex, Race, and Ethnicity

CT-3.2 Medical History

CT-4.1 Concomitant Medications: Medication Taken Prior to Start of Trial Therapy

CT-4.2 Concomitant Medications: Medication Taken During Trial Therapy

CT-4.3 Concomitant Medications: Medication Taken Post Trial Therapy

CT-5.1 Summary of the Percentage of Subjects Who Achieve Clinical Remission at Week 12

CT-5.2 Sensitivity Analysis of the Percentage of Subjects Who Achieve Clinical Remission at Week 12 Using Placebo-based Pattern Imputation

CT-5.3 Supplemental Analysis of the Percentage of Subjects Who Achieve Clinical Remission at Week 12 Using Cochran-Mantel-Haenszel (CMH) Test

CT-5.4 Subgroup Analysis of the Percentage of Subjects Who Achieve Clinical Remission at Week 12

CT-6.1 Summary of the Percentage of Subjects with Endoscopic Response at Week 12

CT-6.2 Summary of Change from Baseline in the SES-CD Score

CT-6.3.1 Summary of the Percentage of Subjects with PRO-2 Remission at Week 12

CT-6.3.2 Sensitivity Analysis of the Percentage of Subjects with PRO-2 Remission at Week 12 Using Placebo-based Pattern Imputation

CT-6.4 Summary of the Percentage of Subjects with Clinical Response at Week 12

CT-6.5 Summary of the Percentage of Subjects Who Achieve Endoscopic Remission at Week 12

CT-6.6 Summary of the Percentage of Subjects with a Decrease from Baseline of ≥ 100 Points in the CDAI Score

CT-7 Extent of Exposure to Trial Medication

CT-8.1 Adverse Events(All Causalities)

CT-8.2.1 Incidence of Treatment-Emergent Adverse Events by System Organ Class and MedDRA Preferred Term

CT-8.2.2 Incidence of Treatment-Emergent Adverse Events by System Organ Class, MedDRA Preferred Term and Severity

CT-8.3.1 Incidence of Potentially Drug-Related Treatment-Emergent Adverse Events by System Organ Class and MedDRA Preferred Term

CT-8.3.2 Incidence of Potentially Drug-Related Treatment-Emergent Adverse Events by System Organ Class, MedDRA Preferred Term and Severity

CT-8.4. Incidence of Deaths due to Treatment-Emergent Adverse Events by System Organ Class and MedDRA Preferred Term

CT-8.5.1 Incidence of Serious Treatment-Emergent Adverse Events by System Organ Class and MedDRA Preferred Term

CT-8.5.2 Incidence of Serious Treatment-Emergent Adverse Events by System Organ Class, MedDRA Preferred Term and Severity

CT-8.6.1 Incidence of Treatment-Emergent Adverse Events Resulting in Discontinuation by System Organ Class and MedDRA Preferred Term

CT-8.6.2 Incidence of Treatment-Emergent Adverse Events Resulting in Discontinuation by System Organ Class, MedDRA Preferred Term and Severity

CT-8.7.1 Incidence of Non-Serious Treatment-Emergent Adverse Events by System Organ Class and MedDRA Preferred Term

CT-8.7.32 Incidence of Non-Serious Treatment-Emergent Adverse Events by System Organ Class, MedDRA Preferred Term and Severity

CT-8.8.1 Incidence of Non-Serious Treatment-Emergent Adverse Events of at Least 5% in Any Group by System Organ Class and MedDRA Preferred Term

CT-8.8.2 Incidence of Non-Serious Treatment-Emergent Adverse Events of at least 5% in Any Group by System Organ Class, MedDRA Preferred Term and Severity

CT-9.1 Listing of Deaths

CT-9.2 Listing of Serious Adverse Events

CT-9.3 Listing of Trial Discontinuations Due to Adverse Events

CT-10.1 Criteria for Laboratory Values of Potential Clinical Relevance

CT-10.2.1 Listing of Laboratory Values of Potential Clinical Relevance by Subject

CT-10.2.2 Listing of Laboratory Values of Potential Clinical Relevance by Test

CT-10.3 Incidence of Laboratory Test Values of Potential Clinical Relevance

CT-10.4.1 Laboratory Parameters - Mean Change from Baseline - Serum Chemistry

CT-10.4.2 Laboratory Parameters - Mean Change from Baseline - Hematology

CT-10.4.3 Laboratory Parameters - Mean Change from Baseline – Urinalysis

CT-10.5.1 Shift Tables of Clinical Laboratory Test Results - Serum Chemistry

CT-10.5.2 Shift Tables of Clinical Laboratory Test Results - Hematology

CT-10.5.3 Shift Tables of Clinical Laboratory Test Results – Urinalysis

CT-10.6.1 Listing of Potentially Liver Injury-related Laboratory Test Abnormalities

CT-10.6.2 Incidence of Potentially Liver Injury-related Laboratory Test Abnormalities

CT-11.1 Criteria for Vital Sign Values of Potential Clinical Relevance

CT-11.2 Listing of Vital Sign Values of Potential Clinical Relevance

CT-11.3 Incidence of Vital Signs of Potential Clinical Relevance

CT-11.4 Vital Sign Parameters - Mean Change from Baseline

CT-12.1 Criteria for Potentially Clinically Relevant Abnormalities in ECG Evaluations

CT-12.2 Listing of ECG Measurements of Potential Clinical Relevance

CT-12.3 Incidence of Potentially Clinically Relevant Changes in ECG Evaluations

CT-12.4 Mean Change from Baseline in Electrocadiogram Results

CT-13.1 By-visit Analysis of CDAI remission

CT-13.2 By-visit Analysis of PRO-2 remission

CT-13.3 Summary of the Percentage of Subjects with a Biological Response in Subjects with Elevated Levels of CRP

CT-13.4 Summary of the Percentage of Subjects with a Biological Response in Subjects with Elevated Levels of FCP

CT-13.5 Summary of Characterization of the Fecal Microbiota Profile and Abundance Based on Taxonomic and Functional Annotation

CT-14 Regression Analysis of QTC Parameters