Statistical Analysis Plan for Study M16-006

A Multicenter, Randomized, Double-Blind, Placebo-Controlled Induction Study of the Efficacy and Safety of Risankizumab in Subjects with Moderately to Severely Active Crohn's Disease

Date: 23 November 2020

Version 5.0

Table of Contents

List of Tables

List of Figures

List of Appendices

1.0 Introduction

This Statistical Analysis Plan (SAP) describes the statistical analyses for risankizumab Study M16-006 "A Multicenter, Randomized, Double-Blind, Placebo Controlled Induction Study of the Efficacy and Safety of Risankizumab in Subjects with Moderately to Severely Active Crohn's Disease."

Pharmacokinetics/pharmacodynamics, pharmacogenetics, and selected biomarkers will be analyzed separately and are not included in this SAP.

The SAP will not be updated in case of administrative changes or amendments to the protocol unless the changes impact the analysis.

Unless noted otherwise, all analyses will be performed using SAS Version 9.4 (SAS Institute Inc., Cary, NC 27513) or later under the UNIX operating system.

This SAP includes changes to analyses described in the protocol. Details are outlined in Section 12.0.

2.0 Study Design and Objectives

2.1 Objectives, Hypotheses and Estimands

The objective of this study is to evaluate the efficacy and safety of risankizumab versus placebo during induction therapy in subjects with moderately to severely active Crohn's Disease (CD).

Primary Efficacy Objective

- **For US-specific protocol:**

The **primary efficacy objective** of the study for US-specific protocol is to demonstrate a higher rate of a) CDAI clinical remission (defined as Crohn's disease activity index [CDAI] < 150), and b) endoscopic response (defined as decreasing in Simple Endoscopic

Score – CD [SES-CD] > 50% from baseline [or for subjects with isolated ileal disease and a baseline SES-CD of 4, at least a 2 point reduction from baseline], as scored by central reviewer) after 12 weeks of treatment with risankizumab when compared to placebo based on Intent-to-Treat (ITT) population ITT1A (see Section 4.0).

Hypotheses corresponding to the primary efficacy objective and endpoints for US-specific protocol are:

- The proportion of subjects achieving CDAI clinical remission treated with risakizumab is greater than that treated with placebo at Week 12
- The proportion of subjects achieving endoscopic response treated with risakizumab is greater than that treated with placebo at Week 12

The **estimands** corresponding to the primary efficacy objective for US-specific protocol are defined as follows:

- Difference in the percentage of subjects achieving CDAI clinical remission at Week 12 regardless of premature discontinuation of study drug and without initiation or dose escalation of CD-related corticosteroids (See Section 8.2.2) in each of the risankizumab 1200 mg intravenous (IV) and risankizumab 600 mg IV groups in comparison with the placebo group in the ITT1A population
- Difference in the percentage of subjects achieving endoscopic response at Week 12 regardless of premature discontinuation of study drug and without initiation or dose escalation of CD-related corticosteroids (See Section 8.2.2) in each of the risankizumab 1200 mg IV and risankizumab 600 mg IV groups in comparison with the placebo group in the ITT1A population.

- **For global protocol outside US:**

The **primary efficacy objective** of the study for global protocol outside US is to demonstrate a higher rate of a) clinical remission (defined as average daily stool frequency [SF] ≤ 2.8 and not worse than baseline AND average daily abdominal pain [AP] score ≤ 1 and not worse than baseline), and b) endoscopic response after 12 weeks of

treatment with risankizumab when compared to placebo based on Intent-to-Treat (ITT) population ITT1A.

Hypotheses corresponding to the primary efficacy objective and endpoints for global protocol outside US is:

- The proportion of subjects achieving clinical remission treated with risakizumab is greater than that treated with placebo at Week 12
- The proportion of subjects achieving endoscopic response treated with risakizumab is greater than that treated with placebo at Week 12

The **estimands** corresponding to the primary efficacy objective for global protocol outside US are defined as follows:

- Difference in the percentage of subjects achieving clinical remission at Week 12 regardless of premature discontinuation of study drug and without initiation or dose escalation of CD-related corticosteroids (See Section 8.2.2) in each of the risankizumab 1200 mg IV and risankizumab 600 mg IV groups in comparison with the placebo group in the ITT1A population
- Difference in the percentage of subjects achieving endoscopic response at Week 12 regardless of premature discontinuation of study drug and without initiation or dose escalation of CD-related corticosteroids (See Section 8.2.2) in each of the risankizumab 1200 mg IV and risankizumab 600 mg IV groups in comparison with the placebo group in the ITT1A population.

Secondary Efficacy Objectives

For both US-specific protocol and global protocol outside US:

The **secondary efficacy objectives** of this study are to demonstrate higher efficacy of treatment with risankizumab when compared to placebo with respect to the ranked secondary endpoints specified in Section 3.2 under each protocol. The ranked secondary efficacy objectives will be assessed based on ITT1A population.

Hypotheses corresponding to the secondary efficacy objectives and endpoints are:

1. For each of the ranked binary secondary endpoints (Section 3.2), greater proportion of subjects with improvement for the endpoint is achieved with risankizumab when compared to that of placebo;
2. For each of the ranked continuous endpoints (Section 3.2), greater mean change from baseline for the endpoint is achieved with risankizumab when compared to that of placebo.

For each of the binary ranked secondary endpoints except for CD-related hospitalization endpoint, the **estimands** corresponding to the secondary efficacy objectives are defined as follows: Difference in the percentage of subjects achieving the endpoint regardless of premature discontinuation of study drug and without initiation or dose escalation of CD-related corticosteroids (See Section 8.2.2) in each of the risankizumab 1200 mg IV and risankizumab 600 mg IV groups in comparison with the placebo group in the ITT1A population.

For the secondary endpoint of "Occurrence of CD-related hospitalization through Week 12," the **estimand** is defined as follows: Difference in the percentage of subjects having CD-related hospitalization through Week 12 regardless of premature discontinuation of study drug and regardless of initiation or dose escalation of CD-related corticosteroids (See Section 8.2.2) in each of the risankizumab 1200 mg IV and risankizumab 600 mg IV groups in comparison with the placebo group in the ITT1A population.

For each of the continuous ranked secondary endpoints, the **estimands** corresponding to the secondary efficacy objectives are defined as follows: Difference in the mean change from baseline for the endpoint regardless of premature discontinuation of study drug and if subjects would not initiate or escalate dose of CD-related corticosteroids (See Section 8.2.2) in each of the risankizumab 1200 mg IV and risankizumab 600 mg IV groups in comparison with the placebo groups in the ITT1A population.

2.2 Study Design Overview

This is a Phase 3, multicenter, randomized, double-blind, placebo-controlled study designed to evaluate the efficacy and safety of risankizumab as an induction therapy in subjects with moderately to severely active CD.

The study consists of the following periods:

1. Screening period of up to a maximum of 35 days
2. 12-week randomized, double-blinded IV therapy induction period (12-Week Induction Period)
3. 12-week randomized, double-blinded, double-dummy Induction Period 2 (Induction Period 2) for subjects who do not achieve clinical response at Week 12
4. 20-week follow up period from the last dose of study drug.

The schematic of the study is shown in Figure 1.

Figure 1. Study Schematic

a. 12-Week Induction Period

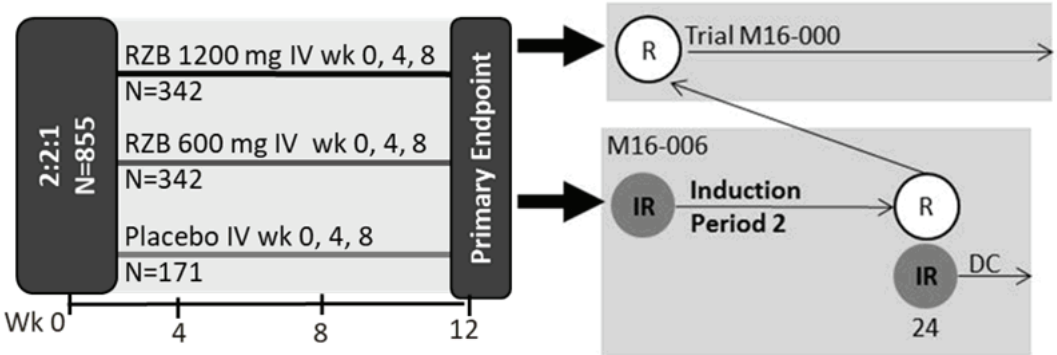
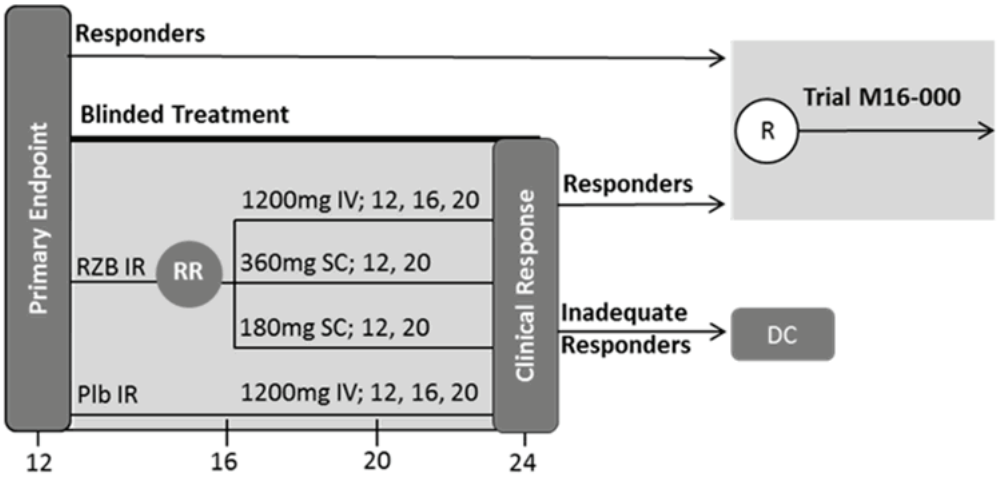


Figure 1. Study Schematic (Continued)

b. Induction Period 2



DC = discontinued; IR = subjects with inadequate clinical response to induction; IV = intravenous; R = subjects with clinical response; RR = re-randomize

Approximately 855 patients with moderate to severe CD who had baseline eligible SES-CD (eligible SES-CD is defined as total SES-CD score excluding presence of narrowing subscore) of ≥ 6 (≥ 4 for isolated ileal disease) are to be randomized in a ratio of 2:2:1 to one of the three following treatment groups in the 12-Week Induction Period.

- Placebo IV (n = 342)
- Risankizumab 600 mg IV (n = 342)
- Risankizumab 1200 mg IV (n = 171)

In addition to the 855 subjects, up to 85 subjects with a lower SES-CD, defined as an eligible SES-CD of ≥ 3 to < 6 for ileocolonic or colonic disease or SES-CD of 3 for isolated ileal disease will be enrolled in order to inform treatment effect in this patient population, but these subjects will not be included in the primary ITT population for efficacy analysis.

Each treatment group receives the corresponding dose of risankizumab or placebo IV infusion at Week 0, Week 4, and Week 8. At Week 12, patients are evaluated for clinical response, defined as ≥ 30% decrease in average daily SF and/or ≥ 30% decrease in average daily AP score and both not worse than baseline.

Visits for clinical evaluation will occur at baseline/Week 0, Weeks 4, 8, and 12 or Premature Discontinuation (PD).

Subjects, who do not achieve clinical response at Week 12 will be randomized by Interactive Response Technologies (IRT) to the Induction Period 2, a double-blind, double dummy 12-week treatment period. Subjects who received risankizumab induction treatment during the 12-Week Induction Period will be randomized 1:1:1 to:

- Risankizumab 1200 mg IV
- Risankizumab 360 mg subcutaneous (SC)
- Risankizumab 180 mg subcutaneous (SC)

Subjects who received placebo induction treatment during the 12-Week Induction Period will receive:

- Risankizumab 1200 mg IV

The IV risankizumab dose or matching IV placebo will be given at Weeks 12, 16, and 20; The SC risankizumab dose or matching SC placebo will be given at Weeks 12, and 20.

At Week 24, subjects who participated in the blinded Induction Period 2 will be reassessed and undergo a third endoscopy for evaluation of mucosal inflammation. Subjects who achieve clinical response at Week 24 may be eligible to enter the Study M16-000. Subjects without clinical response at Week 24, as well as all subjects who are terminated the study early (including subjects who are eligible for but do not participate in the blinded Induction Period 2), will be discontinued and have a follow-up call 140 days

from the last dose of study drug to obtain information on any new and/or ongoing adverse events (AEs).

2.3 Treatment Assignment and Blinding

At baseline, eligible patients will be randomized to three parallel groups with 1200 mg IV and 600 mg IV dose of risankizumab treatments or placebo in a 2:2:1 ratio. The 12-Week Induction Period randomization will be stratified by number of prior biologics failed (0, 1, > 1), baseline steroid use (Yes, No), and baseline eligible SES-CD (original, alternative), where the stratum of "original" includes the patients with baseline eligible SES-CD of ≥ 6 (or ≥ 4 for subjects with isolated ileal disease), and the stratum of "alternative" includes the patients with baseline eligible SES-CD of ≥ 3 to < 6 for ileocolonic or colonic disease or SES-CD of 3 for isolated ileal disease.

Subjects with clinical non-response to risankizumab at Week 12 will be re-randomized in a 1:1:1 ratio to risankizumab 1200 mg IV dose group, risankizumab 360 mg SC dose group, or risankizumab 180 mg SC dose group in the Induction Period 2. Subjects with clinical non-response to placebo at Week 12 will blindly receive risankizumab 1200 mg IV dose. No stratification will be made for the Induction Period 2 randomization.

2.4 Sample Size Determination

For US-specific protocol, the co-primary endpoints are the achievement of CDAI clinical remission at Week 12 and the achievement of endoscopic response at Week 12. For global protocol outside US, the co-primary endpoints are the achievement of clinical remission at Week 12 and the achievement of endoscopic response at Week 12.

Sample size calculation is based on the larger sample size needed to detect treatment difference between the risankizumab treatment group and the placebo group for each of the co-primary endpoints. Since historical data show slightly lower event rate and similar treatment difference for the endoscopic response rate compared to the clinical remission rate at Week 12, as well as the larger treatment difference for the CDAI clinical remission rate compared to the clinical remission rate at Week 12, clinical remission rates at Week

12 are used for power calculation. A total of 855 subjects will be randomized into the two risankizumab treatment groups and the placebo group in a 2:2:1 ratio (342 subjects for the risankizumab 600 mg dose group, 342 subjects for the risankizumab 1200 mg dose group, and 171 subjects for the placebo group). Assuming the clinical remission rate at Week 12 will be 27.8% for one of the risankizumab treatment groups and 12% for the placebo group, a sample size of 342 subjects for the risankizumab group and 171 for the placebo group will have 97% power to detect the treatment difference between risankizumab and placebo using a Fisher's exact test at a 0.025 significant level (two-sided). Assuming the Week 12 CDAI clinical remission rate will be 37% for one of the risankizumab treatment groups and 17% for the placebo group, this sample size will have 99% power to detect the treatment difference between risankizumab and placebo using a Fisher's exact test at a 0.025 significant level (two-sided). Assuming the Week 12 endoscopic response rate will be 25.5% for one of the risankizumab treatment groups and 8% for the placebo group, this sample size will have 99% power to detect the treatment difference between risankizumab and placebo using a Fisher's exact test at a 0.025 significant level (two-sided).

In addition, with sample size of approximately 540 bio-IR subjects, the study will have approximately 76% power to detect the treatment difference between one of the risankizumab treatment groups and the placebo group in clinical remission rates at Week 12 using a Fisher's exact test at a 0.025 significant level (two-sided) for the bio-IR population, assuming the Week 12 clinical remission rate will be 23.5% for one of the risankizumab treatment groups and 10% for the placebo group for bio-IR subjects. The study will have approximately 92% power for the bio-IR population to detect the treatment difference between one of the risankizumab dose groups and placebo in CDAI clinical remission rates at Week 12 using Fisher's exact test at a 0.025 significant level (two-sided) for the bio-IR population, assuming the Week 12 CDAI clinical remission rate will be 34% for one of the risankizumab treatment groups and 15% for the placebo group. The rates are based on the results from the Phase 2 Study 1311.6, where the study population comprised approximately 95% bio-IR subjects, and account for lower rates adjustment for ustekinumab failure patients.

Similarly, with sample size of approximately 315 non-bio-IR subjects, the study will have 72% power to detect the treatment difference between one of the risankizumab treatment groups and placebo in clinical remission rates at Week 12 using a Fisher's exact test at a 0.025 significant level (two-sided) for the non-bio-IR population, assuming the Week 12 clinical remission rate will be 35% for one of the risankizumab treatment groups and 15% for the placebo group for non-bio-IR subjects. The study will have 70% power to detect the treatment difference between one of the risankizumab treatment groups and the placebo group in CDAI clinical remission rates at Week 12 using Fisher's exact test at a 0.025 significant level (two-sided) for the non-bio-IR population, assuming the Week 12 CDAI clinical remission rate will be 42% for one of the risankizumab treatment groups and 21% for the placebo group for non-bio-IR subjects. The non-bio-IR population is expected to have higher remission rates and a wider treatment difference than the bio-IR population. Similar trends for greater remission rates and treatment differences by prior biologic use were observed in prior pivotal studies with ustekinumab, vedolizumab and adalimumab.

3.0 Endpoints

Below are the definitions of the endpoints used in the efficacy analysis:

- **Clinical remission:** average daily SF ≤ 2.8 and not worse than baseline AND average daily AP score ≤ 1 and not worse than baseline
- **Enhanced clinical response:** ≥ 60% decrease in average daily SF and/or ≥ 35% decrease in average daily AP score and both not worse than baseline, and/or clinical remission
- **Clinical response:** ≥ 30% decrease in average daily SF and/or ≥ 30% decrease in average daily AP score and both not worse than baseline
- **Endoscopic response:** decrease in SES-CD > 50% from baseline (or for subjects with isolated ileal disease and a baseline SES-CD of 4, at least a 2 point reduction from baseline), as scored by central reviewer

- **Ulcer-free endoscopy:** SES-CD ulcerated surface subscore of 0 in subjects with SES-CD ulcerated surface subscore ≥ 1 at baseline, as scored by a central reviewer
- **Endoscopic remission:** SES-CD ≤ 4 and at least a 2 point reduction versus baseline and no subscore greater than 1 in any individual variable, as scored by a central reviewer
- **CDAI clinical response:** reduction of CDAI ≥ 100 points from baseline
- **CDAI clinical remission:** CDAI < 150
- **SF remission:** average daily SF ≤ 2.8 and not worse than baseline
- **AP remission:** average daily AP score ≤ 1 and not worse than baseline

3.1 Primary Endpoint(s)

For US-specific protocol, the co-primary efficacy endpoints are

- The achievement of CDAI clinical remission at Week 12
- The achievement of endoscopic response at Week 12

For global protocol outside US, the co-primary efficacy endpoints are

- The achievement of clinical remission at Week 12
- The achievement of endoscopic response at Week 12

3.2 Secondary Endpoint(s)

For US-specific protocol, the secondary endpoints <u>under overall type I error control</u> are:

1. The achievement of clinical remission at Week 12
2. The achievement of CDAI clinical response at Week 4
3. The achievement of CDAI clinical response at Week 12

4. Change from baseline in Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-Fatigue) at Week 12
5. The achievement of CDAI clinical remission at Week 4
6. The achievement of CDAI clinical response and endoscopic response at Week 12
7. The achievement of SF remission at Week 12
8. The achievement of AP remission at Week 12
9. The achievement of endoscopic remission at Week 12
10. The achievement of enhanced clinical response at Week 4
11. The achievement of ulcer-free endoscopy at Week 12
12. The achievement of enhanced clinical response at Week 12
13. The achievement of resolution of extra-intestinal manifestations (EIMs) at Week 12, in subjects with any EIMs at baseline
14. Occurrence of CD-related hospitalization through Week 12
15. The achievement of no draining fistulas at Week 12 in subjects with draining fistulas at baseline

For global protocol outside US, the secondary endpoints <u>under overall type I error control are:</u>

1. The achievement of CDAI clinical remission at Week 12
2. The achievement of CDAI clinical response at Week 4
3. The achievement of clinical remission at Week 4
4. The achievement of CDAI clinical response at Week 12
5. Change from baseline in FACIT-Fatigue at Week 12

6. Change from baseline in Inflammatory Bowel Disease Questionnaire (IBDQ) total score at Week 12
7. The achievement of enhanced clinical response and endoscopic response at Week 12
8. The achievement of endoscopic remission at Week 12
9. The achievement of enhanced clinical response at Week 4
10. The achievement of ulcer-free endoscopy at Week 12
11. The achievement of enhanced clinical response at Week 12
12. The achievement of resolution of EIMs at Week 12, in subjects with any EIMs at baseline
13. Occurrence of CD-related hospitalization through Week 12
14. The achievement of no draining fistulas at Week 12 in subjects with draining fistulas at baseline
15. Change from baseline in Work Productivity and Impairment Questionnaire – Crohn's disease (WPAI-CD) Overall Work Impairment at Week 12
16. Change from baseline in Short Form-36 (SF-36) Physical Component Summary (PCS) score at Week 12

3.3 Other Efficacy Endpoint(s)

The primary and secondary efficacy endpoints are listed in Section 3.1 and Section 3.2, respectively. The additional efficacy endpoints for each induction period are:

12-Week Induction Period:

- Change from baseline in IBDQ at Week 4, Week 12 respectively
- Change from baseline in individual IBDQ domain scores (bowel, emotional, social, systemic) at Week 4, Week 12 respectively

- The achievement of IBDQ remission (IBDQ ≥ 170 points) at Week 4, Week 12 respectively
- The achievement of IBDQ response (increase in IBDQ ≥ 16 points from baseline) at Week 4, Week 12 respectively
- Change from baseline in WPAI-CD at Week 4, Week 12 respectively
- Change from baseline in European Quality of Life 5 Dimensions (EQ-5D-5L) at Week 4, Week 12 respectively
- Change from baseline in FACIT-Fatigue total score and individual items at Week 4, Week 12 respectively
- The achievement of increase in FACIT-Fatigue total score from baseline ≥7 at Week 12
- The achievement of increase in FACIT-Fatigue total score from baseline ≥9 at Week 12
- The achievement of increase in FACIT-Fatigue total score from baseline ≥11 at Week 12
- Change from baseline in Short Form-36 (PCS and Mental Component Summary score [MCS] and all 8 domain scores) at Week 4, Week 12 respectively
- Change from baseline in Fecal calprotectin (FCP) at Week 4, Week 12 respectively
- Change from baseline in high-sensitivity C-reactive protein (hs-CRP) at Week 4, Week 8, Week 12 respectively
- Change from baseline in average daily AP score at Week 4, Week 8, Week 12 respectively
- Change from baseline in average daily SF at Week 4, Week 8, Week 12 respectively
- The achievement of clinical remission at Week 4, Week 8, Week 12 respectively
- The achievement of enhanced clinical response at Week 4, Week 8, Week 12 respectively
- The achievement of clinical response at Week 4, Week 8, Week 12 respectively

- Change from baseline in SES-CD at Week 12
- The achievement of a SES-CD ulcerated surface subscore ≤ 1 at Week 12 in subjects with a SES-CD ulcerated surface subscore ≥ 2 at baseline
- Patient Global Impression of Severity (PGIS) response at Week 4, Week 8, Week 12 respectively
- The achievement of subjects who report symptoms to be "Minimal" or "Absent" on the PGIS at Week 4, Week 8, Week 12 respectively, for subjects who did not report symptoms to be "Minimal" or "Absent" at baseline
- Patient Global Impression of Change (PGIC) response at Week 4, Week 8, Week 12 respectively
- The achievement of subjects who are "Very much improved" or "Much improved" on the Patient Global Impression of Change (PGIC) at Week 4, Week 8, Week 12 respectively
- Change from baseline in PGIS at Week 4, Week 8, Week 12 respectively
- Occurrence of CD-related surgeries through Week 12
- The achievement of any reduction in SES-CD at Week 12
- The achievement of CDAI clinical remission at Week 4, Week 8, Week 12 respectively
- The achievement of CDAI clinical response at Week 4, Week 8, Week 12 respectively
- The achievement of SF remission at Week 4, Week 8, Week 12 respectively
- The achievement of AP remission at Week 4, Week 8, Week 12 respectively
- Change from baseline in CDAI at Week 4, Week 8, Week 12 respectively
- Change from baseline in Crohn's Symptom Severity (CSS) at Week 4, Week 8, Week 12 respectively
- The achievement of resolution of extra-intestinal manifestations (EIMs) at Week 4, Week 8, Week 12 respectively, in subjects with any EIMs at baseline
- The achievement of no draining fistulas at Week 4, Week 8, Week 12 respectively, in subjects with draining fistulas at baseline
- The achievement of no non-draining fistulas at Week 4, Week 8, Week 12 respectively, in subjects with non-draining fistulas at baseline

- The achievement of no anal fissure at Week 4, Week 8, Week 12 respectively, in subjects with anal fissure at baseline

Induction Period 2:

- The achievement of clinical remission at Week 16, 20, 24 respectively
- The achievement of endoscopic response at Week 24
- The achievement of CDAI clinical remission at Week 16, 20, 24 respectively
- The achievement of CDAI clinical response at Week 16, 20, 24 respectively
- The achievement of SF remission at Week 16, 20, 24 respectively
- The achievement of AP remission at Week 16, 20, 24 respectively
- Change from baseline in FACIT-Fatigue at Week 24
- Change from baseline in IBDQ total score at Week 24
- The achievement of enhanced clinical response and endoscopic response at Week 24
- The achievement of CDAI clinical response and endoscopic response at Week 24
- The achievement of endoscopic remission at Week 24
- The achievement of enhanced clinical response at Week 16, 20, 24 respectively
- The achievement of ulcer-free endoscopy at Week 24
- The achievement of resolution of extra-intestinal manifestations (EIMs) at Week 16, 20, 24 respectively, in subjects with any EIMs at baseline
- Occurrence of CD-related hospitalization from Week 12 through Week 24
- The achievement of no draining fistulas at Week 16, 20, 24 respectively, in subjects with draining fistulas at baseline
- Change from baseline in WPAI-CD at Week 24
- Change from baseline in SF-36 (PCS, MCS and all 8 domain scores) at Week 24

3.4 Safety Endpoint(s)

- Treatment-emergent adverse event
- Laboratory data
- Vital signs

3.5 Additional Endpoint(s)

Pharmacokinetics/pharmacodynamics, pharmacogenetics, and selected biomarkers will be analyzed separately and are not included in this SAP.

4.0 Analysis Populations

The following population sets are defined and will be used for the statistical analyses.

Significant non-compliance was identified at an investigational site (Investigator ID [redacted]). As a result of this finding, efficacy data for the subjects enrolled at this investigational site will be excluded from the statistical analyses. Safety data for those subjects will be included in the statistical analyses. There were 5 subjects enrolled at the site with Investigator ID [redacted] in this study.

Intent-to-Treat Population

The Intent-to-Treat (ITT) population includes subjects who received at least one dose of study drug.

12-Week Induction Period:

The Intent-to-Treat population for the 12-Week Induction Period (denoted as ITT1) includes randomized subjects who received at least one dose of study drug during the 12-Week Induction Period.

- **ITT1A:** This Intent-to-Treat population for the 12-Week Induction Period includes the subjects in ITT1 who had baseline eligible SES-CD of ≥ 6 (≥ 4 for

isolated ileal disease). This population will be the primary population for efficacy analysis of the 12-Week Induction Period.

- **ITT1B:** This Intent-to-Treat population for the 12-Week Induction Period includes the subjects who are in ITT1 set but did not have baseline eligible SES-CD of ≥ 6 (≥ 4 for isolated ileal disease). This population will be used for exploratory efficacy analysis purpose.

Induction Period 2:

The Intent-to-Treat population for the Induction Period 2 (denoted as ITT2) includes subjects who are randomized and received at least one dose of risankizumab during the Induction Period 2, and the subjects who received placebo induction treatment during 12-Week Induction Period and entered the Induction Period 2 (denoted as placebo/risankizumab). The placebo/risankizumab subjects will be analyzed as a separate treatment group in the Induction Period 2.

- **ITT2A:** This Intent-to-Treat population for Induction Period 2 includes subjects in ITT2 who had baseline eligible SES-CD of ≥ 6 (≥ 4 for isolated ileal disease). This population will be used for exploratory efficacy analysis purpose.
- **ITT2B:** This Intent-to-Treat population for Induction Period 2 includes the subjects who are in ITT2 set but did not have baseline eligible SES-CD of ≥ 6 (≥ 4 for isolated ileal disease). This population will be used for exploratory efficacy analysis purpose.

For the corresponding ITT populations, subjects were assigned to a treatment group based on the randomization or re-randomization schedule, regardless of the treatment actually received. The demographics/baseline and efficacy analysis will be based on the corresponding ITT population.

Safety Population

For the safety populations, subjects were assigned to a treatment group based on the treatment actually received. The safety population (SA) consists of all subjects who received at least one dose of study medication. The safety analysis in each induction period will be based on the following safety population.

12-Week Induction Period:

- **SA1:** This safety population for the 12-Week Induction Period consists of all subjects who received at least one dose of study medication from the 12-Week Induction Period.
- **SA1A:** This safety population for the 12-Week Induction Period includes the subjects in SA1 set who had baseline eligible SES-CD of ≥ 6 (≥ 4 for isolated ileal disease).

Induction Period 2:

- **SA2:** This safety population for the Induction Period 2 consists of subjects who received at least one dose of risankizumab during the Induction Period 2.
- **SA2A:** This safety population for the Induction Period 2 includes the subjects in SA2 set who had baseline eligible SES-CD of ≥ 6 (≥ 4 for isolated ileal disease).

All Risankizumab:

- **SA-ALL:** This safety population includes all subjects who received at least one dose of risankizumab any time during the study.

5.0 Subject Disposition

The total number of subjects who were enrolled, randomized and treated will be summarized.

A summary of subject accountability will be provided where the number of subjects in each of the following categories will be summarized for each treatment group:

- Subjects enrolled/randomized in the study
- Subjects who took at least one dose of study drug (ITT1A for the 12-Week Induction Period, ITT2 for the Induction Period 2 and ITT1 for the entire study)
- Subjects who completed study drug (SA1A for the 12-Week Induction Period, SA2 for the Induction Period 2 and SA1 for the entire study)
- Subjects who prematurely discontinued study drug (SA1A for the 12-Week Induction Period, SA2 for the Induction Period 2 and SA1 for the entire study)
- Subjects who prematurely discontinued from study (ITT1A for the 12-Week Induction Period, ITT2 for the Induction Period 2 and ITT1 for the entire study)
- Subjects who completed study (ITT1A for the 12-Week Induction Period, ITT2 for the Induction Period 2 and ITT1 for the entire study)

For end of study participation, the number and percentage of subjects who completed the protocol defined follow-up period (or did not with associated reasons) will be summarized overall and by treatment group.

6.0 Study Drug Duration and Compliance

6.1 Study Drug Duration

12-Week Induction Period:

For the safety population SA1A, the duration of exposure to study drug will be summarized for each treatment group in the 12-Week Induction Period. The study drug dose date refers to recorded dates of infusion/injections of study drug.

- Duration of exposure in the 12-Week Induction Period:
  - If the subject enrolled into M16-000 after the 12-Week Induction Period:
    - Duration of exposure = min(the first study drug dose date in M16-000 – the first study drug dose date in the 12-Week Induction Period, the last study drug dose date in the 12-Week Induction Period –

the first study drug dose date in the 12-Week Induction Period +28 days),

- If the subject enrolled into the Induction Period 2 after the 12-Week Induction Period:
  - Duration of exposure = min(the first study drug dose date in the Induction Period 2 – the first study drug dose date in the 12-Week Induction Period, the last study drug dose date in the 12-Week Induction Period – the first study drug dose date in the 12-Week Induction Period +28 days),
- Otherwise:
  - Duration of exposure = the last study drug dose date in the 12-Week Induction Period – the first study drug dose date in the 12-Week Induction Period +28 days.

For each treatment group and total in this induction period, the duration of exposure will be summarized by the number of subjects treated, as well as the mean, standard deviation, median, minimum and maximum values. In addition, the number and percentage of subjects exposed to study drug will be summarized for the following exclusive duration intervals:

- 1 – 28 days
- 29 – 56 days
- 57 – 84 days
- 85 – 112 days
- >112 days

Induction Period 2:

For the safety population SA2A, the duration of exposure to study drug will be summarized for each treatment group in the Induction Period 2. The study drug dose date refers to recorded dates of infusion/injections of study drug.

- Duration of exposure in the Induction Period 2:
  - If the subject enrolled into M16-000 after the Induction Period 2:
    - Duration of exposure = min(the first study drug dose date in M16-000 – the first study drug dose date in the Induction Period 2, the last study drug dose date in the Induction Period 2 – the first study drug dose date in the Induction Period 2 +28 days [for IV dose] or 56 days [for SC dose]),
  - Otherwise:
    - Duration of exposure = the last study drug dose date in the Induction Period 2 – the first study drug dose date in the Induction Period 2 +28 days (for IV dose) or 56 days (for SC dose).

For each treatment group and total in the Induction Period 2, the duration of exposure will be summarized by the number of subjects treated, as well as the mean, standard deviation, median, minimum and maximum values. In addition, the number and percentage of subjects exposed to study drug will be summarized for the following exclusive duration intervals:

- 1 – 28 days
- 29 – 56 days
- 57 – 84 days
- 85 – 112 days
- >112 days

ALL Risankizumab:

For all risankizumab population SA-ALL, the duration of exposure to study drug will be summarized for all risankizumab treatment group over the whole course of the study. The study drug dose date refers to recorded dates of infusion/injections of study drug.

- Duration of exposure for all risankizumab:
  - If the subject enrolled into M16-000:

- Duration of exposure = min(the first study drug dose date in M16-000 – the first risankizumab dose date in M16-006, the last risankizumab dose date in M16-006 – the first risankizumab dose date in M16-006 +28 days [if the last dose is IV dose] or 56 days [if the last dose is SC dose]),

○ Otherwise:

- Duration of exposure = the last risankizumab dose date in M16-006 – the first risankizumab dose date in M16-006 +28 days [if the last dose is IV dose] or 56 days [if the last dose is SC dose].

The duration of exposure will be summarized by the number of subjects treated, as well as the mean, standard deviation, median, minimum and maximum values. In addition, the number and percentage of subjects exposed to risankizumab will be summarized for the following exclusive duration intervals:

- 1 – 28 days
- 29 – 56 days
- 57 – 84 days
- 85 – 112 days
- 113 – 140 days
- 141 – 196 days
- >196 days

6.2 Study Drug Compliance

The treatment compliance (%) of study drug information will be summarized for safety populations SA1 and SA2:

6.2.1 Intravenous Therapy

- Percent of subjects who received full volume
- Average infusion duration (hours)
- Treatment compliance (%)

6.2.2 Subcutaneous Therapy

- Total number of injections received
- Treatment compliance (%)

7.0 Demographics, Baseline Characteristics, Medical History, and Prior/Concomitant Medications

Demographics and baseline or disease characteristics will be summarized for the ITT1A, ITT1B, ITT2A, ITT2B, SA1A population overall and by treatment group. Medical history and prior medications will be summarized for the ITT1A population overall and by treatment group. Concomitant medications will be summarized for the ITT1A and ITT2 population overall and by treatment group. Categorical variables will be summarized with the number and percentage of subjects; percentages will be calculated based on the number of non-missing observations. Continuous variables will be summarized with descriptive statistics (number of non-missing observations, mean and standard deviation, median, minimum and maximum).

7.1 Demographics and Baseline Characteristics

The following demographic and baseline or disease characteristics will be summarized at the baseline of the study, where the baseline refers to the last non-missing observation before the first administration of any study drug.

Demographics

- Age (years)
- Age category [< 18, [18, 40), [40, 65), ≥ 65]
- Sex [male/female]
- Race [American Indian/Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, White, Multiple]
- Ethnicity [Hispanic/Latino, Non-Hispanic/Latino]
- Geographic region (as defined in Appendix I)

Tobacco/Nicotine and Alcohol Use

- Tobacco/Nicotine Use [currently smokes, ex-smoker, never smoked, unknown]
- Alcohol Use [current drinker, former drinker, non-drinker, unknown]

Vital Signs

- Systolic and diastolic blood pressure (mmHg)
- Pulse (bpm)
- Body temperature (°C)
- Body weight (kg) – overall and by sex
- Body weight category (< 60 kg, ≥ 60 kg)
- Height (cm)
- Body Mass Index (kg/m^2)
- BMI Category
  - Underweight [< 18.5 kg/m^2]
  - Normal [≥ 18.5 and < 25 kg/m^2]
  - Overweight [≥ 25 and < 30 kg/m^2]
  - Obese [≥ 30 kg/m^2]

Patient Reported Outcomes at Baseline

- IBDQ – total score and domain scores
- WPAI-CD – 4 sub-scores
- CSS – total score
- FACIT-Fatigue – total score and 13 item scores
- EQ-5D-5L – Index score and Vas Pain
- SF-36 – 8 domain scores, PCS and MCS
- Average Daily Stool Frequency
- Average Daily Abdominal Pain

Other Assessments at Baseline

- CDAI
- SES-CD
- Crohn's Disease Duration (years)
- Fecal calprotectin
- Baseline corticosteroid use (yes, no)
- Baseline immunomodulator use (yes, no)
- Biologics failure history (0, 1, >1)
- TNF failure history (0, 1, >1) within bio-IR population
- Vedolizumab failure history (yes, no) within bio-IR population
- Ustekinumab failure history (yes, no) within bio-IR population
- hs-CRP
- Draining fistulas (yes, no)
- Non-draining fistulas (yes, no)
- Anal fissures (yes, no)
- Crohn's Disease Location Per SES-CD (colonic only, ileal only, ileal-colonic)
- Extra Intestinal Manifestation (yes, no)

In addition, some clinical tests at baseline such as TB test, pregnancy test and electrocardiogram assessment will be summarized. Crohn's Disease history (Montreal classification) and Crohn's Disease surgical history will be summarized separately as well.

7.2 Medical History

Medical history data will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). The actual version of the MedDRA coding dictionary will be noted in the statistical tables and clinical study report. The number and percentage of subjects in each medical history category (by MedDRA system organ class and preferred term) will be summarized overall and by treatment group. The system organ class (SOC) will be presented in alphabetical order, and the preferred terms will be presented in alphabetical

order within each SOC. Subjects reporting more than one condition/diagnosis will be counted only once in each row (SOC or preferred term).

7.3 Prior and Concomitant Medications

Prior and concomitant medications will be summarized by generic name. A prior medication is defined as any medication taken prior to the date of the first dose of any study drug. A concomitant medication for 12-Week Induction Period is defined as any medication other than study drug that (1) was started prior to the first dose of study drug and continued to be taken after the first dose of study drug in the 12-Week Induction Period or (2) was started after the first dose of study drug during the 12-Week Induction Period, but prior to the first dose of Induction Period 2 if the subject enters Induction Period 2 or prior to the first dose of M16-000 if the subject enters M16-000, or prior to the last dose + 140 days if the subject does not enter Induction Period 2 nor M16-000.

A concomitant medication for Induction Period 2 is defined as any medication other than study drug that (1) was started prior to the first dose of study drug in Induction Period 2 and continued to be taken after the first dose of study drug in Induction Period 2 or (2) was started after the first dose of study drug in Induction Period 2, but prior to first dose of M16-000 if the subject enters M16-000, or last dose + 140 days if the subject does not enter M16-000. Information about concomitant medications can be collected up to 140 days after last dose, as per protocol.

The number and percentage of subjects taking medications will be summarized by generic drug name based on the World Health Organization (WHO) Drug Dictionary for both prior and concomitant medications.

8.0 Efficacy Analyses

8.1 General Considerations

The primary and secondary efficacy endpoints as specified in Section 3.1 and Section 3.2 will be analyzed based on ITT1A for 12-Week Induction Period. The difference between

treatment groups for the primary and ranked secondary efficacy endpoints will be tested with graphical multiplicity adjustment[1] to ensure a strong control of family-wise type I error rate at significance level $\alpha = 0.05$ (two-sided) (see Section 11.0). Analyses of other efficacy endpoints for 12-Week Induction Period as specified in Section 3.3 will be performed using ITT1A at the nominal significance level $\alpha = 0.05$ (two-sided) without multiplicity adjustment.

Analyses for other efficacy endpoints as specified in Section 3.3 for Induction Period 2 will be performed using ITT2A. Since the Induction Period 2 is for exploratory purpose, no multiplicity adjustment will be performed.

For all efficacy endpoints, the descriptive statistics will be provided by period and by treatment group in which the subjects were randomized to at the beginning of each period. The statistics include number of observations, mean, standard deviation, minimum, median, and maximum for continuous variables, and number and percent for discrete variables.

Unless otherwise specified, any subject who is randomized based on a wrong stratum will be analyzed according to the actual stratum the subject belongs to.

When all subjects complete their Week 12/PD visit, the database will be locked and un-blinded for the 12-Week Induction Period. The planned analysis for the 12-Week Induction Period will be performed. This is the only and final efficacy analysis for the 12-Week Induction Period. When all the subjects who enter the Induction Period 2 finish Week 24/PD visit, the database will be locked for the whole study and all the planned analysis for the Induction Period 2 will be performed.

Analysis of Binary Variables:

For binary variables, frequencies and percentages will be reported for each treatment group. Pairwise comparison of each risankizumab group and placebo will be performed using the Cochran Mantel-Haenszel (CMH) test adjusting for stratification factors. If there is a stratum for a treatment group that has no subject in it, a value of 0.1 will be

added to all cells in the corresponding table in order to prevent dividing by 0, as suggested in Greenland and Robins (1985)[3]. Point estimates and 95% confidence interval (CI)s for the difference in proportions between each of risankizumab treatment groups and placebo will be provided. Construction of CI for the common risk difference will be based on the Mantel-Haenszel estimate adjusting for stratification factors. For binary endpoints with small number of sample size (such as achievement of no draining fistulas at Week 12 in subjects with draining fistulas at baseline), or small number of responders (such as occurrence of CD-related hospitalization through Week 12), Chi-square test (or Fisher's exact test if ≥ 20% of the cells have expected cell count < 5) will be performed.

Analysis of Continuous Variables:

For continuous variables, the model based mean and standard error will be provided. The baseline and visit means will also be presented for each treatment group. Continuous variables collected longitudinally will be using a Mixed-Effect Model Repeated Measurement (MMRM) model. Continuous efficacy variables which are collected at only one post-baseline visit (such as SES-CD) will be analyzed using an Analysis of Covariance (ANCOVA) model. Point estimates and 95% CIs of mean change from baseline within treatment groups, and between each risankizumab treatment group and placebo will be provided. The baseline refers to the last non-missing observation before the first administration of any study drug or randomization if no study drug is given.

8.2 Handling of Potential Intercurrent Events

Potential intercurrent events considered in this study include 1) premature discontinuation of study drug and 2) initiation or dose escalation of CD-related corticosteroids defined in Section 8.2.2. Intercurrent events will be handled using the following methods for the efficacy analysis:

8.2.1 Premature Discontinuation of Study Drug

Data collected will be used regardless of premature discontinuation of study drug.

8.2.2 CD-Related Corticosteroid Therapy

The CD-related corticosteroids intercurrent event is defined as follows.

- Subjects not on CD-related corticosteroids (systemic or locally acting corticosteroids for CD) at baseline who initiated CD-related corticosteroids during the study;
- Subjects on CD-related systemic corticosteroids at baseline who have dosages increased to greater than the prednisone equivalent dose of corticosteroid taken at baseline, or initiation of any rectal corticosteroids during the study regardless of rectal corticosteroid dose;
- Subjects on CD-related rectal corticosteroids at baseline who have dosages increased to greater than the dose taken at baseline, or initiation of any new type of rectal or any systemic corticosteroids during the study.

The time point of the CD-related corticosteroids intercurrent event is defined as the earlier date when one or more of the scenarios above occurs for a subject. As such, all measurements at or after the occurrence of the CD-related corticosteroids intercurrent event through the end of the study will not be used in the analysis and will be handled as specified in Appendix J.

8.3 Handling of Missing Data

Missing data could occur due to various reasons, including missing visits/assessments, early withdrawal from the clinical trials, or missing due to Coronavirus disease 2019 (COVID-19) infection or logistic restriction.

The COVID-19 pandemic is interfering with the conduct of many ongoing trials, with potential impacts on treatment duration and the collection, analysis, and the interpretation of clinical trial data. Some protocol-specified visits in the clinical trials may be impacted due to COVID-19 infection or logistical restrictions during the pandemic. For example, some scheduled visits may be missed due to self-quarantine or local government restrictions on travel; some visits may also be delayed or canceled due to healthcare resource constraints during the pandemic. Impacted visits due to COVID-19 will be

recorded in the database. The probability of having missed visits and missing data due to COVID-19 infection or logistical restrictions related to the COVID-19 pandemic can be reasonably assumed to be unrelated to the unobserved values. Therefore, for the purpose of statistical analysis, it is reasonable to assume that these missing data are missing at random (MAR) and the statistical models that require MAR assumption are appropriate. The intent is to provide reliable estimates of the treatment effects under the scenario without the impact of COVID-19 pandemic.

Missing data for the efficacy analyses will be handled using the methods described below. Handling of missing data in IBDQ and FACIT-Fatigue is shown in Appendix F and Appendix G.

8.3.1 Categorical Endpoints

For binary efficacy endpoints, missing data will be handled using the following approaches:

- The primary approach for handling missing data in the analysis of binary endpoints will use **N**on-**R**esponder **I**mputation (NRI) while incorporating Multiple Imputation (MI) to handle missing data due to **C**OVID-19 (NRI-C).
  - The NRI-C will categorize any subject who does not have an evaluation during a pre-specified visit window (either due to missing assessment or due to early withdrawal from the study) as a non-responder for the visit. The only exception is that missing data due to COVID-19 infection or logistical restriction will be handled by MI. At each visit, subjects will be characterized as responders or non-responders based on MI imputed values if missing due to Covid-19, otherwise subjects will be considered as non-responder for missing due to other reasons in the NRI-C approach. Of note, subjects will be counted as non-responders thereafter and will not be imputed by MI after the CD-related corticosteroids censoring time point as specified in Section 8.2.2 and Appendix J.
- A sensitivity analysis for binary endpoints will use **NRI** with **N**o special data handling for missing due to **C**OVID-19 (NRI-NC).

- NRI-NC will be performed in the same way as NRI-C without the exception above. Missing due to COVID-19 infection or logistical restriction will also be counted as non-responders. This is the same method as "NRI" as defined in the protocol.

- Details on MI method are described below:
  - Markov Chain Monte Carlo (MCMC) will be first applied to augment data into monotonic missing pattern, where applicable, and PROC MI will be used to generate 30 datasets using the regression method. If the binary endpoints are derived from the continuous variables, PROC MI will be applied to the continuous variables. The variables to be included in the imputation model are: treatment group, randomization stratification factors (Number of prior biologics failed (0, 1, > 1), baseline steroid use (Yes, No)), age, gender, weight and if applicable, baseline measurement and post-baseline measurements at each visit up to the end of the analysis period. The random seed for MCMC and the random seed for PROC MI are specified in Appendix H. The imputed post-baseline measurements will be rounded to the same precision as the observed data before the determination of responder status. Prior to applying CMH model, subjects will be characterized as responders or non-responders based on MI imputed values if missing due to Covid-19, otherwise subjects will be considered as non-responder for missing due to other reasons in the NRI-C approach. Using the CMH model adjusted by randomization stratification factors, the endpoints will be analyzed using each of the 30 imputed datasets. SAS PROC MIANALYZE will be used to generate the final inferences of the risk difference between the risankizumab treatment group and placebo group, using Rubin's rule. The CD-related corticosteroid censoring rule, as specified in Section 8.2.2 and Appendix J, will be applied before and after applying MI for efficacy assessments.

8.3.2 Continuous Endpoints

For continuous efficacy endpoints where Mixed-Effect Model Repeat Measurement (MMRM) analysis or Analysis of Covariance (ANCOVA) is performed, missing data will be handled using the following approaches.

- Mixed-Effect Model Repeat Measurement (MMRM): The repeat measurement analysis will be conducted using a mixed model including observed measurements at all post-baseline visits, after the CD-related corticosteroids censoring rule is applied (see Section 8.2.2 and Appendix J). The mixed model includes the categorical fixed effects of treatment, visit and treatment-by-visit interaction, stratification factors at randomization, and the continuous fixed covariates of baseline measurement. An unstructured variance covariance matrix will be used. Parameter estimation is based on the method of restrictive maximum likelihood (REML). MMRM will be the primary approach in the analysis of continuous variables.
- ANCOVA-C: For continuous efficacy variables that are collected at only one post-baseline visit (such as SES-CD), the missing data due to COVID-19 will be imputed using MI first before applying the ANCOVA model.

In addition, As Observed (AO) and Pattern Mixture Model (PMM) will be performed as sensitivity analysis, if appropriate, for categorical or continuous endpoints.

- As Observed (AO): The AO analysis will not impute values for missing evaluations, and thus a subject who does not have an evaluation on a scheduled visit will be excluded from the AO analysis for that visit. AO will include all values collected in the study regardless of censoring for steroids or not.
- Pattern Mixture Model (PMM): Subjects will be classified into pre-specified missing data patterns and the pattern variables will be defined based on these patterns. The random effect PMM[5] approach will be performed, which includes the pattern variables in the model to examine the effect of missing-data patterns on the outcomes of interest. PROC GLIMMIX will be used in PMM approach for handling categorical endpoints, and PROC MIXED will be used for handling continuous endpoints. The model includes categorical effects of treatment, visit, treatment-by-visit interaction, missing-data pattern and treatment-by-pattern interaction, and the main stratification factors at randomization, and the continuous fixed covariates of baseline measurement, if applicable. The overall treatment effect can be estimated by averaging over the estimates from the different missing-data patterns.

8.4 Primary Efficacy Endpoint(s) and Analyses

8.4.1 Primary Efficacy Endpoint(s)

The co-primary endpoints for the primary analysis of efficacy are:

- **For US specific protocol:**
  - The achievement of CDAI clinical remission at Week 12
  - The achievement of endoscopic response at Week 12
- **For global protocol outside US:**
  - The achievement of clinical remission at Week 12
  - The achievement of endoscopic response at Week 12

The primary estimand for US specific protocol is the following:

- Difference in the percentage of subjects achieving CDAI clinical remission at Week 12 regardless of premature discontinuation of study drug and without initiation or dose escalation of CD-related corticosteroids (See Section 8.2.2) in each of the risankizumab 1200mg IV and risankizumab 600mg IV groups in comparison with the placebo group in the ITT1A population
- Difference in the percentage of subjects achieving endoscopic response at Week 12 regardless of premature discontinuation of study drug and without initiation or dose escalation of CD-related corticosteroids (See Section 8.2.2) in each of the risankizumab 1200mg IV and risankizumab 600mg IV groups in comparison with the placebo group in the ITT1A population.

The primary estimand for global protocol outside US is the following:

- Difference in the percentage of subjects achieving clinical remission at Week 12 regardless of premature discontinuation of study drug and without initiation or dose escalation of CD-related corticosteroids (See Section 8.2.2) in each of the risankizumab 1200mg IV and risankizumab 600mg IV groups in comparison with the placebo group in the ITT1A population

- Difference in the percentage of subjects achieving endoscopic response at Week 12 regardless of premature discontinuation of study drug and without initiation or dose escalation of CD-related corticosteroids (See Section 8.2.2) in each of the risankizumab 1200mg IV and risankizumab 600mg IV groups in comparison with the placebo group in the ITT1A population.

Details in estimand definition are outlined in Appendix J.

8.4.2 Handling of Missing Data for the Primary Efficacy Endpoint(s)

The NRI-C approach for handling missing data will be used for primary analysis for the co-primary endpoints for the 12-Week Induction Period.

The NRI-NC and PMM approaches for handling missing data will be conducted as sensitivity analyses for the co-primary endpoints for the 12-Week Induction Period if applicable. For PMM approach, the pre-specified missing-data patterns are given as below. Other missing data patterns could be explored if deemed necessary.

Missing-data Patterns
Non-completers (missing at Week 4 or at Week 8 or at Week 12)
Completers

8.4.3 Primary Efficacy Analysis

The co-primary endpoints will be analyzed between placebo and each of risankizumab 600 mg IV and 1200 mg IV groups using the CMH test, stratified by the randomization factors of number of prior biologics failed (0, 1, > 1) and baseline steroid use (Yes, No) based on ITT1A. A CMH based two-sided 95% CI for the difference between treatment groups will be constructed.

8.4.4 Additional Analyses of the Primary Efficacy Endpoint(s)

For the co-primary efficacy endpoints, an additional analysis using the same CMH test will be performed using As Observed (AO) data handling without any imputation. The analysis will be conducted using the ITT1A population.

8.5 Secondary Efficacy Analyses

8.5.1 Key Secondary Efficacy Analyses

The estimands corresponding to the secondary efficacy endpoints are defined in Section 2.1. Details in estimand definition for the key secondary endpoints are outlined in Appendix J.

In general, continuous secondary efficacy endpoints with repeat measurements will be analyzed using a MMRM model including factors for treatment group, visit, visit by treatment interaction, stratification variables and the continuous fixed covariates of baseline measurement. The MMRM analysis is considered primary for inferential purposes. Continuous secondary efficacy variables which are collected at only one post-baseline visit (such as SES-CD) will be analyzed using ANCOVA-C approach.

Categorical secondary efficacy variables will be analyzed using the CMH test controlling for stratification variables (prior biologics failed [0, 1, > 1] and baseline steroid use [Yes, No]), except for the endpoints related to draining fistulas and CD-related hospitalization. A CMH based two-sided 95% CI for the difference between each risankizumab treatment group and placebo group will be constructed. The endpoints of "The achievement of no draining fistulas at Week 12 in subjects with draining fistulas at baseline" and "Occurrence of CD-related hospitalization through Week 12" will be analyzed using Chi-square test (or Fisher's exact test if ≥ 20% of the cells have expected cell count < 5).

The NRI-C method will be the primary approach for missing data handling for categorical secondary endpoints except for CD-related hospitalization and draining fistulas. The endpoint of "The achievement of no draining fistulas at Week 12 in subjects with draining fistulas at baseline" will be analyzed using NRI-NC as the primary approach due to the

small sample size. The endpoint of "Occurrence of CD-related hospitalization through Week 12" will be analyzed based on the observed data using AO approach only.

In addition, NRI-NC for missing data handling will be applied as sensitivity analyses for the binary secondary endpoints except for CD-related hospitalization and draining fistulas endpoints.

Analysis of PMM will also be performed as sensitivity analyses for ranked secondary endpoints as appropriate. Analysis of PMM will not be performed for the endpoint of "The achievement of no draining fistulas at Week 12 in subjects with draining fistulas at baseline" due to the small sample size.

8.5.2 Supportive Secondary Efficacy Analyses

For the categorical secondary efficacy endpoints, the same CMH analysis will be repeated using As Observed (AO) data handling without any imputation as an additional analysis. The analysis will be conducted using the ITT1A population.

8.6 Additional Efficacy Analyses

Other continuous and categorical efficacy endpoints for the 12-Week Induction Period will be analyzed using ITT1A similarly as the secondary efficacy endpoints as described in Section 8.5. No multiplicity adjustment will be performed. The chi-square test (or Fisher's exact test if ≥ 20% of the cells have expected cell count < 5 for analyses of binary outcomes) will be used for the following endpoints:

- Patient Global Impression of Severity (PGIS) response at Week 4, Week 8, Week 12 respectively
- Patient Global Impression of Change (PGIC) response at Week 4, Week 8, Week 12 respectively
- The achievement of no draining fistulas at Week 4, Week 8, Week 12 respectively, in subjects with draining fistulas at baseline
- The achievement of no non-draining fistulas at Week 4, Week 8, Week 12 respectively, in subjects with non-draining fistulas at baseline

- The achievement of no anal fissure at Week 4, Week 8, Week 12 respectively, in subjects with anal fissure at baseline

PGIC response and PGIS response endpoints will be analyzed based on the observed data using AO approach. The draining fistulas endpoint, non-draining fistulas endpoint and anal fissure endpoint will be analyzed using NRI-NC approach due to the small sample size.

Analyses for other efficacy endpoints as specified in Section 3.3 for Induction Period 2 will be performed using ITT2A for exploratory purpose. Point estimate and 95% CI for each treatment group as well as point estimate and 95% CI for pairwise treatment differences (i.e., risankizumab 1200mg IV – risankizumab 360mg SC, risankizumab 1200mg IV – risankizumab 180mg SC, risankizumab 360mg SC - risankizumab 180mg SC) will be presented based on the observed data for randomized subjects using AO approach. The subjects in the non-randomized risankizumab 1200mg IV group, who took the placebo in the 12-Week Induction Period, will be displayed separately using descriptive statistics. No p-value will be provided.

In addition, the following selected endpoints will be analyzed using ITT1B for exploratory purpose. Point estimate and 95% CI for each treatment group as well as point estimate and 95% CI for treatment differences between each risankizumab group and the placebo group will be presented based on the observed data using AO approach. No p-value will be provided.

- The achievement of CDAI clinical remission at Week 12
- The achievement of clinical remission at Week 12
- The achievement of endoscopic response at Week 12
- The achievement of SES-CD $\leq$ 2 or 2 points reduction from the baseline at Week 12 among the subjects with baseline eligible SES-CD of >=3 to 6 for ileocolonic or colonic disease or 3 for isolated ileal disease
- The achievement of any reduction in SES-CD at Week 12

Similarly, the following selected endpoints will be analyzed using ITT2B for exploratory purpose. Point estimate and 95% CI for each treatment group as well as point estimate and 95% CI for pairwise treatment differences will be presented based on the observed data for randomized subjects using AO approach. The subjects in the non-randomized risankizumab 1200mg IV group, who took the placebo in the 12-Week Induction Period, will be displayed separately using descriptive statistics. No p-value will be provided.

- The achievement of CDAI clinical remission at Week 24
- The achievement of clinical remission at Week 24
- The achievement of endoscopic response at Week 24
- The achievement of SES-CD ≤ 2 or 2 points reduction from the baseline at Week 24 among the subjects with baseline eligible SES-CD of >=3 to <6 for ileocolonic or colonic disease or 3 for isolated ileal disease
- The achievement of any reduction in SES-CD at Week 24

8.7 Efficacy Subgroup Analyses

To evaluate the consistency of the efficacy over demographic and other baseline characteristics, subgroup analysis will be performed for the co-primary endpoints for the subgroups listed in Table 1 below using ITT1A.

For efficacy subgroup analysis, point estimate and 95% CI for each treatment group as well as point estimate and 95% CI for treatment differences between each risankizumab group and the placebo group will be presented. No p-value will be provided. If any of the resulting subgroups except for age, sex and race has fewer than 10% of the planned study size, the subgroup analyses for that category will not be presented.

Table 1. Subgroups for Efficacy Analysis

Subgroup Factor	Categories
Age	< 18, [18, 40), [40, 65), ≥ 65
Sex	Male or Female
Weight	< 60 kg or ≥ 60 kg
Race	White or non-White
Geographic Region	North America, South/Central America, Western Europe, Eastern Europe, Asia, Other
Prior Biologics Failed	Yes or No, if Yes, please analyze (1, > 1)
Prior Ustekinumab Failure (within bio-IR population)	Yes or No
Prior TNF Failure (within subjects with Prior TNF Failure)	1, >1
Prior Vedolizumab Failure (within bio-IR population)	Yes or No
Baseline Steroids Use	Yes or No
Baseline Immunodulator Use	Yes or No
Baseline CDAI	≤ median or > median
Baseline CDAI	≤ 300 or > 300
Baseline SF	≤ median or > median
Baseline AP Score	≤ median or > median
Baseline SES-CD	≤ median or > median
Baseline SES-CD	≤ 15 or > 15
Disease Duration at Baseline	≤ median or > median
Disease Duration at Baseline	≤ 5 years or > 5 years
Baseline hs-CRP	≤ median or > median
Baseline hs-CRP	≤ 5 mg/L or > 5 mg/L
Baseline Calprotectin	≤ median or > median
Baseline Calprotectin	≤ 250 mg/kg or > 250 mg/kg
Crohn's Disease Location at Baseline	Colonic only, Ileal only, Ileal-colonic

9.0 Safety Analyses

9.1 General Considerations

Safety analyses will include reporting of adverse events, laboratory, and vital signs measurements by treatment group, including a total group for all subjects on

risankizumab. All safety analyses will be performed for the 12-Week Induction Period at the time of database lock. The separate safety analysis for subjects who received at least one dose of risankizumab during the Induction Period 2 will be performed.

The following summary statistics will be presented for subjects who have both baseline and post-baseline values for laboratory parameters and vital signs: the mean value at baseline and at each respective protocol specified visit, and the mean, standard deviation and median for changes from baseline. Categorical data will be summarized using frequencies and percentages. The risk difference and 95% CI between each risankizumab group and placebo group will be presented for safety analysis.

For the safety analysis, subjects are assigned to a treatment group based on the treatment actually received, regardless of the treatment randomized.

For the 12-Week Induction Period analysis, the baseline for safety analysis is defined as the last available measurement before study drug administration for all subjects. For the Induction Period 2 analysis, the baseline for safety analysis will be the last available measurement before the first dose of risankizumab administration in Induction Period 2.

Missing safety data will not be imputed.

9.2 Adverse Events

Adverse events (AEs) will be summarized and presented using primary MedDRA System Organ Classes (SOCs) and preferred terms (PTs) according to the version of the MedDRA coding dictionary used for the study at the time of database lock. The actual version of the MedDRA coding dictionary used will be noted in the AE tables and in the clinical study report. Specific adverse events will be counted once for each subject for calculating percentages, unless stated otherwise. In addition, if the same adverse event occurs multiple times within a subject, the highest severity and level of relationship to investigational product will be reported.

9.2.1 Treatment-Emergent Adverse Events

12-Week Induction Period: Treatment emergent AEs (TEAEs) for the 12-Week Induction Period are defined as events that begin either on or after the first dose of the study drug in the 12-Week Induction Period and until the first dose of study drug in the Study M16-000 if the subject is enrolled into the Study M16-000 or until first dose of study drug in Induction Period 2 if the subject is enrolled into Induction Period 2, or within 140 days after the last dose administration of the study drug in 12-Week Induction Period if the subject does not participate in the Study M16-000 or the Induction Period 2.

Induction Period 2: Treatment-emergent AEs (TEAEs) for the Induction Period 2 are defined as events that begin either on or after the first dose of the study medication in the Induction Period 2 and until first dose of study drug in the Study M16-000 if the subject is enrolled into the Study M16-000 or within 140 days after the last dose of the study drug in the Induction Period 2 if the subject does not participate in the Study M16-000.

All Risankizumab: Treatment-emergent AEs (TEAEs) for "All Risankizumab" are defined as events that begin either on or after the first dose of risankizumab and until first dose of study drug in the Study M16-000 if the subject is enrolled into the Study M16-000 or within 140 days after the last dose of risankizumab in the study if the subject does not participate in the Study M16-000.

Events where the onset date is the same as the study drug start date are assumed to be treatment-emergent, unless the study drug start time and the adverse event start time are collected and the adverse event start time is prior to the study drug start time. If an incomplete onset date was collected for an adverse event, the event will be assumed to be treatment-emergent unless there is other evidence that confirms that the event was not treatment-emergent (e.g., the event end date was prior to the study drug start date).

All treatment-emergent AEs will be summarized overall, as well as by primary MedDRA SOC and Preferred Term. The SOCs will be presented in alphabetical order, and the PTs will be presented in alphabetical order within each SOC.

The number and percentage of subjects experiencing treatment-emergent AEs will be summarized. The risk difference and 95% CI between each risankizumab group and placebo group will be presented.

9.2.2 Adverse Event Overview

An overview of AEs will be presented consisting of the number and percentage of subjects experiencing at least one event for each of the following AE categories:

- Any treatment-emergent AE
- Any treatment-emergent AE related to COVID-19
- Any treatment-emergent AE related to study drug according to the investigator
- Any severe treatment-emergent AE
- Any serious treatment-emergent AE
- Any treatment-emergent AE leading to discontinuation of study drug
- Any treatment-emergent AE leading to death
- Treatment-emergent AEs of Safety Interest including
    - Adjudicated MACE events
    - Serious infections
    - Active Tuberculosis
    - Opportunistic infections excluding tuberculosis and herpes zoster
    - Malignant tumors
    - Malignancies excluding NMSC
    - Hypersensitivity – serious events only
    - Adjudicated anaphylactic reaction
    - Serious anaphylactic reaction
- All deaths
    - Deaths occurring ≤ 140 days after last dose of study drug
    - Deaths occurring > 140 days after last dose of study drug
    - Deaths related to COVID-19

The exposure adjusted rate for AE overview will be presented as well.

9.2.3 Treatment-Emergent Adverse Events by SOC and/or PT

Treatment-emergent adverse events will be summarized by SOC and PT; by maximum relationship to study drug as assessed by the investigator (e.g., reasonable possibility or no reasonable possibility) and SOC and PT; by maximum severity and SOC and PT; and by subject number and SOC and PT. Specific adverse events will be counted once for each subject for calculating percentages, unless stated otherwise. In addition, if the same adverse event occurs multiple times within a subject, the highest severity and level of relationship to investigational product will be reported.

In addition, treatment-emergent adverse events and SAEs will be summarized by PT and sorted by decreasing frequency for the total risankizumab group.

9.2.4 Treatment-Emergent Adverse Events per Patient-Years of Exposure

Exposure-adjusted AEs per 100 patient-years will be provided for all TEAEs, where AEs per 100 patient-years of exposure are defined as the number of AEs divided by the total exposure in 100 patient-years, where the total exposure is calculated as the time from the first study drug dose date in the relevant period/study to the last study drug dose date +140 days in the relevant period/study or the first dose date of the next period/study, whichever occurs earlier.

The exposure adjusted incidence rate (censored at the time of first event) may be conducted for deaths and selected Areas of Safety Interest categories including adjudicated cardiovascular events, malignant tumors, non-melanoma skin cancer (NMSC), and malignant tumors excluding NMSC, as appropriate.

9.2.5 SAEs (Including Deaths) and Adverse Events Leading to Study Drug Discontinuation

SAEs (including deaths) and AEs leading to study drug discontinuation will be summarized by SOC and PT and in listing format. Exposure-adjusted SAEs per 100 patient-years by SOC and PT will be provided as well.

9.2.6 Areas of Safety Interest

The Areas of Safety Interest (ASI) categories are listed in Appendix B, which include serious infections, tuberculosis, opportunistic infections, herpes zoster, malignancies, hypersensitivity reaction, anaphylactic reactions, hepatic events, injection site reactions, adjudicated cardiovascular events. These will be summarized by PT and will be based on standardized or company MedDRA queries (SMQs or CMQs), or based on adjudication results. The risk difference and 95% CI between each risankizumab group and placebo group will be presented. Exposure-adjusted ASI per 100 patient-years by PT or adjudication terms will be provided as well.

Additional ASI may be considered for tabulation/summary based on recommendations from Clinical and Safety teams, as deemed appropriate.

9.3 Analysis of Laboratory Data

Data collected from central and local laboratories, including additional laboratory testing due to an SAE, will be used in all analyses, except for baseline where SAE-related laboratory assessments on or before the first dose of study drug will be excluded. The clinical laboratory tests defined in the protocol (e.g., hematology and clinical chemistry) will be summarized.

Each laboratory variable will be summarized for all time points (starting with baseline) with the number of non-missing observations, mean and standard deviation, median, minimum and maximum. Mean change from baseline to each applicable post-baseline visit will be summarized for selected laboratory variables, with the number of

observations, baseline mean, and visit mean. The change from baseline mean, standard error, and 95% CI will be presented for the mean change from baseline within each treatment group and difference between each risankizumab group and placebo group.

Changes in laboratory parameters will be tabulated using shift tables either based on NCI CTC criteria or categorized as low, normal, or high based on the normal ranges of the laboratory used for each sample. A shift table from baseline either to the worse value (based on NCI CTC criteria) during treatment or to minimum and maximum value (based on normal range), will be created. A similar shift table will be provided to summarize shifts from baseline to the final post-baseline value.

Laboratory abnormalities will be evaluated based on Potentially Clinically Significant (PCS) criteria (Appendix C). For each laboratory PCS criterion, the number and percentage of subjects who have a laboratory value meeting the criteria will be summarized by treatment group. Listings will be provided to summarize subject-level laboratory data for subjects meeting PCS criteria.

The liver-specific laboratory tests include the serum glutamic pyruvic transaminase (SGPT/ALT), serum glutamic-oxaloacetic transaminase (SGOT/AST), alkaline phosphatase (ALP), and total bilirubin (TBL). The frequencies and percentages of subjects with post baseline liver specific function test values that meet the following criteria of potential clinical interest will be summarized by treatment group:

- ALT $\geq$ 3 × ULN
- ALT $\geq$ 5 × ULN
- ALT $\geq$ 10 × ULN
- ALT $\geq$ 20 × ULN
- AST $\geq$ 3 × ULN
- AST $\geq$ 5 × ULN
- AST $\geq$ 10 × ULN
- AST $\geq$ 20 × ULN
- TBL $\geq$ 2 × ULN

- Alkaline phosphatase ≥ 1.5 × ULN
- (ALT and/or AST ≥ 3 × ULN) and (TBL ≥ 1.5 × ULN)
- (ALT and/or AST ≥ 3 × ULN) and (TBL ≥ 2 × ULN)

where ULN is the upper limit of normal.

A listing of potentially clinically significant liver function laboratory values will be provided. The listing will include all subjects who met any of the following 4 criteria at any post-baseline visit:

- ALT ≥ 3 × ULN, or
- AST ≥ 3 × ULN, or
- Alkaline phosphatase ≥ 1.5 × ULN, or
- Total bilirubin ≥ 1.5 × ULN.

A listing of possible Hy's Law cases, defined as those who meet all of the following conditions at any post-baseline visit will be provided:

- ALT of > 3 × ULN or AST of > 3 × ULN,
- Total bilirubin ≥ 2 × ULN.

9.4 Analysis of Vital Signs

Vital sign measurements of systolic and diastolic blood pressure, pulse rate and body weight will be summarized.

Each vital sign variable will be summarized for all time points (starting with baseline) with the number of non-missing observations, mean and standard deviation, median, minimum and maximum. Mean change from baseline to each applicable post-baseline visit will be summarized for each vital sign variable, with the number of observations, baseline mean, and visit mean. The change from baseline mean, standard error, and 95%

CI will be presented for the mean change from baseline within each treatment group and difference between each risankizumab group and placebo group.

Key vital sign variables will be evaluated based on potentially clinically significant (PCS) criteria (Appendix C). For each vital sign PCS criterion, the number and percentage of subjects who have a vital sign value meeting the criteria will be summarized. Listings will be provided to summarize subject-level vital sign data for subjects meeting PCS criteria.

10.0 Interim Analyses

There will be no efficacy interim analysis planned for this study.

10.1 Data Monitoring Committee

An external independent data monitoring committee (DMC) composed of persons independent of AbbVie and with relevant expertise in their field will review unblinded safety data from the ongoing study. The primary responsibility of the DMC will be to protect the safety of the subjects participating in this study.

A separate DMC charter describes the roles and responsibilities of the DMC members, frequency of data reviews, relevant data to be assessed, and general operations.

Since there are no efficacy analyses for early stopping, no alpha adjustment is needed.

11.0 Overall Type-I Error Control

Both of the co-primary efficacy endpoints will be tested at two-sided significance level of 0.025 to control the overall type I error rate at alpha = 0.05 (2-sided) for the multiple comparisons between each risankizumab dose versus placebo using a multiple testing procedure.[1] In addition, the overall type I error rate of efficacy evaluation based on the co-primary and ranked secondary endpoints for the two doses will be strongly controlled at 0.05 (2-sided) level. Specifically, the testing will utilize the sequence of hypothesis testing for the co-primary endpoints followed by the ranked secondary endpoints in the

order as specified in Section 3.2, and will begin with testing each of the co-primary endpoints using α of 0.025 (2-sided) for each dose compared to placebo. If both co-primary endpoints achieve statistical significance within a dose level, continued testing will follow a pre-specified weight of α allocation between the single hypothesis within the family, as well as between the families of hypotheses across the doses (denoted as node). The iterative graphical testing procedures are provided in Figure 1 for US specific protocol and Figure 2 for global protocol outside US.

In the graphs, the arrows specify the weight of α allocation between nodes. Once a hypothesis is rejected (i.e., deemed the endpoint is significant) at its assigned significance level, its significance level will be allocated to the subsequent node. If more than one arrow originates from a node, the significance level will be split between multiple subsequent nodes following the pre-specified weight. The numbers on the arrows denote the weights. For example, the weight 1 denotes 100% transfer of significance level to the next node, and the weight 1/2 denotes 50% and 50% splitting of significance level between the 2 nodes connected to it.

The two co-primary endpoints will be tested at the same α level in order to allocate the significance level to the next sequence of ranked secondary endpoints in the graph. The last group of ranked secondary endpoints in the graph will be tested using the Holm procedure.[2]

Figure 2. Graphical Multiple Testing Procedure for US Specific Protocol

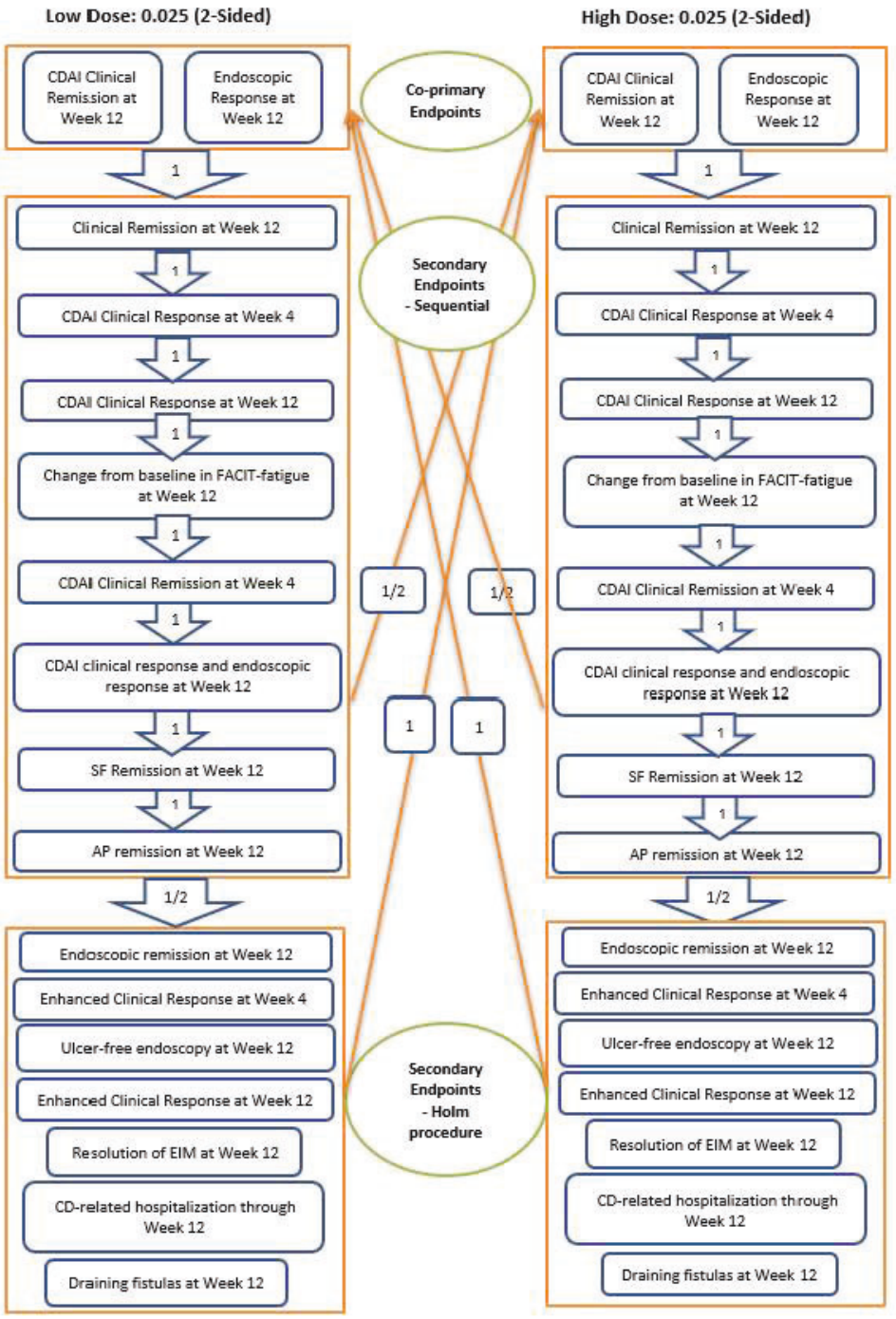
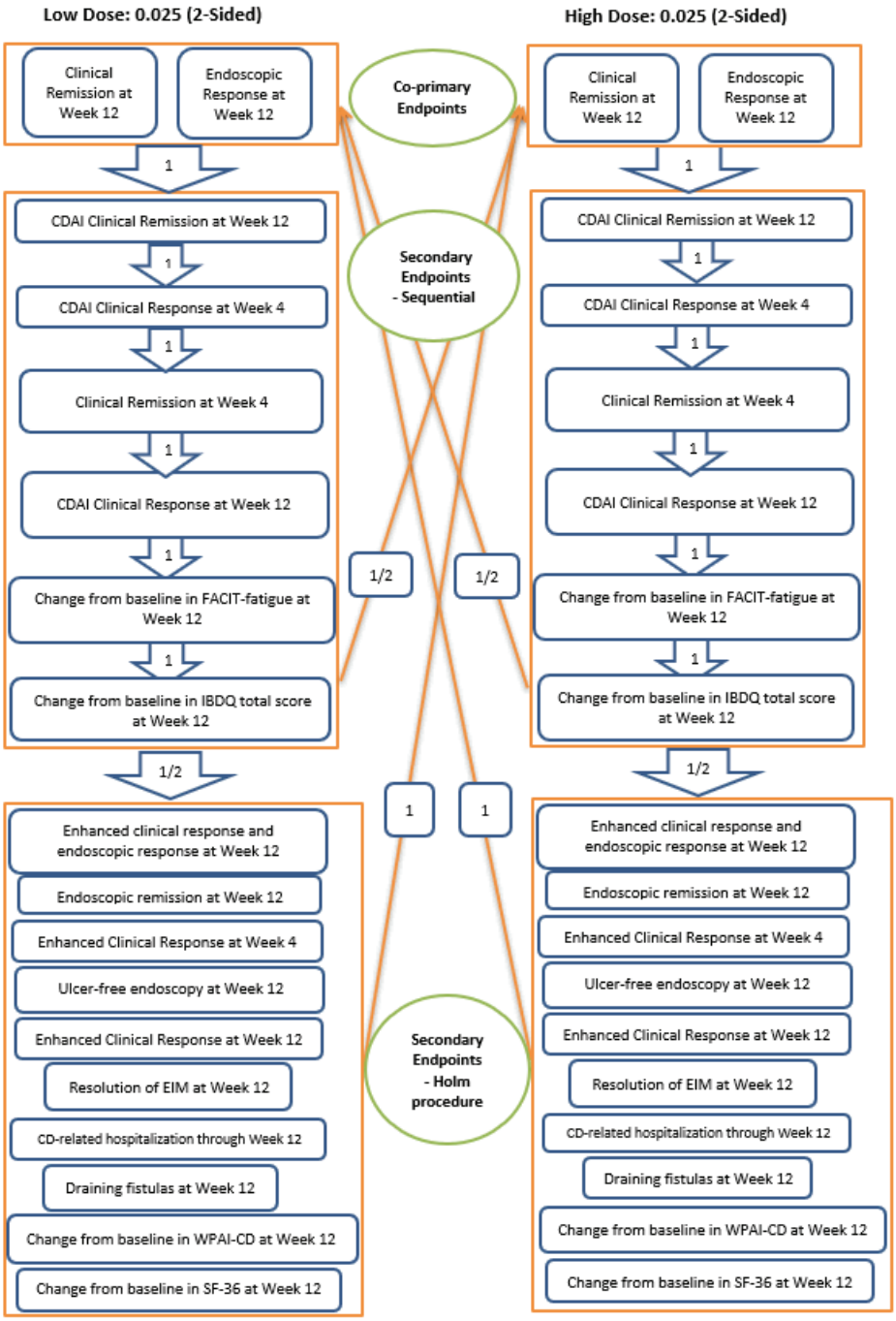


Figure 3. Graphical Multiple Testing Procedure for Global Protocol Outside US



12.0 Version History

Table 2. SAP Version History Summary

SAP Version	Date
1.0	06 OCT 2017
2.0	13 NOV 2018
3.0	25 FEB 2019
4.0	07 FEB 2020

Summary of Changes Between the Previous Version and the Current Version

1. This version of SAP (Version 5) incorporated estimand language according to the guidance provided in ICH E9 R1 Addendum.
2. The co-primary endpoints and ranked secondary endpoints have been updated for US/FDA regulatory purpose in Section 3.0 based on FDA's comments.
3. Two endpoints (WPAI-CD and SF-36) have been moved from exploratory endpoints to ranked secondary endpoints for outside US/FDA regulatory purpose in Section 3.0 based on EMA's comments.
4. The endpoints related to non-draining fistula and anal fissure, and the endpoints related to FACIT-Fatigue have been added as other efficacy endpoints in Section 3.0 for clinical importance.
5. The sample size justification has been updated in Section 2.4 to clarify the sample size calculation for different regulatory purpose.
6. The exclusion of the data from non-compliance sites has been added in Section 4.0.
7. The graphic multiple testing procedures have been updated in Section 11.0 for US/FDA regulatory purpose and outside US/FDA regulatory purpose respectively.
8. Missing imputation methods have been updated in Section 8.3 to include more details for multiple imputation method, as well as introducing NRI-C method for handling missing data due to Covid-19 pandemic.

9. CD-related corticosteroid therapy censoring convention has been added in Section 8.2.2 to clarify the rules for efficacy censoring.
10. The primary, secondary and other efficacy analyses have been updated in Section 8.0 to use NRI-C method as the primary approach for handling missing data.
11. The language of "The risk difference and 95% CI between risankizumab group and placebo group will be presented for safety analysis" has been added in Section 9.0 based on agency's feedback.
12. The language of "The exposure adjusted incidence rate (censored at the time of first event) may be conducted for selected ASI categories as appropriate." has been added in Section 9.2.4 based on agency's feedback.
13. Details about FACIT-Fatigue has been added in Appendix G.
14. Areas of Safety Interest (ASI) have been updated in Appendix B to revise the ASI categories following new standard.
15. Random seeds have been added in Appendix H.

Summary of Changes Between the Current Version of SAP and Protocol

1. Missing imputation methods have been updated to include more details for multiple imputation method, as well as introducing NRI-C method for handling missing data due to Covid-19 pandemic.
2. Two endpoints (WPAI-CD and SF-36) have been moved from exploratory endpoints to ranked secondary endpoints for outside US/FDA regulatory purpose based on EMA's comments.

13.0 References

1. Bretz F, Maurer W, Brannath W, et al. A graphical approach to sequentially rejective multiple test procedures. Stat Med. 2009;28(4):586-604.

2. Holm S. A simple sequentially rejective multiple test procedure. Scand J Stat. 1979;6:65-70.
3. Greenland S, Robins JM. Estimation of a common effect parameter from sparse follow-up data. Biometrics. 1985;41(1):55-68.
4. Berglund PA. An introduction to multiple imputation of complex sample data using SAS® v9.2. 2010. Report No: 21-2010.
5. Hedeker D, Gibbons RD. Application of random-effects pattern-mixture models for missing data in longitudinal studies. Psychological methods. 1997;2(1):64-78.

Appendix A. Protocol Deviations

The number and percentage of subjects who reported at least one of the following protocol deviation categories will be provided.

- Subject entered into the study even though s/he did not satisfy entry criteria.
- Subject developed withdrawal criteria during the study and was not withdrawn.
- Subject received wrong treatment or incorrect dose of study.
- Subject took prohibited concomitant medication.

Appendix B. Areas of Safety Interest (ASI)

ASI Grouping	Categories (ASI)	Search Criteria	Terms to Display	Include in AE Overview (Y/N)
Adjudicated CV Events	MACE	Adjudicated terms will be identified using CECAT and CETERM from the CE SDTM dataset.	Display underlined terms defined by the following adjudicated terms: • CV Death which includes all CETERM values with CEDECOD = "Cardiovascular death" or CETERM="Death due to stroke" • Non-fatal Myocardial infarction • Non-fatal Stroke	Y
	Extended MACE	Adjudicated terms will be identified (for MACE +) using CECAT and CETERM from the CE SDTM dataset.	Display underlined terms from MACE and underlined terms below: • Hospitalization for Unstable Angina • Coronary Revascularization Procedures	N

ASI Grouping	Categories (ASI)	Search Criteria	Terms to Display	Include in AE Overview (Y/N)
Serious infections, Tuberculosis, opportunistic infections and herpes zoster	Serious infections	Serious AEs within the Infections and Infestations SOC	PTs	Y
	Active Tuberculosis	Active Tuberculosis CMQ (code 80000188)	PTs	Y
	Opportunistic infections excluding tuberculosis and herpes zoster	Opportunistic infections excluding tuberculosis and herpes zoster CMQ (code 80000189)	PTs	Y
	Herpes Zoster	Herpes zoster CMQ (code 80000175)	PTs	N
Malignancies	Malignant Tumours	Narrow Malignant tumours (SMQ 20000194)	PTs	Y
	Non-melanoma skin cancer (NMSC)	Broad Skin malignant tumours (SMQ 20000204) excluding terms identified by the Melanoma CMQ (code 80000119)	PTs	N
	Malignancies excluding NMSC	'Malignancies excluding NMSC' is identified by the 'Malignant Tumours' search excluding terms identified by the 'Non-melanoma skin cancer' (NMSC) search.	PTs	Y

ASI Grouping	Categories (ASI)	Search Criteria	Terms to Display	Include in AE Overview (Y/N)
Hypersensitivity Reaction	Hypersensitivity	Narrow Hypersensitivity (SMQ 20000214)	PTs	N
	Serious hypersensitivity	Serious AE within narrow Hypersensitivity (SMQ 20000214)	PTs	Y
Anaphylactic Reaction	Adjudicated Anaphylactic Reaction	Adjudicated terms will be identified using SDTM data (e.g., CE and PR domains).	Adjudicated term defined in the charter	Y
	Serious anaphylactic reactions	Serious AE within narrow Anaphylactic reaction (SMQ 20000021);	PTs	Y

ASI Grouping	Categories (ASI)	Search Criteria	Terms to Display	Include in AE Overview (Y/N)
Hepatic events	Hepatic events	Broad Hepatic failure, fibrosis and cirrhosis and other liver damage-related conditions (SMQ 20000013) Broad Hepatitis, non-infectious (SMQ 20000010) Broad Cholestasis and jaundice of hepatic origin (SMQ 20000009) Broad Liver related investigations, signs and symptoms (SMQ 20000008) Narrow Liver-related coagulation and bleeding disturbances (SMQ 20000015)	PTs	N
Injection site reactions		Injection site reaction CMQ (code 80000019)	PTs	N
		Injection site reaction assessment and infusion site reaction assessment CRFs	Terms on the CRF	

Appendix C. Potentially Clinically Significant Criteria for Safety Endpoints

The criteria for Potentially Clinically Significant (PCS) laboratory findings are described in Table C-1 and Table C-2, and the PCS criteria for vital sign findings are described in Table C-3.

Table C-1. Criteria for Potentially Clinically Significant Hematology Values

		Definition of Grade 3 or Greater
Hematology Variables	**Units**	**Very Low**
Hemoglobin	g/dL	< 8.0
Platelets count	10^9/L	< 50.0
WBC count	10^9/L	< 2.0
Neutrophils	10^9/L	< 1.0
Lymphocytes	10^9/L	< 0.5

Note: A post-baseline value must be a higher grade than the baseline CTC grade to be considered a potentially clinically significant finding.

Table C-2. Criteria for Potentially Clinically Significant Chemistry Values

		Definition of Grade 3 or Greater	
Chemistry Variables	**Units**	**Very Low**	**Very High**
TBL	mcmol/L		> 3.0 × ULN
SGOT/AST	U/L		> 5.0 × ULN
SGPT/ALT	U/L		> 5.0 × ULN
Albumin	g/L	< 20	
Glucose	mmol/L	< 2.2	> 13.9
Triglycerides	mmol/L		> 5.7
Creatinine	mcmol/L		> 3.0 × ULN (> 3.0 × BL)
Sodium	mmol/L	< 130	> 155
Potassium	mmol/L	< 3.0	> 6.0
Calcium (corrected for albumin)	mmol/L	< 1.75	> 3.1
Total Cholesterol	mmol/L		> 10.34
GGT			> 5.0 × ULN
ALP			> 5.0 × ULN

Table C-2. Criteria for Potentially Clinically Significant Chemistry Values (Continued)

Note: A post-baseline value must be a higher grade than the baseline CTC grade to be considered a potentially clinically significant finding.

Table C-3. Criteria for Potentially Clinically Significant Vital Sign Values

Vital Sign	Category	Criteria for Potential Clinically Significant Vital Signs
Systolic blood pressure (mmHg)	Low	Value ≤ 90 mmHg and decrease ≥ 20 mmHg from baseline
	High	Value ≥ 160 mmHg and increase ≥ 20 mmHg from baseline
Diastolic blood pressure (mmHg)	Low	Value ≤ 50 mmHg and decrease ≥ 10 mmHg from baseline
	High	Value ≥ 100 mmHg and increase ≥ 10 mmHg from baseline

Appendix D. Crohn's Disease Activity Index (CDAI)

The Crohn's Disease Activity Index (CDAI) is a composite instrument that includes patient symptoms evaluated over 7 days (abdominal pain, stool frequency and general well-being), as well as physical and laboratory findings. These items are scored individually, weighted, and do not contribute equally to the overall score. The CDAI is derived from summing up the weighted individual scores of eight items as outlined below:

			Factor	Subtotal
1. Number of liquid or very soft stools (Record the frequency per day)	__+___+___+___+___+___+___+___=___ Days: 1 2 3 4 5 6 7 Sum	×	2	
2. Abdominal pain rating: 0 = none, 1 = mild, 2 = moderate, 3 = severe	__+___+___+___+___+___+___+___=___ Days: 1 2 3 4 5 6 7 Sum	×	5	
3. General well-being: 0 = generally well, 1 = slightly underpar, 2 = poor, 3 = very poor, 4 = terrible	__+___+___+___+___+___+___+___=___ Days: 1 2 3 4 5 6 7 Sum	×	7	
4. Number of 6 listed categories the subject now has **Check all items that apply:** ☐ Arthritis/arthralgia ☐ Iritis/uveitis ☐ Erythema nodosum/pyoderma gangrenosum/aphthous stomatitis ☐ Fissure, abscess and/or anal fistula (draining/non-draining) ☐ Other cutaneous fistula (draining/non-draining) Fistula ☐ Fever over 100°F (37.8°C) during past week	_____ _____ Record "0" if no categories checked	×	20	
5. Taking Lomotil/Imodium/ Loperamide/opiates for diarrhea 0 = no, 1 = yes	_____	×	30	
6. Abdominal mass 0 = none, 2 = questionable, 5 = defined	_____	×	10	
7. Hematocrit: ___.__	Male: (47 – hematocrit) = Female: (42 – hematocrit) = Subtotal If hematocrit > normal, enter "0"	×	6	

			Factor	Subtotal
8. Body weight: ___.__(kg) Standard weight: ___.__(kg)	100 × [1 – (Body wt/Standard wt)] = Percent below standard weight: _____ If body wt > std. wt, enter "0"	×	1	
			Total	

CDAI with higher score indicates more severe disease. For the calculation of CDAI and components at each visit, the following rules will be applied:

- The 7 most recent useable days out of the 14 days preceding the visit will be used.
- If 7 useable days are not available, an average will be calculated based on the number of days with available data as follows:
  - Data for days with missing diary entries will be imputed using the average of the non-missing days, in order to calculate a CDAI and components based on 7 days.
  - An average for the most recent 6 days will be calculated if data for only 6 days are available
  - An average for the most recent 5 days will be calculated if data for only 5 days are available
  - An average for the most recent 4 days will be calculated if data for only 4 days are available
- If the minimum number of days of diary data (i.e., 4 days for CDAI and components) are not available, then the subject's score for that visit will be considered missing.

The CDAI score is set to missing under the following conditions:

1. If there are less than 4 days of any diary data.
2. If any component other than the diary data is missing.

Appendix E. Simple Endoscopic Score – CD (SES-CD)

SES-CD is calculated based the sum of individual segment values for four endoscopic variables (presence and size of ulcers, ulcerated surface, affected surface and presence of narrowing). Each variable in each segment be scored 0 to 3 resulting in SES-CD values ranging from 0 to 56 with higher scores indicating more severe disease.

Variable	Score			
	0	1	2	3
Size of ulcers (cm)	None	Aphthous ulcers (diameter 0.1–0.5)	Large ulcers (diameter 0.5–2)	Very large ulcers (diameter >2)
Ulcerated surface (%)	None	<10	10–30	>30
Affected surface (%)	Unaffected segment	<50	50–75	>75
Presence of narrowings	None	Single, can be passed	Multiple, can be passed	Cannot be passed

*Total SES-CD: sum of the values of the 4 variables for the 5 bowel segments. Values are given to each variable and for every examined bowel segment (for example, rectum, left colon, transverse colon, right colon and ileum).

Medscape Source: Nat Rev Gastroenterol Hepatol ©2009 Nature Publishing Group

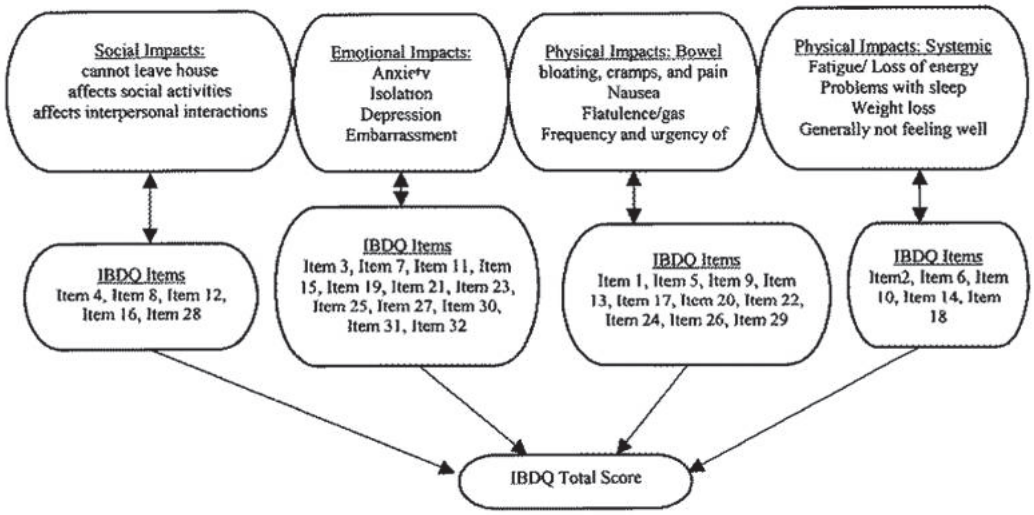
SES-CD Scoring:

	Rectum	Sigmoid and Left Colon	Transverse Colon	Right Colon	Ileum	Total
Size of Ulcers Enter: 0 if none 1 if aphthous ulcers (Ø 0.1 to 0.5 cm) 2 if large ulcers (Ø 0.5 to 2 cm) 3 if very large ulcers (Ø > 2 cm)						
Ulcerated Surface Enter: 0 if none 1 if < 10% 2 if 10% – 30% 3 if > 30%						
Affected Surface Enter: 0 if unaffected segments 1 if < 50% 2 if 50% – 75% 3 if > 75%						

	Rectum	Sigmoid and Left Colon	Transverse Colon	Right Colon	Ileum	Total
Presence of Narrowing Enter: 0 if none 1 if single, can be passed 2 if multiple, can be passed 3 if cannot be passed						
					TOTAL =	

Appendix F. Inflammatory Bowel Disease Questionnaire (IBDQ)

The IBDQ is a 32-item (ranges 1 – 7) self-report questionnaire for patients with IBD to evaluate the patient reported outcomes across 4 dimensions: bowel symptoms (loose stools, abdominal pain), systemic symptoms (fatigue, altered sleep pattern), social function (work attendance, need to cancel social events), and emotional function (anger, depression, irritability). The IBDQ total Score ranges from 32 to 224 with a higher score indicating better outcome.



The derivation of the IBDQ total score and the four IBDQ domain scores are as follows:

Domain Scores:

- S1. BOWEL SYMPTOM: Q1 + Q5 + Q9 + Q13 + Q17 + Q20 + Q22 + Q24 + Q26 + Q29
  [score ranges from 10 - 70],
- S2. SOCIAL FUNCTION: Q4 + Q8 + Q12 + Q16 + Q28
  [score ranges from 5 - 35],
- S3. SYSTEMIC SYMPTOM: Q2 + Q6 + Q10 + Q14 + Q18
  [score ranges from 5 - 35],

- S4. EMOTIONAL FUNCTION: Q3 + Q7 + Q11 + Q15 + Q19 + Q21 + Q23 + Q25 + Q27 + Q30 + Q31 + Q32
  [score ranges from 12 - 84].

Total Score:

IBDQ Total Score = SUM of (bowel symptom domain score, social function domain score, systemic symptom domain score, emotional function domain score).

The following convention applies to Inflammatory Bowel Disease Questionnaire (IBDQ):

When not more than 20% of items in a domain of IBDQ were missing, it was substituted with the mean values from the items completed in the particular domain; otherwise, they were treated as missing. The 20% threshold in each domain is: bowel symptom domain: 2 items, systemic symptom domain: 1 item, social function domain: 1 item, emotional function domain: 2 items. If any of the 4 domain scores is missing, the total IBDQ score will be set to missing.

Appendix G. FACIT-Fatigue

The Functional Assessment of Chronic Illness Therapy (FACIT) measurement system is a comprehensive compilation of questions that measure health related quality of life in patients with cancer and other chronic diseases (www.facit.org). The FACIT-Fatigue scale is a symptom-specific subscale of FACIT and is composed of 13 fatigue-related questions as follows. The response to each question is as follows: 0=Not at all, 1=A little bit, 2=Somewhat, 3=Quite a bit, 4=Very much.

1. I feel fatigued
2. I feel weak all over
3. I feel listless ("washed out")
4. I feel tired
5. I have trouble starting things because I am tired
6. I have trouble finishing things because I am tired
7. I have energy
8. I am able to do my usual activities
9. I need to sleep during the day
10. I am too tired to eat
11. I need help doing my usual activities
12. I am frustrated by being too tired to do the things I want to do
13. I have to limit my social activity because I am tired

In order to have higher values representing lower level of fatigue, all of the items, except for item 7 "I have energy" and item 8 "I am able to do my usual activities," are assigned reversed scores: reversed score = 4 – raw score. The total FACIT-Fatigue score will then be calculated as follows:

- FACIT-Fatigue score = 13×[Sum of answered-item scores / Number of items answered]

The score ranges from 0 to 52, 52 being the lowest level of fatigue.

When there are missing data, provided that more than 50% of the items (i.e., at least 7 of 13 items) were answered in FACIT-Fatigue questionnaire, the total score will not be deemed as missing but be calculated as the mean response for all the non-missing-value items. If less than 50% of the items were answered, the total score will have a missing value.

Appendix H. Random Seeds

In case of non-convergence, the random seed will be updated by adding 100000 at each attempt until convergence of model happens.

Table H-1. Random Seeds for NRI-C

Variables	Random Seed	
	MCMC Procedure	PROC MI
CDAI	50001	60001
SES-CD (for 20 individual variables)	50002	60002
SF score	50003	60003
AP score	50004	60004
EIM	50005	60005
FACIT-Fatigue	50006	60006
IBDQ total score	50009	60009
PGIS	50010	60010
PGIC	50011	60011

Appendix I. Geographic Region

Below table lists the countries/regions considered for each geographic region.

Geographic Region	Countries/Regions
Asia	Japan (JPN), China (CHN), Hong Kong (HKG), Malaysia (MYS), Singapore (SGP), South Korea (KOR), Taiwan (TWN)
Eastern Europe	Belarus (BLR), Bosnia and Herzegovina (BIH), Bulgaria (BGR), Croatia (HRV), Czech Republic (CZE), Estonia (EST), Hungary (HUN), Kazakhstan (KAZ), Latvia (LVA), Lithuania (LTU), Poland (POL), Romania (ROU), Russia (RUS), Serbia (SRB), Slovakia (SVK), Slovenia (SVN), Turkey (TUR), Ukraine (UKR)
North America	Canada (CAN), United States (USA), Puerto Rico (PRI)
Other	Australia (AUS), Egypt (EGY), Israel (ISR), New Zealand (NZL), South Africa (ZAF), Tunisia (TUN)
South/Central America	Argentina (ARG), Brazil (BRA), Chile (CHL), Colombia (COL), Mexico (MEX), Guatemala (GTM)
Western Europe	Austria (AUT), Belgium (BEL), Denmark (DNK), Finland (FIN), France (FRA), Germany (DEU), Greece (GRC), Ireland (IRL), Italy (ITA), Netherlands (NLD), Norway (NOR), Portugal (PRT), Spain (ESP), Switzerland (CHE), Sweden (SWE), United Kingdom (GBR)

Appendix J. Definition of Estimand for Primary and Ranked Secondary Endpoints

Estimand	Attributes of the Estimand				
	Population	**Endpoint(s)**	**Treatment**	**Intercurrent Events**	**Statistical Summary**
Primary	ITT1A population	For US specific protocol: Achievement of CDAI clinical remission at Week 12. Achievement of endoscopic response at Week 12. For global protocol outside US: Achievement of clinical remission at Week 12. Achievement of endoscopic response at Week 12.	Risankizumab 1200 mg IV vs. placebo Risankizumab 600 mg IV vs. placebo	IE1: premature discontinuation of study drug IE2: Initiation or dose escalation of CD-related corticosteroids All data after IE1 will be used. All subjects will be considered as non-responders at or after IE2	Proportion of subjects achieving CDAI clinical remission. Proportion of subjects achieving clinical remission. Proportion of subjects achieving endoscopic response.
Binary Key Secondary	ITT1A population	Binary secondary endpoints as defined in Section 3.2 except for CD-related hospitalization endpoint	Risankizumab 1200 mg IV vs. placebo Risankizumab 600 mg IV vs. placebo	IE1: premature discontinuation of study drug IE2: Initiation or dose escalation of CD-related corticosteroids All data after IE1 will be used. All subjects will be considered as non-responders at or after IE2	Proportion of subjects achieving each binary secondary endpoint

Estimand	Attributes of the Estimand				
	Population	**Endpoint(s)**	**Treatment**	**Intercurrent Events**	**Statistical Summary**
Continuous Key Secondary	ITT1A population	Change from Baseline in IBDQ total score/FACIT-F score at Week 12	Risankizumab 1200 mg IV vs. placebo Risankizumab 600 mg IV vs. placebo	IE1: premature discontinuation of study drug IE2: Initiation or dose escalation of CD-related corticosteroids All data after IE1 will be used. All data after IE2 will not be used. MMRM will be the primary approach for the analysis.	Mean change from Baseline in IBDQ total score and FACIT-F score
CD-related hospitalization	ITT1A population	Occurrence of CD-related hospitalization through Week 12	Risankizumab 1200 mg IV vs. placebo Risankizumab 600 mg IV vs. placebo	IE1: premature discontinuation of study drug IE2: Initiation or dose escalation of CD-related corticosteroids All data after IE1 will be used. All data after IE2 will be used	Proportion of subjects having at least one CD-related hospitalization through Week 12

In addition, a supplementary analysis corresponding to the AO analysis specified in Section 8.3.2 will be conducted for the primary and key binary secondary endpoints in which all data after IE1 and IE2 will be included in the analysis.