

**Janssen Research & Development**

**Statistical Analysis Plan**

---

**Phase 2, Multicenter, Randomized, Double-blind, Placebo-controlled, Platform Study  
Evaluating the Efficacy and Safety of Interventions in Subjects with Moderately to  
Severely Active Crohn's Disease - Intervention Specific Appendix to Master Clinical  
Protocol PLATFORMPACRD2001PRISM-SCARLET**

---

**Protocol 67864238PACRD2001; Phase 2a**

**JNJ-67864238**

<b>Status:</b>	Approved
<b>Date:</b>	June 29th 2021
<b>Prepared by:</b>	[Janssen Research & Development, LLC]
<b>Document No.:</b>	EDMS-RIM-201693

**Compliance:** The study described in this report was performed according to the principles of Good Clinical Practice (GCP).

**Confidentiality Statement**

The information in this document contains trade secrets and commercial information that are privileged or confidential and may not be disclosed unless such disclosure is required by applicable law or regulations. In any event, persons to whom the information is disclosed must be informed that the information is privileged or confidential and may not be further disclosed by them. These restrictions on disclosure will apply equally to all future information supplied to you that is indicated as privileged or confidential.

## TABLE OF CONTENTS

<b>TABLE OF CONTENTS .....</b>	<b>2</b>
<b>VERSION HISTORY .....</b>	<b>4</b>
<b>1. INTRODUCTION.....</b>	<b>5</b>
1.1. Objectives and Endpoint.....	5
1.2. Study Design.....	6
<b>2. STATISTICAL HYPOTHESES .....</b>	<b>7</b>
<b>3. SAMPLE SIZE DETERMINATION .....</b>	<b>8</b>
<b>4. POPULATIONS (ANALYSIS SETS) FOR ANALYSIS.....</b>	<b>8</b>
<b>5. STATISTICAL ANALYSES .....</b>	<b>8</b>
5.1. General Considerations .....	8
5.1.1. Analysis Window.....	9
5.2. Participant Dispositions.....	9
5.3. Primary Endpoint Analysis .....	10
5.3.1. Definition of Endpoint.....	10
Treatment Failure Rules:.....	11
5.3.2. Estimand.....	12
5.3.2.1. Primary Estimand (Estimand 1).....	12
5.3.3. Analysis Methods.....	14
5.3.3.1. Main Estimator (Analysis) for the Primary Estimand .....	14
5.3.3.2. Sensitivity and Supplementary Estimand for Primary Endpoint .....	14
5.3.3.3. Supplementary Estimand for Primary Endpoint .....	15
Estimand 2 .....	15
Estimand 3 .....	15
Estimand 4 .....	16
5.3.4. Subgroup Analysis.....	17
5.3.5. Multiplicity Adjustment for Secondary Endpoints .....	17
5.4. Secondary Endpoint(s) Analysis .....	17
5.4.1. Key Confirmatory Secondary Endpoint .....	17
5.4.1.1. Definition of Endpoint(s) .....	17
5.4.1.2. Estimands .....	18
5.4.1.3. Analysis Methods for the Estimands for the Major Secondary Endpoints .....	20
5.4.2. Supportive Endpoint(s) .....	20
5.4.2.1. Analysis Methods .....	21
5.5. Tertiary/Exploratory Endpoint(s) Analysis.....	21
5.6. Safety Analyses .....	21
5.6.1. Extent of Exposure .....	21
5.6.2. Adverse Events.....	22
5.6.3. Additional Safety Assessments .....	22
5.6.3.1. Clinical Laboratory Tests .....	22
5.6.3.2. Vital Signs and Physical Examination Findings.....	25
5.6.3.3. Electrocardiogram .....	25
5.7. Other Analyses.....	25
5.7.1. Pharmacokinetics .....	25
5.7.1.1. PTG-200 Concentrations.....	25
5.7.1.2. Data Handling Rules.....	26
5.7.1.3. PK vs Efficacy.....	26
5.7.2. Immunogenicity.....	27
5.7.2.1. Antibodies to PTG-200 .....	27
5.7.2.2. Neutralized Antibodies to PTG-200 .....	27

5.7.2.3.	Antibody vs Efficacy/PK.....	27
5.7.3.	Pharmacogenomic.....	27
5.7.4.	Pharmacokinetic/Pharmacodynamic Relationships .....	27
5.7.5.	Biomarker Analyses.....	28
5.7.6.	Health Economics.....	28
5.7.6.1.	PROMIS 29.....	28
5.7.6.2.	Analysis Method .....	28
5.7.7.	Other Variables and/or Parameters.....	28
5.7.8.	Definition of Subgroups .....	29
5.8.	Interim Analyses.....	29
5.8.1.	Data Monitoring Committee (DMC) or Other Review Board .....	30
<b>6.</b>	<b>SUPPORTING DOCUMENTATION .....</b>	<b>31</b>
6.1.	Appendix 1 List of Abbreviations.....	31
6.2.	Appendix 2 Changes to Protocol-Planned Analyses .....	32
6.3.	Appendix 3 Demographics and Baseline Characteristics .....	33
6.4.	Appendix 4 Protocol Deviations .....	35
6.5.	Appendix 5 Prior and Concomitant Medications .....	36
6.6.	Appendix 6 Medical History .....	37
6.7.	Appendix 7 Intervention Compliance .....	38
6.8.	Appendix 8 Adverse Events of Special Interest.....	39
6.9.	Appendix 9 Medications of Special Interest.....	40
6.10.	Appendix 10 Laboratory Toxicity Grading.....	41
<b>7.</b>	<b>REFERENCES.....</b>	<b>42</b>
Attachment 1	Decision-Making Framework .....	43
Attachment 2:	Data Simulation .....	45
Attachment 3:	Crohn's Disease Activity Index.....	46
Attachment 4:	Sample score sheet and scoring definitions for the Simple Endoscopic Score for Crohn's Disease (SES-CD).....	47
Attachment 5:	STANDARD WEIGHT TABLE .....	48

**VERSION HISTORY**

Table [xx] – SAP Version History Summary

<b>SAP Version</b>	<b>Approval Date</b>	<b>Change</b>	<b>Rationale</b>
1		Not Applicable	Initial release

## 1. INTRODUCTION

This statistical analysis plan (SAP) contains the definitions of analysis sets, key derived variables, and statistical methods to evaluate the safety and efficacy of JNJ-67864238 (also known as PTG-200), an experimental drug, in adult participants with moderately to severely active Crohn's disease for Intervention-Specific Appendix (ISA) Protocol JNJ-67864238PACRD2001 that refers to Master Protocol PLATFORMPACRD2001. Analysis results titles, mock-ups and programming instructions for all statistical outputs (tables, figures and listings) will be provided in a separate document entitled Data Presentation Specifications (DPS).

### 1.1. Objectives and Endpoint

#### Objectives

##### Primary Objective

- The primary objective is to evaluate the efficacy of JNJ-67864238 as measured by the change in the Crohn's Disease Activity Index (CDAI) score from baseline at Week 12.

##### Secondary Objectives

- To investigate the safety and tolerability of JNJ-67864238 in patients with moderately to severely active Crohn's disease
- To evaluate the efficacy of JNJ-67864238 to reduce the Simplified Endoscopic Score for Crohn's disease (SES-CD), induce clinical remission, clinical response, and endoscopic healing of the mucosa
- To evaluate the pharmacokinetics (PK), pharmacodynamics (PD), immunogenicity (if applicable), and biomarker response of JNJ-67864238

##### Other Objectives

- to evaluate the efficacy of JNJ-67864238 on patient-related outcomes as measured by the Patient-Reported Outcomes Measurement Information System (PROMIS)-29 score

#### Endpoints

##### Primary Endpoint

- Change from baseline in the CDAI score at Week 12.

##### Secondary Endpoints

- Change in the SES-CD from baseline
- Clinical response as measured by CDAI ( $\geq 100$ -point reduction from baseline in CDAI or CDAI  $< 150$ )
- Clinical remission rate as measured by CDAI (CDAI  $< 150$ )
- PRO-2 remission defined as abdominal pain (AP) mean daily score (AP component of the CDAI) at or below 1 AND stool frequency (SF) mean daily score at or below 3, i.e., AP  $\leq 1$  and SF  $\leq 3$

- Endoscopic response (defined as at least a 50% improvement from baseline in the SES-CD)
- Endoscopic remission (defined as an SES-CD score  $\leq 2$ )

### Other Endpoints

Other endpoints to be evaluated across all post-baseline visits include but are not limited to:

- Change in AP from baseline based on a 0 to 10 numerical rating scale (NRS)
- Change in number of daily loose or watery stools
- Changes from baseline in the CDAI, PRO-2 (the sum of AP and SF subscores of the CDAI score), PRO-3 (the sum of AP, stool frequency, and general well-being subscores of the CDAI score), C-reactive protein (CRP), and fecal calprotectin
- Clinical remission (CDAI  $< 150$ ), PRO-2 remission (AP  $\leq 1$  and SF  $\leq 3$ ), and clinical response based on CDAI
- Endoscopic improvement based on a reduction from baseline in the SES-CD score ( $\geq 3$  points, 25%, 50%)
- Endoscopic healing (defined as the absence of mucosal ulcerations)
- Patient's Global Assessment of Crohn's Disease Activity
- Change from baseline in multivariate score of biomarkers measured in mucosal biopsy samples, whole blood, serum, and/or stool
- Change in PROs as measured by the PROMIS-29 score

## **1.2. Study Design**

The clinical development program for experimental drug in Crohn's disease will be conducted under a platform protocol, which consists of a Master protocol and multiple ISAs. The Master protocol describes the design elements that are common across the whole platform study without any compound information. The ISA describes the specific JNJ-active compound information to evaluate the safety and efficacy of JNJ-active in Participants with moderately to severely active Crohn's disease who have demonstrated an inadequate response or failure to tolerate previous conventional therapy or biologic therapy. This SAP will be in-line follow platform protocol starting from the first ISA phase 2a Proof-of-concept study first. The first ISA is a Phase 2a, multicenter, randomized, double-blind, placebo-controlled study designed to assess the effects of a predominantly gastrointestinal-restricted anti-IL 23R compound for moderately to severely active Crohn's disease. This 12-week Phase 2a study is focused on proof-of-concept and will assess the safety and efficacy of an orally administered high dose of JNJ-67864238 compared with placebo in participants with Crohn's disease. Since the inhibition of the IL-23 pathway has been validated as a mechanism for treating Crohn's disease, a high dose was chosen to assess the efficacy of a locally acting anti-IL-23 inhibitor.

Participants may be biologic intolerant or refractory (Bio-IR) or biologic nonfailures (Bio-NF) populations as defined below:

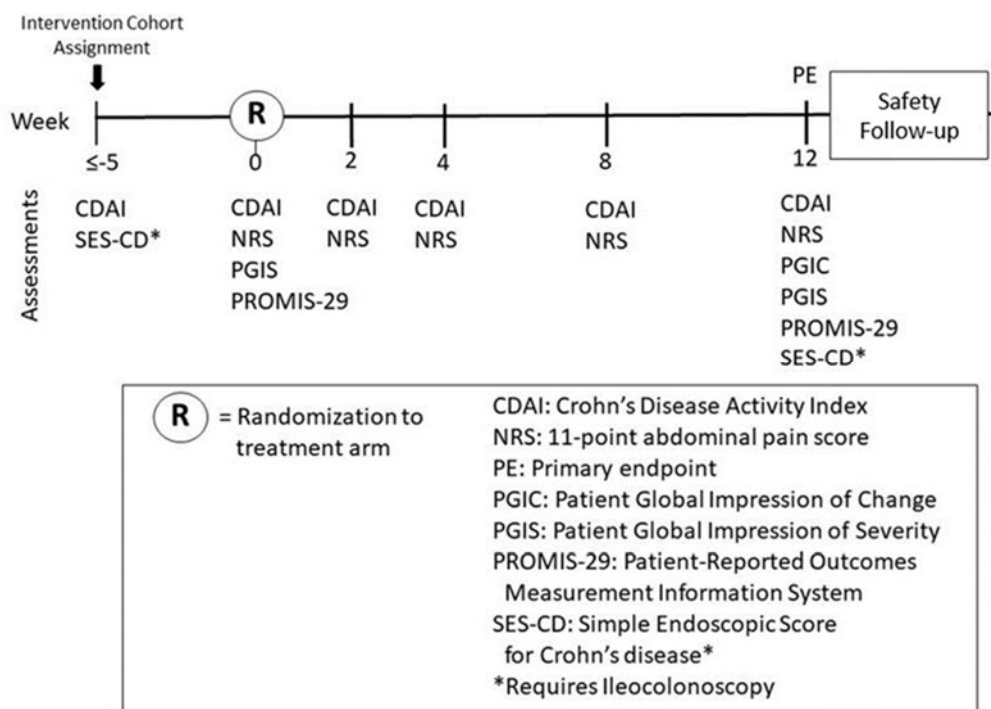
- Bio-IR Population: Those participants who have received infliximab, adalimumab, certolizumab pegol, vedolizumab, or ustekinumab at a dose approved for the treatment of Crohn's disease, and either did not respond initially, responded initially but then lost response, or were intolerant to the medication.

- **Bio-NF Population:** Those participants who have demonstrated an inadequate response to or have failed to tolerate corticosteroids or the immunomodulators 6-MP, AZA, or MTX. Participants who have demonstrated corticosteroid dependence (ie, an inability to successfully taper corticosteroids without a return of the symptoms of Crohn's disease) are also eligible. Bio-NF participants may also have received biologic therapy but only if it was discontinued for reasons other than lack of efficacy or intolerance (eg, a drug holiday).

A target of 90 Bio-IR and Bio-NF participants will be randomized to receive JNJ 67864238 CCI [REDACTED] or placebo in a 3:2 ratio using permuted block randomization, stratified by baseline Crohn's Disease Activity Index (CDAI) score ( $\leq 300$ ,  $>300$ ) and biological refractory status (Bio-IR, Bio-NF). The treatment arms for this study will be as follows: JNJ-67864238 CCI [REDACTED] (N=54) and Placebo (N=36). Of note, the sample size has been chosen to assess a level of efficacy consistent with that observed with systemically available biological therapies for Crohn's disease.

An overview of this study design is presented in [Figure 1](#) below:

**Figure 1:** Schematic Overview of Study Design



## 2. STATISTICAL HYPOTHESES

The hypothesis is that JNJ-67864238 is superior to placebo as measured by the reduction from baseline in the CDAI at Week 12 in Participants with moderately to severely active Crohn's disease.

### Primary Objective

The primary objective is to evaluate the efficacy of experimental drug JNJ-67864238 as measured by the change in the CDAI score from baseline at Week 12.

## Secondary Objectives

The major secondary objectives are as follows:

- To investigate the safety and tolerability of experimental drug JNJ-67864238 in participants with moderately to severely active Crohn's disease
- To evaluate the efficacy of experimental drug JNJ-67864238 to reduce the SES-CD, induce clinical remission, clinical response, and endoscopic healing of the mucosa
- To evaluate the PK, PD, immunogenicity (if applicable), and biomarker response of JNJ-67864238

### 3. SAMPLE SIZE DETERMINATION

The sample size determination is based on the overall population, including both Bio-IR and Bio-NF participants. To examine the hypothesis that JNJ-active is superior to placebo as measured by CDAI, sample size is planned to achieve desirable operating characteristics of a predefined decision-making framework through simulation. A 90 Bio-IR or Bio-NF participants will provide about 90% power when the true treatment difference is 60 (placebo – JNJ-67864238) with standard deviation of 93 in the change in CDAI at Week 12 for both Bio-IR and Bio-NF populations.

### 4. POPULATIONS (ANALYSIS SETS) FOR ANALYSIS

For purposes of analysis, the following analysis sets are defined:

Analysis Sets	Description
Randomized	The randomized analysis set includes all participants who were randomized in the study.
Full Analysis Set (FAS)	All participants who are randomized to a treatment arm within an intervention cohort and receive at least 1 dose of the study intervention
Per Protocol (PP)	The per protocol analysis set (PP) includes a subset of participants in the full analysis set (FAS) who were in compliance with the protocol. Compliance is defined after reviewing by study physician and statistician
Safety	All randomized participants who take at least 1 dose of the study intervention within an intervention cohort. Participants will be analyzed according to the study intervention they actually received
Pharmacokinetics Analysis Set	The PK analysis set is defined as participants who received at least 1 dose JNJ-67864238 and have at least 1 valid blood sample drawn for PK analysis.
Immunogenicity Analysis Set	The Immunogenicity analysis set is defined as participants who received at least 1 dose JNJ-67864238 and have at least 1 valid blood sample drawn for Immunogenicity analysis.

### 5. STATISTICAL ANALYSES

#### 5.1. General Considerations

Descriptive statistics (eg, mean, median, SD, minimum, and maximum) will be used to summarize continuous variables. Counts and percentages will be used to summarize categorical variables. Graphic data displays may also be used to summarize the data. Analyses suitable for categorical data (eg, chi-square tests, Cochran-Mantel-Haenszel chi-square tests, or logistic regression, as appropriate) will be used to compare the proportions of participants achieving selected endpoints



(eg, clinical remission). In cases of sparse events, Fisher's exact test will be considered for treatment comparisons. Continuous response parameters will be compared using a longitudinal model, an analysis of variance (ANOVA) model, or an analysis of covariance (ANCOVA) model. If non-normality is observed, then van der Waerden normal scores will be used to fit a model. For primary endpoint change from baseline in the CDAI score will be fit through a MMRM with categorical effects of population, treatment, time and treatment-by-time, and baseline CDAI score as a continuous covariate as needed, an appropriate external prior information will be incorporated in the form of a prior probability distribution function to evaluate various study hypotheses in a Bayesian framework. In such a framework, inference will be based on posterior probability distributions.

### 5.1.1. Analysis Window

The study date and corresponding study visit/week will be captured on each case report form (CRF). Visit windows will be created around the study day of each scheduled visit. It will be used to aggregate data, which are to be summarized by visit. The visit windows and the targeted study day are indicated in Table 1. Details will be added later if more than one assessment falls within the same visit window. If an assessment is not scheduled for every visit, windows will be combined so that the interval between targeted study days is split evenly and consistently between visits.

**Table 1. Visit Windows**

Scheduled Study Day	Visit Window
Screening	-999, -1
Week 0 Day 1*	1, 7
Week 2 Day 1 (Day 15)	8, 21
Week 4 Day 1 (Day 29)	22, 42
Week 8 (Day 57)	43, 70
Week 12 (Day 85)	71, 98
Safety Follow-up (Week 16)	99, 999
* Study Day 1 begins on the day of randomization. Each Week consists of 7 days.	

## 5.2. Participant Dispositions

The number of participants in the following disposition categories will be summarized throughout the study by intervention group and overall:

- Participants who received study intervention
- Participants who completed the study
- Participants who discontinued study intervention
- Reasons for discontinuation of study intervention
- Participants who terminated study prematurely
- Reasons for termination of study

The distribution of the time to [study termination/discontinuation of study intervention] will be displayed with Kaplan-Meier curves. Participants who terminate study participation

prematurely/discontinue study intervention at any time will be considered an ‘Event’ and their date of study termination/study intervention discontinuation will be used in the time to event calculation. Participants who complete [the study/study intervention] will be censored and the date of [study completion/last dose of study intervention] will serve as the time of censoring.

A listing of participants will be provided for the following categories:

- Participants who discontinued study intervention
- Participants who terminated study prematurely
- Participants who were unblinded during the study period
- Participants who were randomized yet did not receive study intervention.

### 5.3. Primary Endpoint Analysis

**Level of Significance:** In general, all tests will be performed at a 2-sided significance level of 0.1, unless otherwise specified. All interval estimations will be reported using 2-sided 90% confidence intervals (CIs).

Multiplicity adjustment consideration will be based on an overall type I error rate of 0.1 (2-sided). The hypothesis testing will be conducted in a hierarchical manner with the test on the primary endpoint conducted first and the tests on major secondary endpoints conducted next. Within the major secondary endpoints, the sequential gatekeeping approach in conjunction with graphical multiplicity control procedures will be applied. As the interim analysis (IA) will not lead to an early completion due to success, multiplicity adjustment is not required. Nominal p-values will be calculated for all treatment comparisons.

The efficacy analysis of the primary will be based on the full analysis set. Regardless of the actual treatments the participants received, efficacy data will be analyzed according to the initial randomization assignment for the treatment phase.

#### 5.3.1. Definition of Endpoint

The primary endpoint will be change from baseline in the CDAI score at Week 12 based on full analysis set. This will apply for both interim analysis and final analysis.

The CDAI is a validated multi-item measure of severity of illness derived as a weighted sum of 8 different Crohn’s disease-related variables<sup>1</sup>. These 8 variables are:

- extra-intestinal manifestations
- abdominal mass
- weight
- hematocrit
- use of antidiarrheal drug(s) and/or opiate
- total number of liquid or very soft stools
- abdominal pain/cramps

- general well-being.

The last 3 variables are scored over 7 days by the participant on a diary card. For the total number of liquid or very soft stools, abdominal pain/cramps, and general well-being, if only 5 days or 6 days of data are available for the calculation, the weights of 7/5 and 7/6, respectively, will be used for the calculation; if the values are recorded for less than 5 days, the component will not be calculated.

At each timepoint, the CDAI score will only be calculated if  $\geq 4$  of the 8 components are available. If the CDAI score cannot be calculated (ie,  $<4$  components available), the CDAI score will be considered missing. When at least 4 of the 8 components are available, any missing components will be imputed by carrying forward the last non-missing component.

The weight component of the CDAI will be based on the standard weight table (Attachment 5).

Subjects who do not return for evaluation or have a missing CDAI score at Week 12 will be considered to have no change at Week 12 from their baseline CDAI score.

### **Treatment Failure Rules:**

Treatment failure rules will be applied for all efficacy endpoints unless otherwise specified. In such a case, baseline values (at Week 0) will be assigned from the point of treatment failure onward, regardless of the observed data, for continuous endpoints, and participants will be considered as not achieving the respective endpoints for dichotomous endpoints. Treatment failure rules override other data handling rules.

For dichotomous endpoints, participants who have a treatment failure are considered as not achieving the respective endpoints after treatment failure. For continuous endpoints, participants who have a treatment failure will have their baseline values carried forward after treatment failure onwards.

The treatment failure rules will be applied to determine the final CDAI score. The treatment failure rules override the actual observation of the CDAI score. Participants who have any of the following events prior to the Week 12 visit will be considered a treatment failure and will be considered to have no change at Week 12 from their baseline CDAI score, regardless of the actual CDAI score:

- A Crohn's disease-related surgery (with the exception of drainage of an abscess or seton placement)
- Discontinuation of study intervention due to lack of efficacy or due to an adverse event (AE) of worsening Crohn's disease
- Specified changes in concomitant Crohn's disease medications described below.

The treatment failure rules for physician prescribed concomitant medication changes between Week 0 and Week 12 are:

1. Corticosteroids:
  - a. Initiation of oral corticosteroids (excluding budesonide) due to worsening Crohn's disease.
  - b. Initiation of oral corticosteroids (excluding budesonide) for more than 7 days due to reasons other than worsening Crohn's disease.
  - c. Initiation of oral budesonide due to worsening Crohn's disease.
  - d. Initiation of oral budesonide for more than 7 days due to reasons other than worsening Crohn's disease.
  - e. Increase by  $\geq 5$  mg above baseline in the prednisone equivalent dosage of oral corticosteroids (excluding budesonide) due to worsening Crohn's disease.
  - f. Increase by  $\geq 5$  mg above baseline in the oral prednisone equivalent dosage of oral corticosteroids (excluding budesonide) for more than 7 days due to reasons other than worsening Crohn's disease.
  - g. Increase by  $\geq 3$  mg above baseline in the dose of oral budesonide due to worsening Crohn's disease.
  - h. Initiation of IV corticosteroids due to worsening Crohn's disease.
  - i. Initiation of rectal corticosteroids due to worsening Crohn's disease.
2. 5-ASA compounds:
  - a. Initiation or increase of oral 5-ASA compounds due to worsening Crohn's disease.
3. Immunomodulator agents:
  - a. Initiation of oral 6-MP/AZA due to worsening Crohn's disease.
  - b. Initiation of oral, subcutaneous, or intramuscular MTX due to worsening Crohn's disease.
  - c. Increase above baseline in the dose of oral 6-MP/AZA due to worsening Crohn's disease.
  - d. Increase above baseline in the dose of oral, subcutaneous, or intramuscular MTX due to worsening Crohn's disease (with the same route).
4. Protocol-prohibited medications:

An initiation of any of the following after baseline due to worsening Crohn's disease:

  - a. Immunomodulatory agents other than 6-MP/AZA or MTX (including but not limited to 6-TG, cyclosporine, tacrolimus, sirolimus, mycophenolate mofetil).
  - b. Immunomodulatory biologic agents (including but not limited to TNF antagonists, natalizumab, abatacept).
  - c. Experimental Crohn's disease medications (including but not limited to thalidomide, briakinumab, vedolizumab, traficet, AMG 827).
5. Crohn's disease related antibiotics:
  - a. Initiation or change of antibiotics due to worsening Crohn's disease.

### **5.3.2. Estimand**

#### **5.3.2.1. Primary Estimand (Estimand 1)**

The primary estimand uses a Composite Strategy that considers meeting the treatment failure (TF) criteria as an unfavorable outcome. This strategy aims to achieve an objective treatment effect.

The primary estimand, the main clinical quantity of interest to be estimated in the study, is defined by the following components:

**Treatment:** Oral tablets CCI [REDACTED] for a 12 week period

- Experimental: JNJ-67864238 (PTG-200), CCI [REDACTED]
- Control: Matching Placebo tablets CCI [REDACTED]

**Population:** Participants with moderately to severely active Crohn's disease, The target population consists of men or women 18 to 75 years of age with moderately to severely active Crohn's disease of at least 3 months' duration, defined as a CDAI score  $\geq 220$  but  $\leq 450$  at Week 0, with an SES-CD score  $\geq 3$  by central endoscopy read. Participants with a screening SES-CD score  $< 3$  can be randomized and received at least one study dose if they have an elevated CRP ( $> 0.3$  mg/dL or 3.0 mg/L) or elevated fecal calprotectin ( $> 250$   $\mu\text{g/g}$ ). Participants must have colitis, ileitis, or ileocolitis previously confirmed at any time in the past by radiography, histology, and/or endoscopy. Participants must be biologic intolerant or refractory (Bio-IR) or biologic nonfailures (Bio-NF) and must be on a stable dose of at least 1 protocol-permitted therapy for the treatment of Crohn's disease prior to baseline and not receiving any protocol-prohibited therapies. For more details and for a complete list of inclusion/exclusion criteria, please refer to the protocol.

**Variable (Endpoint):** Change from baseline in the CDAI score at Week 12. Participants who have intercurrent events in categories 1-3 (defined below) prior to the Week 12 visit will have their baseline CDAI score assigned, regardless of the observed data.

### Intercurrent Events and Corresponding Strategies:

The following are the intercurrent events considered for this trial:

1. A Crohn's disease-related surgery (with the exception of minor procedures such as drainage of a superficial abscess or seton placement, etc.);
2. A prohibited change in CD medication;
3. Discontinuation of study intervention due to lack of efficacy or an AE of worsening Crohn's disease;
4. Discontinuation of study intervention due to COVID-19 related reasons (like missing the assessment visit due to COVID-19 quarantine, but excluding COVID-19 infection)
5. Discontinuation of study intervention due to reasons other than lack of efficacy or an AE of worsening Crohn's disease (including AE/SAE due to COVID-19 infection).

Intercurrent events (ICEs) in categories 1-3 will be handled by the composite strategy, and ICE category 4 will be handled by the hypothetical strategy (as if subjects would have not experienced this intercurrent event), and ICE category 5 will be handled by the treatment policy strategy. The composite strategy assesses the treatment effects not only based on the variable measurements, but also based on intercurrent events. This estimand acknowledges that having an intercurrent event in categories 1-3 is an unfavorable outcome, ie. baseline value will be used after ICEs. For subjects experiencing ICE category 4, CDAI data at all visits after an ICE will be set to missing with

assumption of missing at random. For ICE category 5, treatment policy strategy, the observed values after ICEs will be used, if available.

**Population-level summary:** The difference of change in the CDAI score at Week 12 between the Experimental treatment group and the placebo treatment group.

### 5.3.3. Analysis Methods

#### 5.3.3.1. Main Estimator (Analysis) for the Primary Estimand

The primary analysis will be based on the full analysis set which will include all participants randomized and who receive at least 1 dose of the study intervention. Participants in the efficacy analyses will be summarized according to their assigned treatment arm regardless of whether or not they receive the assigned study intervention. The primary endpoint will be analyzed based on the primary estimand.

Data used in the primary analysis (i.e. the main estimator for the primary estimand) include all values for change from baseline in CDAI score collected up to and including Week 12 for all participants in the primary efficacy analysis set (ie, those who were randomly assigned to study intervention and received at least one dose of study intervention). Any CDAI score collected after experiencing intercurrent event categories 1-3 will be replaced by the baseline observation for CDAI. For subjects experiencing ICE category 4, CDAI data at all visits after an ICE will be set to missing with assumption of missing at random. For ICE category 5, treatment policy strategy, the observed values after ICEs will be used, if available.

The main estimator for the primary estimand is based on the Mixed Effects Model for Repeated Measures (MMRM), under the Missing at Random (MAR) assumption for the missing responses, the missing data will be accounted for through correlation of repeated measures in the model, to test the difference between the experimental group and the placebo group. The analysis will use change in CDAI scores at post-baseline time points through Week 12 as dependent variables. The explanatory variables of the MMRM model will include the fixed, categorical effects of treatment, time, and treatment-by-time interaction, as well as the continuous, fixed covariates of baseline and baseline-by-time interaction. If the intervention cohort includes both Bio-IR and Bio-NF participants, then the biologic therapy status will be included in the MMRM model as well. An unstructured variance-covariance structure will be used to model the within-participant errors. The treatment comparison at Week 12 in terms of the least square mean difference and confidence intervals for decision making will be estimated based on this model.

The estimates for the primary estimand of the treatment difference between each experimental group and the placebo group will be provided by the difference in the least squares means (LSmeans). The 90% confidence interval (CI) for the differences in LSmeans and p-values will be calculated based on the MMRM. An unstructured covariance matrix for repeated measures within a subject will be used.

#### 5.3.3.2. Sensitivity and Supplementary Estimand for Primary Endpoint

To examine the robustness of the primary estimator under the different assumption from the main estimator. Sensitivity estimators will be conducted for primary estimand

**Sensitivity Estimator 1:**

- **Assumption:** Consider different analysis population
- **Sensitivity Estimator** Per-protocol analysis: participants with major deviations/identified by SRP (study responsible physician) will be excluded.

**Sensitivity Estimator 2:**

- **Assumption:** Consider different statistical model ANCOVA
- **Sensitivity Estimator** An ANCOVA model will be applied with the change from baseline in the CDAI at Week 12 as the response variable, treatment and population status (Bio-IR and Bio-NF) as fixed factors, and baseline CDAI score as a covariate.

**Sensitivity Estimator 3:**

- **Assumption:** Consider background concomitant medication steroids tapering as treatment failure
- **Sensitivity Estimator:** Imputing CDAI score based on last value carried forward (LVCF) method for participants who have any steroids tapering prior to week 8.

Other exploratory estimators may be conducted as necessary including but are not limited to

- Completer analysis for participant who completed week 12 treatment
- Imputing the CDAI score using the values prior to the treatment failure (ICE category 1-3) or prior to COVID-19 related ICE categories 4 and 5,
- Observed CDAI without incorporating treatment failure

**5.3.3.3. Supplementary Estimand for Primary Endpoint****Estimand 2**

**Composite strategy:** This supplementary estimand will be same as Estimand 1 except for ICE categories 4 and 5, ICE categories 4 and 5 will be handled same as ICEs categories 1-3: baseline CDAI values will be used at all visits after an ICE.

**Analysis for Estimand 2:**

For testing of the primary endpoint using the supplementary estimand 3, the change from baseline in the CDAI score at Week 12 will be compared between experimental group and the placebo group using MMRM with categorical effects of population, treatment, time and treatment-by-time, and baseline CDAI score as a continuous covariate, with assumption of missing at random.

**Estimand 3**

**Hypothetical strategy:** This supplementary estimand will be handled by the hypothetical strategy for ICE categories 1-5: CDAI data at all visits after an ICE will be set to missing and assuming missing at random.

**Analysis for Estimand 3:**

For testing of the primary endpoint using the supplementary estimand 3, the change from baseline in the CDAI score at Week 12 will be compared between experimental group and the placebo group using MMRM with categorical effects of population, treatment, time and treatment-by-time, and baseline CDAI score as a continuous covariate, with assumption of missing at random.

#### Estimand 4

**Tipping point method:** This supplementary estimand will be used to assess the robustness of primary estimand using tipping point method to assess COVID-19 ICE impact as appropriate. For the change from baseline in CDAI, a delta adjustment multiple imputation method will be performed to evaluate the impact of missing data when deviating from the MAR assumption.

This method will employ the following 3 steps:

##### Step 1 - Multiple Imputation (MI)

Preparing monotone missing data pattern, and then run MAR-based multiple imputation with the regression option to impute missing values. For subjects with a non-monotone missing data pattern, ie, subjects may have ‘intermediate missing’ data between non-missing observations, imputing the intermediate missing values using the MCMC (Markov Chain Monte Carlo) method. This will be done using SAS PROC MI.

##### Step 2 - Delta Adjustments

The imputed values for subjects who discontinued will be adjusted by adding Delta ( $\delta$ ) to the results from step 1. Delta-adjusted fully imputed datasets will be generated for different combinations of  $\delta$  values for active treatment group and placebo group as defined below:

- $\delta = 0$  to tipping point value of  $\Delta$  by incremental of 1 to active treatment only
- $\delta = 0$  to tipping point value of  $\Delta$  by incremental of 1 to active treatment; for control group,  $\delta =$  half of active treatment adjustment
- $\delta = 0$  to tipping point value of  $\Delta$  by incremental of 1 for both active treatment and placebo groups

##### Step 3 – Analysis and Pooling

For each  $\delta$  combination described in step 2:

- Run MMRM model for each set of the adjusted datasets, the p-values for testing the treatment difference will be obtained.
- Multiple imputation combining rules in PROC MIANALYZE will be applied to the MMRM

The delta adjustment method will be applied for two scenarios:

- To all subjects who discontinued due to ICE categories 4 and 5.
- To all subjects who discontinued, except those with a discontinuation reason that was not considered related to the study drug, ie. lost to follow-up.



## Analysis for Estimand 4:

For testing of the treatment benefit impacted by COVID-19, a turning point threshold will be explored based on PTG-200 treatment effect.

### 5.3.4. Subgroup Analysis

The consistency of treatment effect for the primary endpoint will be evaluated for the subgroups defined in Section 5.7.8. For each of these subgroups, the difference of experimental group vs placebo in the change from baseline in the CDAI score at Week 12 and the associated 90% confidence interval will be provided. The same data handling rules will be applied as the primary efficacy analysis.

### 5.3.5. Multiplicity Adjustment for Secondary Endpoints

Inferential hypothesis testing will be conducted in a hierarchical manner with the test on the primary endpoint conducted first. Major secondary endpoints will be conducted only if the primary analysis of CDAI change from baseline achieves statistical significance. Within the major secondary endpoints, the gatekeeping approach in conjunction with graphical multiplicity control procedures will be applied. All statistical significance levels are defined at a 2-sided alpha of 0.1.

## 5.4. Secondary Endpoint(s) Analysis

### 5.4.1. Key Confirmatory Secondary Endpoint

#### 5.4.1.1. Definition of Endpoint(s)

The major secondary endpoints include the following at Week 12:

- Change in the SES-CD from baseline
- Clinical response as measured by CDAI ( $\geq 100$ -point reduction from baseline in CDAI or CDAI  $< 150$ )
- Clinical remission rate as measured by CDAI (CDAI  $< 150$ )
- PRO-2 remission defined as abdominal pain (AP) mean daily score (AP component of the CDAI) at or below 1 AND stool frequency (SF) mean daily score at or below 3, ie, AP  $\leq 1$  and SF  $\leq 3$
- Endoscopic response (defined as at least a 50% improvement from baseline in the SES-CD)
- Endoscopic remission (defined as an SES-CD score  $\leq 2$ )

The (SES-CD) is a scoring system developed to provide a more granular evaluation of endoscopic disease severity in patients with Crohn's disease. It is constructed based on the evaluation of 4 endoscopic components across 5 predefined ileocolonic segments. The 4 endoscopic components within each segment are: the presence/ size of ulcers, the proportion of mucosal surface covered by ulcers, the proportion of mucosal surface affected by any other lesions, and the presence/ type of narrowing (also commonly referred to as strictures/ stenosis clinically). Each endoscopic component is scored from 0 to 3 for each segment, and a total score is calculated as a sum of all the component scores across all the segments, as outlined in [Attachment 4](#). The total SES-CD score ranges from 0 to 56.

Calculation of the SES-CD score:

Main approach (SES-CD calculated based on all segments available):

The total SES-CD score at a visit will be calculated based on all segments scored at the visit. If the total SES-CD score cannot be calculated (i.e., no segment is scored) at a visit, the total SES-CD score will be considered missing.

Alternative approach (baseline segments matched approach):

To calculate the SES-CD score at a visit, the sum of the segments that were present at baseline will be used. For segments that were present at baseline but missing post-baseline, the baseline score for the missing segment(s) will be carried forward. In the event that a segment is missing at baseline but non-missing at post-baseline, the non-missing post-baseline score is not used in the calculation of SES-CD.

The Week 12 SES-CD Total score will be calculated only if the participant has a Week 12 endoscopy and at least one segment is evaluable. For segments evaluated and scored at baseline but not at Week 12, associated baseline segment scores will be carried forward to calculate the Total SES-CD score at Week 12. Otherwise, the Week 12 Total SES-CD will be missing.

Participants who meet 1 or more treatment failure rules (as specified for the primary endpoint) before Week 12 will be considered not to be in clinical remission, clinical response or clinical-biomarker response, either based on CDAI or PRO-2. For endoscopic response, the same treatment failure rules will be applied as appropriate. Any missing values at week 12 will be considered as no responder/remitter.

#### 5.4.1.2. Estimands

The following describes the attributes of the estimands for the major secondary endpoints:

**Treatment by Week 12:** Same as Estimand 1, ie:

**Treatment:** CCI [REDACTED] for a 12 weeks period

- Experimental: JNJ-67864238 (PTG-200), CCI [REDACTED]
- Control: Matching Placebo tablets CCI [REDACTED]

**Population:** Participants with moderately to severely active Crohn's disease,

#### Variables and Population-level Summary (for Estimands 5-10)

**Table 2: Variables and Population-level Summary for the estimand for each Major Secondary Endpoint**

Estimand	Variable (Endpoint)	Population-level summary
5	Change in the SES-CD from baseline (both main approach and <u>alternative approach analyses will be performed</u> )	Change from baseline in the SES-CD score at Week 12 or early withdrawal visit between experimental group and placebo.

6	Clinical response as measured by CDAI ( $\geq 100$ -point reduction from baseline in CDAI or CDAI $< 150$ )	Difference in percentage of subjects who achieved clinical response at Week 12 between experimental group and placebo.
7	Clinical remission rate as measured by CDAI (CDAI $< 150$ )	Difference in percentage of subjects who achieved clinical remission at Week 12 between experimental group and placebo.
8	PRO-2 remission defined as abdominal pain (AP) mean daily score (AP component of the CDAI) at or below 1 AND stool frequency (SF) mean daily score at or below 3, ie, AP $\leq 1$ and SF $\leq 3$	Difference in percentage of subjects who achieved PRO-2 remission at Week 12 between experimental group and placebo.
9*	Endoscopic response (defined as at least a 50% improvement from baseline in the SES-CD) (both main approach and <u>alternative approach analyses will be performed</u> )	Difference in percentage of subjects who achieved endoscopic response at Week 12 between experimental group and placebo.
10	Endoscopic remission (defined as an SES-CD score $\leq 2$ ) (both main approach and <u>alternative approach analyses will be performed</u> )	Difference in percentage of subjects who achieved endoscopic remission at Week 12 between experimental group and placebo.

\* For Endoscopic response, please refer to below prohibited medications for ICE category 2

- Immunomodulatory agents other than AZA, 6-MP, or MTX (including, but not limited to, 6-TG, cyclosporine, mycophenolate mofetil, tacrolimus, and sirolimus).
- Immunomodulatory biologic agents (including, but not limited to, TNF antagonists, natalizumab, ustekinumab, rituximab, vedolizumab). Ustekinumab is permitted in this study only in participants randomly assigned to ustekinumab and only as stipulated in this protocol.
- Experimental Crohn's disease medications (including, but not limited to, upadacitinib, filgotinib, ozanimod, etrolizumab, brazikumab, mirakizumab [formerly LY-3074828], risankizumab, GS-5745)
- Thalidomide or related agents.

### Intercurrent Events and Corresponding Strategies:

The intercurrent events for all major secondary endpoints other than endoscopic response at Week 12 will be same as those used in the primary estimand, and are as follows:

1. A Crohn's disease-related surgery (with the exception of minor procedures such as drainage of a superficial abscess or seton placement, etc.);
2. A prohibited change in CD medication;
3. Discontinuation of study intervention due to lack of efficacy or an AE of worsening Crohn's disease;
4. Discontinuation of study intervention due to COVID-19 related reasons (like missing the assessment visit due to COVID-19 quarantine, but excluding COVID-19 infection)
5. Discontinuation of study intervention due to reasons other than lack of efficacy or an AE of worsening Crohn's disease (including AE/SAE due to COVID-19 infection).

ICEs will be handled by the same Estimands strategies as the primary Estimands strategies for the major secondary endpoints including clinical remission, clinical response, PRO-2 remission, endoscopic response (refer to Table 2 footnotes for ICE category 2), or endoscopic remission.

#### **5.4.1.3. Analysis Methods for the Estimands for the Major Secondary Endpoints**

The major secondary endpoints of clinical remission (either based on CDAI, SES-CD, or PRO-2), clinical response, endoscopic response, and clinical-biomarker response at Week 12 will be compared between PTG-200 group and the placebo group.

A logistic regression model will be used in the responder analyses with the treatment arm and population status (Bio-IR or Bio-NF), if applicable, as fixed factors, at 2-sided significance level of 0.1. In this model, baseline measurements may be included as covariates as appropriate. The change from baseline in the SES-CD score at Week 12 will be evaluated using MMRM model in comparing treatment difference by LS Mean (SD), LS mean (90% CI). P-Value was based on comparison with placebo model. And other statistical analysis as appropriate.

#### **5.4.2. Supportive Endpoint(s)**

Other endpoints to be evaluated across all post-baseline visits as appropriate:

- Change in AP from baseline based on a 0 to 10 numerical rating scale (NRS)
- Change in number of daily loose or watery stools
- Changes from baseline in the CDAI, PRO-2 (the sum of AP and SF subscores of the CDAI score), PRO-3 (the sum of AP, stool frequency, and general well-being subscores of the CDAI score), C-reactive protein (CRP), and fecal calprotectin
- Clinical remission (CDAI <150), PRO-2 remission (AP  $\leq$ 1 and SF  $\leq$ 3), and clinical response based on CDAI
- Endoscopic improvement based on a reduction from baseline in the SES-CD score ( $\geq$ 3 points, 25%, 50%), (both main approach and alternative approach analyses will be performed)
- Endoscopic healing (defined as the absence of mucosal ulcerations), (both main approach and alternative approach analyses will be performed)
- Patient's Global Assessment of Crohn's Disease Activity
- Change from baseline in multivariate score of biomarkers measured in mucosal biopsy samples, whole blood, serum, and/or stool
- Change in PROs as measured by the PROMIS-29 score

ICEs will be handled by the same Estimands strategies as the primary Estimands strategies for the supportive secondary endpoints. Participants who do not return for evaluation or have missing AP or SF scores at Week 12 will be considered not to be in PRO-2 remission (ie, clinical remission measured by AP and SF). Participants who do not return for evaluations or have missing CRP or fecal calprotectin values at Week 12 will be considered not to be in clinical-biomarker response unless the non-missing values at Week 12 already confirm a clinical-biomarker response. Any missing values at week 12 will be considered as no responder/remitter.

#### **5.4.2.1. Analysis Methods**

Descriptive statistics (i.e., mean, median, SD, IQ range, minimum, and maximum) will be used to summarize continuous variables. Counts and percentages will be used to summarize categorical variables. Graphical data displays (e.g., line plots) may also be used to summarize the data. The statistical model specified in the major secondary section will be applied for the other efficacy endpoints as appropriate.

#### **5.5. Tertiary/Exploratory Endpoint(s) Analysis**

Exploratory analysis will be performed to understand the potential factors of regional and/or subpopulation variability such as covariate-adjusted analyses, subgroup summaries if it is necessary.

#### **5.6. Safety Analyses**

Safety summaries will be based on the safety analysis set and participants will be analyzed based on the treatment the participants received.

For all continuous safety variables, descriptive statistics by intervention group will include the N, mean, standard deviation, median, minimum, and maximum. Categorical variables will be summarized by intervention group using frequency counts and percentages.

Safety analyses will be based on observed data (i.e., no data imputation).

##### **5.6.1. Extent of Exposure**

The number and percentage of participants who receive study intervention will be summarized by intervention group.

Study intervention duration is defined as (date of last dose of study intervention – date of first dose of study intervention) +1. Descriptive statistics for duration of study intervention (N, mean, SD, median, and range (minimum, maximum)) will be summarized by intervention group.

Total dose days of intervention is defined as the total number of days that study intervention was administered to the participant (excluding days off study intervention).

The number (%) of participants with a dose adjustment will be summarized by intervention group. Reasons for dose adjustments will also be summarized intervention group.

Descriptive statistics will be presented for the following parameters:

- Number of administrations
- Cumulative total dose
- Mean daily dose

The mean daily dose of study intervention is calculated as (sum of total daily dose during the intervention phase)/total dose days of intervention.

## 5.6.2. Adverse Events

The verbatim terms used in the CRF by investigators to identify adverse events will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). Any AE occurring at or after the initial administration of study intervention through the day of last dose plus 30 days is considered to be treatment emergent. If the event occurs on the day of the initial administration of study intervention, and either event time or time of administration are missing, then the event will be assumed to be treatment emergent. If the event date is recorded as partial or completely missing, then the event will be considered to be treatment emergent unless it is known to be prior to the first administration of study intervention based on partial onset date or resolution date. All reported treatment-emergent adverse events will be included in the analysis. For each adverse event, the number and percentage of participants who experience at least 1 occurrence of the given event will be summarized by intervention group.

Summary tables will be provided for treatment-emergent adverse events:

- AEs
- Serious adverse events (SAEs)
- AEs leading to discontinuation of study agent
- Reasonably related AEs.
- Infections, including infections requiring with an oral or parenteral antibiotic treatment.
- Serious infections.

An infection is defined as any AE that was characterized by the investigator as an infection on the eCRF.

In addition to the summary tables, listings will be provided for participants who:

- Had SAEs
- Had AEs leading to discontinuation of study intervention/termination of study participation

A listing of participants who died will be provided.

Incidence of other treatment-emergent adverse events of [clinical, special] interest will be summarized as necessary.

## 5.6.3. Additional Safety Assessments

### 5.6.3.1. Clinical Laboratory Tests

Clinical laboratory tests will be displayed for the participants included in the safety analysis set. Clinical laboratory values are to be graded using National Cancer Institute Common Terminology Criteria for Adverse Events (NCI-CTCAE) version 5.0 (see [Table 3](#)). Descriptive statistics and graphical displays will be presented for chemistry and hematology at scheduled time points.

Change/percent change from baseline to time point will be summarized for chemistry, hematology, tests and displayed by intervention group.

- Summary of the observed values and changes from baseline over time through the final efficacy & safety visit for clinical laboratory parameters.
- Summary of maximum NCI-CTCAE toxicity grade for postbaseline laboratory values through Week 12 and through the final efficacy & safety visit for clinical laboratory parameters listed in [Table](#) . For toxicity grades, treatment emergent (TE) will be concluded if the postbaseline value is worse than the baseline value.
- Shift tables for selected clinical laboratory parameters (ALT, AST, alkaline phosphatase and total bilirubin) at Week 12.
- For abnormalities based on normal range and/or criteria: If the postbaseline value is above the upper limit and the baseline value is below the upper limit (eg, Normal or Low), then the postbaseline abnormality will be considered TE. The same applies to the postbaseline value being below the lower limit with the baseline value being above the lower limit (eg, Normal or High).
- If the baseline value is missing, a postbaseline abnormality will always be considered as TE.

Listings of participants with any abnormal postbaseline laboratory values of CTCAE grade  $\geq 2$  will also be provided.

The laboratory data to be summarized are as follows:

- **Hematology assessments:** hematocrit, hemoglobin, platelet count, total leukocytes and differential WBC count (neutrophils, lymphocytes, monocytes, eosinophils, and basophils)
- **Blood chemistry assessments:** ALT, AST, alkaline phosphatase, albumin, total and direct bilirubin, blood urea nitrogen, calcium, chloride, creatinine, phosphate, potassium, total protein, and sodium.

The baseline value for a participant is the value closest to but prior to the first dose of study agent (Week 0). Change from baseline is defined to be the assessment at the postbaseline visit minus the assessment at baseline. Summaries of laboratory data will be completed using all the available laboratory data at the time point of interest without imputing missing data. Shift tables summarize the number of participants with low, normal, and high values (determined by the laboratory normal ranges) at the post baseline visit for each of the classifications of low, normal, and high at baseline.

Table 3: Grading Criteria for Clinical Laboratory Tests [CTCAE Version 5.0]

Table 3: Grading Criteria for Clinical Laboratory Tests [CTCAE Version 5.0]					
Hematology Tests		Criteria			
Test	Direction	1	2	3	4
Hemoglobin (g/dL)	Increase	>0 - 2	>2 - 4	>4	
Hemoglobin (g/dL)	Decrease	<LLN - 10.0	<10.0 - 8.0	<8.0	
Lymphocytes (/mm <sup>3</sup> )	Increase		>4000 - 20,000	>20,000	
Lymphocytes (/mm <sup>3</sup> )	Decrease	<LLN - 800	<800 - 500	<500 - 200	<200
Neutrophils (/mm <sup>3</sup> )	Decrease	<LLN - 1500	<1500 - 1000	<1000 - 500	<500
Platelets (/mm <sup>3</sup> )	Decrease	<LLN - 75,000	<75,000 - 50,000	<50,000 - 25,000	<25,000
Total WBC count (/mm <sup>3</sup> )	Increase			>100,000	
Total WBC count (/mm <sup>3</sup> )	Decrease	<LLN - 3000	<3000 - 2000	<2000 - 1000	<1000
Chemistry Tests		Criteria			
Test	Direction	1	2	3	4
ALT	Increase	>ULN - 3.0 x ULN if baseline was normal; 1.5 - 3.0 x baseline if baseline was abnormal	>3.0 - 5.0 x ULN if baseline was normal; >3.0 - 5.0 x baseline if baseline was abnormal	>5.0 - 20.0 x ULN if baseline was normal; >5.0 - 20.0 x baseline if baseline was abnormal	>20.0 x ULN if baseline was normal; >20.0 x baseline if baseline was abnormal
AST	Increase	>ULN - 3.0 x ULN if baseline was normal; 1.5 - 3.0 x baseline if baseline was abnormal	>3.0 - 5.0 x ULN if baseline was normal; >3.0 - 5.0 x baseline if baseline was abnormal	>5.0 - 20.0 x ULN if baseline was normal; >5.0 - 20.0 x baseline if baseline was abnormal	>20.0 x ULN if baseline was normal; >20.0 x baseline if baseline was abnormal
Albumin (g/L)	Decrease	≥30 - <LLN	≥20 - <30	<20	
Alkaline Phosphatase	Increase	>ULN - 2.5 x ULN if baseline was normal; 2.0 - 2.5 x baseline if baseline was abnormal	>2.5 - 5.0 x ULN if baseline was normal; >2.5 - 5.0 x baseline if baseline was abnormal	>5.0 - 20.0 x ULN if baseline was normal; >5.0 - 20.0 x baseline if baseline was abnormal	>20.0 x ULN if baseline was normal; >20.0 x baseline if baseline was abnormal
Bilirubin (total)	Increase	>ULN - 1.5 x ULN if baseline was normal; > 1.0 - 1.5 x baseline if baseline was abnormal	>1.5 - 3.0 x ULN if baseline was normal; >1.5 - 3.0 x baseline if baseline was abnormal	>3.0 - 10.0 x ULN if baseline was normal; >3.0 - 10.0 x baseline if baseline was abnormal	>10.0 x ULN if baseline was normal; >10.0 x baseline if baseline was abnormal
Corrected Calcium* (mmol/L)	Increase	>ULN - ≤2.9	>2.9 - ≤3.1	>3.1 - ≤3.4	>3.4



Corrected Calcium (mmol/L)	Decrease	$\geq 2.0 - < \text{LLN}$	$< 2.0 - \geq 1.75$	$< 1.75 - \geq 1.5$	$< 1.5$
Creatinine	Increase	$> \text{ULN} - \leq 1.5 \times \text{ULN}$	$> 1.5 - 3.0 \times \text{baseline}; > 1.5 - 3.0 \times \text{ULN}$	$> 3.0 \times \text{baseline}; > 3.0 - 6.0 \times \text{ULN}$	$> 6.0 \times \text{ULN}$
Glucose (mmol/L)	Decrease	$< \text{LLN} - 3.0$	$< 3.0 - 2.2$	$< 2.2 - 1.7$	$< 1.7$
Potassium (mmol/L)	Increase	$> \text{ULN} - \leq 5.5$	$> 5.5 - 6.0$	$> 6.0 - 7.0$	$> 7.0$
Potassium (mmol/L)	Decrease	$< \text{LLN} - 3.0$		$< 3.0 - 2.5$	$< 2.5$
Sodium (mmol/L)	Increase	$> \text{ULN} - 150$	$> 150 - 155$	$> 155 - 160$	$> 160$
Sodium (mmol/L)	Decrease	$< \text{LLN} - 130$	$129 - 125$	$124 - 120$	$< 120$

### 5.6.3.2. Vital Signs and Physical Examination Findings

Vital signs are measured at every visit. The listing for participants with clinically important results will be provided as appropriate.

Continuous vital sign parameters collected will be summarized at each assessment time point. Descriptive statistics (mean, standard deviation, median, minimum and maximum) will be presented by intervention groups.

### 5.6.3.3. Electrocardiogram

The listing for participants with clinically important results will be provided as appropriate.

## 5.7. Other Analyses

### 5.7.1. Pharmacokinetics

PK analyses will be based on the Pharmacokinetics Analysis Set (Section 4).

#### 5.7.1.1. PTG-200 Concentrations

Blood, colonic tissue and fecal samples for determining PTG-200 concentrations will be collected from all participants according to the schedule in the Protocol Schedule of Activities. Plasma, colonic tissue and fecal PTG-200 concentrations will be used to evaluate local and systemic drug exposure.

Plasma, colonic tissue and fecal PTG-200 concentrations will be summarized using descriptive statistics (i.e., n, arithmetic mean, standard deviation [SD], coefficient of variation [%CV], median, range [minimum and maximum], and interquartile [IQ] range) at each visit. PK data may be displayed graphically.

The following analyses will be performed as appropriate.

- Summary of plasma PTG-200 concentration over time
- Proportion of participants with plasma PTG-200 concentration below the lowest quantifiable concentration in a sample at each visit
- Summary of plasma PTG-200 concentrations over time by baseline body weight (quartiles). Other covariates may also be applied.
- Plots of median plasma PTG-200 concentrations over time

- Plots of median plasma PTG-200 concentrations over time by baseline body weight ( $\leq$ median,  $>$ median).

As applicable, a similar approach as described above for plasma PTG-200 concentration will also be applied to colonic tissue and fecal PTG-200 concentrations.

### 5.7.1.2. Data Handling Rules

Unless otherwise specified, the following data handling rules will apply to PK sample analyses:

- Plasma concentration summaries will be based on the actual treatment received.
- All plasma, colonic tissue and fecal concentration summaries for a particular timepoint will include data obtained from treated participants at the timepoint of interest without imputing any missing data.
- A concentration not quantifiable (below the lower limit of quantification) will be treated as 0 in the summary statistics and shown as the lower limit of quantification ( $<$  LLOQ) in the data listings.
- The data from a participant who meets 1 of the following dosing deviation criteria will be excluded from the by-visit data analyses from that point onwards:
  - Discontinued study agent administrations
  - Skipped an administration
  - Received an incomplete/ incorrect dose
  - Received an incorrect study agent
  - Received an additional dose
  - Received an administration outside the dosing window specified in [Table](#) .

**Table 4: Dosing Window**

Visit	Window
Week 0 through Week 12	$\pm$ 5 days from scheduled visit day
Final Safety and Efficacy Follow-up visits	$\pm$ 14 days from scheduled visit day

All participants and samples excluded from the analysis will be clearly documented in the Clinical Study Report.

### 5.7.1.3. PK vs Efficacy

To explore the relationship between PTG-200 plasma, colonic tissue or fecal concentrations and efficacy endpoints, the following analyses may be explored:

- The relationship between PTG-200 plasma, colonic tissue or fecal concentrations (quartiles) and change from baseline in CDAI score, clinical response and clinical remission status at

Week 8 or Week 12 may be explored. Summaries by Bio-IR status (yes or no) (i.e., Bio-IR and Bio-NF) and by SNP status (positive or negative) may also be provided.

## **5.7.2. Immunogenicity**

### **5.7.2.1. Antibodies to PTG-200**

Immune response analyses will be based on the Immunogenicity Analysis Set.

The antibody to PTG-200 status (positive at any time, negative) and titers will be summarized by treatment group for participants who receive a dose of PTG-200 and have appropriate samples for detection of antibodies to PTG-200 (ie, participants with at least 1 sample obtained after their first dose of PTG-200). The maximum titers of antibodies to PTG-200 will be provided for participants who are positive for antibodies to PTG-200.

A listing of participants who are positive for antibodies to PTG-200 will be provided. The sample antibody status, the titer, and the neutralizing antibody status to PTG-200 will be listed by visit. This listing will also provide information regarding dose administered, PTG-200 plasma concentration, and CDAI for all applicable visits. In addition, a list of antibody to PTG-200 status in participants who discontinued study agent early will be provided.

### **5.7.2.2. Neutralized Antibodies to PTG-200**

The incidence of neutralizing antibodies (NAbs) to PTG-200 will be summarized for participants who are positive for antibodies to PTG-200 and have samples evaluable for NAbs to PTG-200.

### **5.7.2.3. Antibody vs Efficacy/PK**

Plasma PTG-200 concentrations, clinical remission/response status, and change from baseline in the CDAI score over time may be summarized by antibody to PTG-200 status if sufficient numbers of participants are positive for antibodies. Plots of median plasma PTG-200 concentrations over time by antibody to PTG-200 status may also be provided. The evaluation between colonic tissue PTG-200 concentration and the identified efficacy parameters will be performed as appropriate.

## **5.7.3. Pharmacogenomic**

A whole blood sample collection for SNPs and DNA analysis will be collected from all participants at Screening who have signed an optional pharmacogenomics consent form. DNA testing will be done to search for links of specific genes to disease or response to JNJ-67864238 drug. Exploratory genetic analyses on DNA collection from participants who signed the optional DNA consent will be presented in a separate technical report.

## **5.7.4. Pharmacokinetic/Pharmacodynamic Relationships**

The relationship between plasma concentration of PTG-200 and PD may be analyzed graphically. Similarly, the relationship between colonic tissue or fecal PTG-200 concentration and PD endpoints listed above may be analyzed graphically. If any visual trend is observed, a suitable PK/PD model may be developed to describe the exposure-response relationship. In the latter case, the results of the PK/PD modeling analysis will be presented in a separate technical report.

### **5.7.5. Biomarker Analyses**

Biomarker assessments will be conducted to examine the biological response to treatment and to identify biomarkers that are relevant to PTG-200 treatment and/or Crohn's disease. Assessments include the evaluation of relevant biomarkers in serum, whole blood (for RNA analysis), stool, and mucosal biopsy samples (for protein, RNA and histology analysis) collected according to the Schedule of Activities in the protocol.

Biomarker analyses will characterize the effects of PTG-200 on the measured biomarkers to identify biomarkers relevant to treatment and to determine if these biomarkers can predict response to PTG-200. Results will be reported in separate technical reports.

### **5.7.6. Health Economics**

#### **5.7.6.1. PROMIS 29**

The PROMIS-29 profile instrument is intended for adults (ages 18 years and older). It is a collection of short forms containing 4 items for each of 7 PROMIS domains: Depression, Anxiety, Physical Function, Pain Interference, Fatigue, Sleep Disturbance, and Ability to Participate in Social Roles and Activities. The PROMIS-29 instrument also includes a Pain Intensity 0-10 NRS in addition to the 7 domains. PROMIS-29 is a universal rather than disease-specific instrument. It assesses all domains over the past 7 days, except for the Physical Function domain, which has no specified timeframe. The raw score of each domain is converted into a standardized score with a mean of 50 and a standard deviation (SD) of 10 (T-Score). The standardized T-score is reported as the final score for each participant. Pain Intensity is presented as raw responses (0-10). For the PROMIS domains of Depression, Anxiety, Physical Function, Pain Interference, and Fatigue, a score of 50 is the mean for the United States general population with an SD of 10 (validated with a large sample of the general US population). Since the other 2 domains (Ability to Participate in Social Roles and Activities and Sleep Disturbance) were not validated with a national sample, a score of 50 represents the mean of the calibration sample (which was enriched for chronic illness so a score of 50 represents a sicker population). A higher PROMIS T-score represents more of the domain being measured. For negatively-worded domains like Anxiety, a T-score of 60 is 1 SD worse than average. By comparison, an Anxiety T-score of 40 is 1 SD better than average. However, for positively-worded domains like Physical Function, a T-score of 60 is better than average.

PROMIS surveys will only be collected in countries where translations are available.

#### **5.7.6.2. Analysis Method**

The analysis will be performed at Week 12. The analysis will include descriptively summarizing change from baseline in PROMIS-29 score. Using MMRM model in comparing treatment difference by LS Mean (SD), LS mean (90% CI). p-Value was based on comparison with placebo model. And other statistical analysis as appropriate.

### **5.7.7. Other Variables and/or Parameters**

Not applicable.

### 5.7.8. Definition of Subgroups

Subgroups of interest in this study include but not limited to:

1. Population
  - a. BIO-IR vs BIO-NF
2. Demographics at baseline:
  - a. Age ( $\leq$  median age,  $>$ median age)
  - b. Sex (male, female)
  - c. Race (White, non-White)
  - d. Weight at baseline ( $\leq$  median,  $>$  median;  $\leq$  1<sup>st</sup> quartile,  $>$ 1<sup>st</sup> quartile and  $\leq$  2<sup>nd</sup> quartile,  $>$ 2<sup>nd</sup> quartile and  $\leq$  3<sup>rd</sup> quartile,  $>$ 3<sup>rd</sup> quartile)
  - e. Center location (North America, Rest of World)
3. Disease characteristics at baseline:
  - a. Crohn's disease duration ( $\leq$  5 years,  $>$  5 years to  $\leq$  15 years, or  $>$  15 years)
  - b. Involved gastrointestinal areas (ileum only, colon only, ileum & colon)
  - c. CDAI score ( $<$ 300,  $\geq$ 300 )
  - d. CRP ( $\leq$  3 mg/L,  $>$  3 mg/L)
  - e. Fecal Calprotectin ( $\leq$ 250  $\mu$ g/g,  $>$  250  $\mu$ g/g)
4. Concomitant Crohn's-disease medication at baseline:
  - a. Oral corticosteroids (including budesonide and beclomethasone dipropionate) (receiving, not receiving)
  - b. Oral 5-ASA compounds (receiving, not receiving)
  - c. Immunomodulators (6-MP/AZA/MTX) (receiving, not receiving)
  - d. Oral corticosteroids (including budesonide and beclomethasone dipropionate) and Immunomodulators (receiving, not receiving)
5. Other:
  - a. Current fistula at baseline (yes, no)

### 5.8. Interim Analyses

An interim analysis is planned when approximately 50% of planned Participants reach their Week 12 visit, which is approximately 45 Participants who received at least one administration of study agent (including a partial dose), have completed their Week 12 visit, or have terminated their study participation before Week 12. Additional ad hoc IA(s) may be conducted if deemed necessary. The objective of the IA(s) is to enable early termination of the intervention cohort if JNJ-67864238 is futile and plan for future intervention cohorts. At the time of the IA(s), the study team will remain blinded. No DBL is planned for the IA. The statistical model specified in master protocol Section 9.4.2 and this document Section 5 will be used to evaluate the treatment difference at Week

12 using all available data. The interim analysis criteria are discussed in attachment 1. Interim analyses will be based on data that are as clean as possible.

An internal interim analysis committee independent of the study team will be established to review the interim data and formulate recommended decisions/actions in accordance with predefined decision rules. The details of interim analyses procure will be included in a separate interim analysis plan.

#### Interim Analysis Decision-Making Rules:

If neither change in CDAI nor change in SES-CD of the JNJ-67864238 treatment arm meets the success criteria, the futility of the JNJ-67864238 treatment arm will be considered. The futility is defined as follows:

CCI

CCI

- Futility in change in CDAI: if upper bound (UB) (95%)  $< 60$  and lower bound (LB) (10%)  $< 40$ , ie, the 95% upper confidence bound of the true mean treatment difference at Week 12 is less than 60 and its 90% lower confidence bound less than 40; and
- Futility in change in SES-CD: if UB (95%)  $< 1.6$  and LB (10%)  $< 0.4$ , ie, the 95% upper confidence bound of the true mean treatment difference at Week 12 is less than 1.6 and its 90% lower confidence bound less than 0.4.

If the futility criteria are met, the decision to terminate enrollment will be made based on a benefit risk assessment of the totality of data, including overall efficacy assessments, analysis of biomarker and PD data, and a safety review by the external data monitoring committee (DMC). If the futility criteria are not met and the DMC has not raised any safety issue, enrollment will continue until the planned sample size is reached. To augment decision-making, all available data at this IA, including from those participants without Week 12 assessments, will be analyzed. An Interim Analysis Committee will review the unblinded interim data and formulate recommendations (Section 9.5.1 of the Master protocol PLATFORMPACRD2001).

#### 5.8.1. Data Monitoring Committee (DMC) or Other Review Board

An external DMC will be established to monitor safety data on an ongoing basis to ensure the continuing safety of participants enrolled in the study as noted in Committees Structure in platform protocol Appendix 5, Regulatory, Ethical, and Study Oversight Considerations. Please refer to DMC charter for details.

## 6. SUPPORTING DOCUMENTATION

### 6.1. Appendix 1 List of Abbreviations

5-ASA	5-aminosalicylate
6-MP	6-mercaptopurine
AE	adverse event
ANCOVA	analysis of covariance
AZA	Azathioprine
Bio-IR	biologic intolerant or refractory
Bio-NF	biologic nonfailure, ie, inadequate response to or failed to tolerate corticosteroids or immunomodulators, but not a biologic
CDAI	Crohn's Disease Activity Index
CI	confidence interval
CRF	case report form
CRP	C-reactive protein
DPS	Data presentation specifications
MedDRA	Medical Dictionary for Regulatory Activities
MTX	Methotrexate
SAE	serious adverse event
SAP	Statistical Analysis Plan
SDC	Sponsor Decision Committee
SES-CD	Simple Endoscopic Score for Crohn's Disease
SNP	Single Nucleotide Polymorphism
SSG	Statistical Support Group
TNF	tumor necrosis factor
ISA	Intervention-Specific Appendix
DPS	Data Presentation Specifications
PK	pharmacokinetics
PD	pharmacodynamics
DBL	database lock
DMC	data monitoring committee
IA	interim analysis

## **6.2. Appendix 2 Changes to Protocol-Planned Analyses**

Based on protocol amendment to incorporate COVID-19 infection event, for primary endpoint point analysis, intercurrent events (ICEs) caused by COVID-19 infection will be addressed by treatment police strategy for primary estimator. In addition, hypothetical strategy to assume missing at random after this ICE and tipping point method will be applied to assess the impact after COVID-19 ICEs in a sensitivity analysis as appropriate. Please refer to section 5.3.2 for details. For subject information and safety analyses, COVID-19 infection event will be tabulate in the corresponding tables and listings.



### 6.3. Appendix 3 Demographics and Baseline Characteristics

The number of participants in each analysis set will be summarized and listed by intervention group, and overall. In addition, the distribution of participants by region, country, and site ID will be presented unless otherwise noted.

Table 5 presents a list of the demographic variables that will be summarized by intervention group, and overall. Demographics will also be summarized by region using the full analysis set.

To assess the comparability of participants, the following data will be summarized by treatment group based on the randomized analysis set:

- Participants' demographic data (i.e. age, gender, race, height, weight, and region)
- Baseline disease characteristics (including endoscopic disease), and baseline use and history of Crohn's disease medications
- Medical history of interest

**Table 5: Demographic and Baseline Disease Characteristic Variables**

<b>Continuous Variables:</b>	<b>Summary Type</b>
Age (years)	Descriptive statistics (N, mean, standard deviation [SD], median and range [minimum and maximum], and IQ range).
Weight (kg)	
Height (cm)	
Crohn's Disease Duration (yrs)	
CDAI	
PRO-2 (defined as the sum of the abdominal pain and stool frequency subscores of the CDAI score)	
SES-CD	
CRP Concentration (mg/L)	
Fecal Calprotectin (mg/kg)	
<b>Categorical Variables</b>	
Race ([White, Black or African American, Asian, American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, Other, Not Reported and Unknown])	
Sex ([Male, Female])	
Involved GI areas ([Ileum only, Colon only, Ileum and colon, Proximal gastrointestinal tract, and Perianal])	
Extra intestinal manifestations ([Absent, Present])	
Endoscopic disease severity (SES-CD score) ([Remission (0-2), Mild inflammation (3-6), Moderate inflammation (7-16) and Severe inflammation (>16)])	
Abnormal CRP (>3 mg/L) ([Yes, No])	
Abnormal fecal calprotectin (>250 mg/kg) ([Yes, No])	
Crohn's disease complications (Intra-abdominal abscess, Sinus tracts /perforation, Fistula, Bowel Strictureing]	

Study Day and Relative Day: Study Day 1 or Day 1 refers to the day of the first study agent administration on or after the randomization date. All efficacy and safety assessments at all visits will be assigned a day relative to this date.

Study day or relative day for a visit is defined as:

- Visit date - (date of Study Day 1) +1, if visit date is  $\geq$  date of Day 1
- Visit date - Date of Day 1, if visit date < date of Day 1

There is no 'Day 0'.

Baseline: Baseline is defined as the last observation prior to the first study agent administration unless otherwise specified.

#### **6.4. Appendix 4 Protocol Deviations**

Major protocol deviations will be summarized by treatment group in the following categories: participants who entered the trial but did not satisfy entry criteria, participants who received the wrong medication or incorrect dose, participants who received disallowed medication, and other as appropriate.

**6.5. Appendix 5 Prior and Concomitant Medications**

Summaries of Crohn's-specific concomitant medications at baseline (including corticosteroids, Immunomodulators, antibiotics and Aminosalicylates) will be summarized.

**6.6. Appendix 6 Medical History**

Medical History will be summarized as needed..

## 6.7. Appendix 7 Intervention Compliance

The number of Participants receiving each scheduled administration will be summarized by treatment group.

Drug compliance is calculated as the percent of number of tablets actually taken during the course of study. Specifically, for patient  $i$

$$Compliance_i = \left( \frac{\text{Actual number of tablets taken}}{\text{Total number of tablets supposed to be taken}} \right) \times 100$$

The number and percentage of patients who completed the specified range of compliance rate, starting at  $\leq 75\%$  compliance to a maximum of 100% in increments of 5% will be summarized by treatment as appropriate. For Participants who discontinue study agent prior to week 12, the expected number of tablets is counted through the time of study agent discontinuation.

## 6.8. Appendix 8 Adverse Events of Special Interest

Adverse Events of Special Interest will be summarized as needed. Any newly identified malignancy or case of active TB occurring after the first study intervention administration(s) in participants participating in this clinical study are considered AEs of special interest.

A **possible Hy's law case** is defined by the occurrence of ALT/AST  $\geq 3$  x the upper limit of normal (ULN), alkaline phosphatase  $< 2$  x ULN together with Tbili  $\geq 2$  x ULN or international normalized ratio  $> 1.5$  (if measured). Any possible Hy's Law case is considered an important medical event and should be reported to the sponsor in an expedited manner using the adverse events of special interest form, even before all other possible causes of liver injury have been excluded ([FDA 2009](#)). A confirmed Hy's law case must be reported as a SAE.

**6.9. Appendix 9 Medications of Special Interest**

Not applicable.



## **6.10. Appendix 10 Laboratory Toxicity Grading**

Laboratory Toxicity Grading related analysis will be performed as needed.

The grading scale use for lab assessments is based on ‘Common Terminology Criteria for Adverse Events (CTCAE) v5.0’.

If a laboratory value falls within the grading as specified below but also within the local laboratory normal limits, the value is considered to be normal and will be reset to grade 0.

Pre-baseline measurements will use the same grading ranges as applied to baseline measurements. In case a test has two sets of ranges – one for baseline normal and one for baseline abnormal, the one for baseline normal will be applied for all measurements taken pre-baseline and on baseline.

## 7. REFERENCES

1. Best WR, Beckett JM, Singleton JW, Kern F Jr. Development of a Crohn's disease activity index. National Cooperative Crohn's Disease Study. *Gastroenterology*. 1976;70(3):439-444.
2. D'haens G, Van Deventer S, Van Hogezaand R, et al. Endoscopic and histological healing with infliximab anti-tumor necrosis factor antibodies in Crohn's disease: A European multicenter trial. *Gastroenterology*. 1999;116(5):1029-1034.
3. D'Haens GR, Geboes K, Peeters M, Baert F, Penninckx F, Rutgeerts P. Early lesions of recurrent Crohn's disease caused by infusion of intestinal contents in excluded ileum. *Gastroenterology*. 1998;114(2):262-267.
4. Irvine EJ, Feagan B, Rochon J, et al. Quality of life: a valid and reliable measure of therapeutic efficacy in the treatment of inflammatory bowel disease. Canadian Crohn's Relapse Prevention Trial Study Group. *Gastroenterology*. 1994; 106(2):287-296.
5. Laharie D, Reffet A, Belleannée G, et al. Mucosal healing with methotrexate in Crohn's disease: a prospective comparative study with azathioprine and infliximab. *Aliment Pharmacol Ther*. 2011;33(6):714-721.
6. Mantzaris GJ, Christidou A, Sfakianakis M, et al. Azathioprine is superior to budesonide in achieving and maintaining mucosal healing and histologic remission in steroid-dependent Crohn's disease. *Inflamm Bowel Dis*. 2009;15(3):375-382.
7. Mojtahed A, Khanna R, Sandborn WJ, et al. Assessment of histologic disease activity in Crohn's disease: a systematic review. *Inflamm Bowel Dis*. 2014;20(11):2092-2103.

## Attachment 1 Decision-Making Framework

There are 2 decision points planned for this ISA protocol:

- At the IA: This IA will take place when approximately 45 participants have completed their Week 12 assessments or have withdrawn from the study before Week 12. If the futility criteria are met, the decision to terminate enrollment will be made based on a benefit risk assessment of the totality of data, including overall efficacy assessments, analysis of biomarker and PD data, and a safety review by the external DMC. If the futility criteria are not met and the DMC has not raised any safety issue, enrollment will continue until the planned sample size is reached. To augment decision-making, all available data at this IA, including from those participants without Week 12 assessments, will be analyzed. An Interim Analysis Committee will review the unblinded interim data and formulate recommendations (Section 9.5.1 of the Master protocol PLATFORMPACRD2001).
- At the completion of this intervention cohort: The overall study success or futility is defined in the Master protocol. This decision-making framework is suggestive but not binding. The final decision will be made based on the totality of data.

### Decision-Making Specifics at the Interim Analysis

If neither change in CDAI nor change in SES-CD of the PTG-200 CCI treatment arm meets the success criteria, the futility of the PTG-200 CCI treatment arm will be considered. The futility is defined as follows:

- Futility in change in CDAI: if upper bound (UB) (95%)  $< 60$  and lower bound (LB) (10%)  $< 40$ , ie, the 95% upper confidence bound of the true mean treatment difference at Week 12 is less than 60 and its 90% lower confidence bound less than 40; and
- Futility in change in SES-CD: if UB (95%)  $< 1.6$  and LB (10%)  $< 0.4$ , ie, the 95% upper confidence bound of the true mean treatment difference at Week 12 is less than 1.6 and its 90% lower confidence bound less than 0.4.

The thresholds chosen for futility at the IA correspond to PTG-200 profile with limited benefit.

### Decision-Making Specifics at the Completion of the Intervention Cohort

The success of the PTG-200 CCI treatment arm is defined as follows:

- Success in change in CDAI: if LB (20%)  $\geq 40$ , ie, the 80% lower confidence bound of the true mean treatment difference at Week 12 is at least 40; or
- Success in change in SES-CD: if LB (20%)  $\geq 0.4$ , ie, the 80% lower confidence bound of the true mean treatment difference at Week 12 is at least 0.4.

The threshold of 40 is selected to provide confidence that a therapy is at least as effective as currently approved therapies (STELARA® [ustekinumab] or adalimumab), where a CDAI change of 40 represents the lower confidence bound for success of the treatment difference between JNJ-67864238 and placebo.

If neither change in CDAI nor change in SES-CD of the PTG-200 CCI treatment arm meets the success criteria, the futility of the PTG-200 CCI treatment arm will be considered. The futility of is defined as follows:

- Futility in change in CDAI: if UB (80%)  $< 60$  and LB (20%)  $< 40$ , ie, the 80% upper confidence bound of the true mean treatment difference at Week 12 is less than 60 and its 80% lower confidence bound less than 40; and
- Futility in change in SES-CD: if UB (90%)  $< 1.6$  and LB (20%)  $< 0.4$ , ie, the 90% upper confidence bound of the true meantreatment difference at Week 12 is less than 1.6 and its 80% lower confidence bound less than 0.4.

**Attachment 2: Data Simulation**

To evaluate the operating characteristics of the study design, the following assumptions are made based on 3 completed clinical studies (C0743T26/CNTO1275 Ph2b, CNTO1275CRD3001, and CNTO1275CRD3002) conducted by the sponsor:

- The compositions of Bio-IR and Bio-NF populations are 50% each
- For Bio-IR population, the CDAI scores of the JNJ-67864238 [CCI] treatment arm and placebo treatment arm at Weeks 0, 2, 4, 8, and 12 are assumed to be (323,285,259,247,238) and (323,305,300,299,298), respectively. Similarly, for Bio-NF population, the CDAI scores of the JNJ-67864238 [CCI] treatment arm and placebo treatment arm at Weeks 0, 2, 4, 8 and 12 are assumed to be (302,245,208,180,176) and (302,271,254,241,236), respectively.
- The covariance matrix is assumed to be

$$\begin{pmatrix} 3844 & 3069 & 3100 & 3193 & 3255 \\ 3069 & 8100 & 7200 & 6489 & 5670 \\ 3100 & 7200 & 10000 & 7210 & 5250 \\ 3193 & 6489 & 7210 & 10609 & 7030 \\ 3255 & 5670 & 5250 & 7030 & 11025 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 3600 & 2970 & 2565 & 3646 & 2520 \\ 2970 & 8100 & 5558 & 5670 & 4725 \\ 2565 & 5558 & 9025 & 6484 & 4988 \\ 3646 & 5670 & 6484 & 11025 & 7166 \\ 2520 & 4725 & 4988 & 7166 & 11025 \end{pmatrix}$$

for the Bio-IR and Bio-NF populations, respectively.

- In CNTO1275CRD3001 and CNTO1275CRD3002, an optional endoscopy substudy was conducted in a subset of study participants. Based on the collected data, change in SES-CD is assumed to have a mixture distribution - excessive zeros (no change) and a normal distribution. Therefore, 2 steps were utilized to simulate the change in SES-CD data.
  - Step 1: no change in SES-CD is simulated for 19% of JNJ-67864238 [CCI] treated participants and 32% placebo treated participants.
  - Step 2: changes in SES-CD based on the normal distributions with mean changes from baseline of -3.26 and -0.22 are generated for the JNJ-67864238 [CCI] and placebo arms, respectively. The common standard deviation is 3.8. As the change in SES-CD of a participant is an integer, the simulated value then is rounded to the nearest integer. The above assumptions are for the Bio-IR population. The change in SES-CD responses for the Bio-NF population can be simulated similarly with the following assumptions: the proportion of the excessive zeros is 12% for both the active and placebo treatment arm in Step 1. The means of the normal distributions are -4.68 and -2.71 for the PTG-200 [CCI] and placebo arms, respectively. The common standard deviation is 5.9.
- The correlation coefficients between the change in SES-CD and the CDAI scores are assumed to be -0.15, 0, 0, 0.15, and 0.15 at Weeks 0, 2, 4, 8, and 12 for the Bio-IR population. They are assumed to be 0, 0, 0, 0.15 and 0.15 at Weeks 0, 2, 4, 8, and 12 for the Bio-NF population.

## Attachment 3: Crohn's Disease Activity Index

<u>DISEASE ACTIVITY INDEX</u>	<u>SUM</u>	<u>X FACTOR</u>	<u>SUBTOTAL</u>
Total number of liquid or very soft stools in the previous 7 days		x 2 =	
Sum abdominal pain/cramps ratings (total for previous 7 days): 0 = none      2 = moderate 1 = mild      3 = severe		x 5 =	
General well being (total for previous 7 days): 0 = generally well      3 = very poor 1 = slightly under par      4 = terrible 2 = poor		x 7 =	
Categories currently present and presumed to be related to Crohn's disease: 0 = no; 1 = yes			
Y = arthritis/arthralgia		x 20 =	
Y = iritis/uveitis		x 20 =	
Y = erythema nodosum/pyoderma gangrenosum/aphthous stomatitis		x 20 =	
Y = anal fissure, fistula or abscess		x 20 =	
Y = other fistula		x 20 =	
Y = fever over 100° F (37.8° C) during the previous 7 days.		x 20 =	
During the previous 7 days has participant received antidiarrheal therapy at least once:			
<b>OR</b>			
During the previous 7 days has participant received opiate therapy on each of the 7 days: 0 = no 1 = yes		x 30 =	
Abdominal mass: 0 = none    2 = questionable    5 = definite		x 10 =	
Hematocrit: Males: (47-Hct) = SUM Females: (42-Hct) = SUM		x 6 = (add or subtract by sign)	
		*	
$\frac{(\text{Standard Weight} - \text{Actual Body Weight})}{\text{Standard Weight}} \times 100 =$		x 1 =	
		(add or subtract by sign, round to 3 decimal places)	
* If this value is less than -10 then enter -10 here. Standard weight and actual weight must be in same units (kg or lb)			
		TOTAL =	
		(round total to integer)	

**Attachment 4: Sample score sheet and scoring definitions for the Simple Endoscopic Score for Crohn's Disease (SES-CD)**

	Ileum	Right Colon	Transverse Colon	Left Colon	Rectum	Total
<b>1. Presence and size of ulcers (0-3)</b>						15 max
<b>2. Extent of ulcerated surface (0-3)</b>						15 max
<b>3. Extent of affected surface (0-3)</b>						15 max
<b>4. Presence and type of narrowings (0-3)</b>						11 max*
Total 1 + 2 + 3 + 4 =						SES-CD (56 max)

	Score = 0	Score = 1	Score = 2	Score = 3
<b>Size of ulcers</b>	None	Aphthous ulcers ( $\varnothing$ 0.1 – 0.5 cm)	Large ulcers ( $\varnothing$ 0.5 – 2.0 cm)	Very large ulcers ( $\varnothing$ > 2.0 cm)
<b>Ulcerated surface</b>	None	<10%	10-30%	>30%
<b>Affected surface</b>	Unaffected segment	<50%	50-75%	>75%
<b>Narrowings</b>	None	Single, can be passed	Multiple, can be passed	Cannot be passed

\* The maximum sub-score for narrowings (i.e. stricturing) is 11 points. The presence of a narrowing that cannot be passed can be only observed once.  
 $\varnothing$  =Diameter.

## Attachment 5: STANDARD WEIGHT TABLE

Actual Height Inches (cm)	Standard Weight in Pounds Men (kg)	Standard Weight in Pounds Women (kg)
58.0 (147.3)		115.0 (52.2)
58.5 (148.6)		116.0 (52.6)
59.0 (149.9)		117.0 (53.1)
59.5 (151.1)		118.3 (53.6)
60.0 (152.4)		119.5 (54.2)
60.5 (153.7)		120.8 (54.8)
61.0 (154.9)		122.0 (55.3)
61.5 (156.2)		123.5 (56.0)
62.0 (157.5)	136.0 (61.7)	125.0 (56.7)
62.5 (158.8)	137.0 (62.1)	126.5 (57.4)
63.0 (160.0)	138.0 (62.6)	128.0 (58.0)
63.5 (161.3)	139.0 (63.0)	129.5 (58.7)
64.0 (162.6)	140.0 (63.5)	131.0 (59.4)
64.5 (163.8)	141.3 (64.1)	132.5 (60.1)
65.0 (165.1)	142.5 (64.6)	134.0 (60.8)
65.5 (166.4)	143.8 (65.2)	135.5 (61.4)
66.0 (167.6)	145.0 (65.8)	137.0 (62.1)
66.5 (168.9)	146.5 (66.4)	138.5 (62.8)
67.0 (170.2)	148.0 (67.1)	140.0 (63.5)
67.5 (171.5)	149.5 (67.8)	141.5 (64.2)
68.0 (172.7)	151.0 (68.5)	143.0 (64.9)
68.5 (174.0)	152.5 (69.2)	144.5 (65.5)
69.0 (175.3)	154.0 (69.8)	146.0 (66.2)
69.5 (176.5)	155.5 (70.5)	147.5 (66.9)
70.0 (177.8)	157.0 (71.2)	149.0 (67.6)
70.5 (179.1)	158.5 (71.9)	150.5 (68.3)
71.0 (180.3)	160.0 (72.6)	152.0 (68.9)
71.5 (181.6)	161.8 (73.4)	153.5 (69.6)
72.0 (182.9)	163.5 (74.1)	155.0 (70.3)
72.5 (184.2)	165.3 (75.0)	
73.0 (185.4)	167.0 (75.7)	
73.5 (186.7)	169.0 (76.6)	
74.0 (188.0)	171.0 (77.5)	* Height in shoes with one-inch heels
74.5 (189.2)	172.8 (78.4)	* Indoor clothing weighing 5 pounds for men and 3 pounds for women
75.0 (190.5)	174.5 (79.1)	
75.5 (191.8)	176.8 (80.2)	* Centimeters x 0.3937 = inches
76.0 (193.0)	179.0 (81.2)	* Pounds x 0.4535 = kilograms