Janssen Research & Development

Statistical Analysis Plan
Amendment 1

A Randomized, Double-blind, Placebo-controlled, Parallel-group, Multicenter Study to Evaluate the Efficacy and Safety of Guselkumab Subcutaneous Induction Therapy in Participants With Moderately to Severely Active Crohn's Disease

Protocol Amendment 1 CNTO1959CRD3004 (GRAVITI); Phase 3

CNTO1959 (guselkumab)

Status:	Approved
Date:	2 October 2023
Prepared by:	Janssen Research & Development, LLC
Document No.:	EDMS-RIM-986653, 2.0

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

TABLE OF CONTENTS

VERSION HISTORY

Table 1 – SAP Version History Summary

SAP Version	Approval Date	Change	Rationale
1	21-Mar-2023	Not Applicable	Initial release
	CCI [redacted]	[redacted]	[redacted]
	4	Added "Note that the 3 randomized subjects at site W30-PL10064 were excluded from all analysis sets because the subject identifications could not be verified, the source data could not be verified, and study intervention administration could not be confirmed."	Modified analysis datasets due to audit findings
	5.3.1	Updated the requirements for the alternative approach for CDAI calculation.	Clarification
	CCI [redacted]	[redacted]	[redacted]
	CCI [redacted]	[redacted]	[redacted]
	CCI [redacted]	[redacted]	[redacted]
	5.4.2.1	Clarified that the ICE 6 criteria are determined by IWRS	Clarification
	5.5.1.1	Provided definitions for the following two endpoints: Corticosteroid-free Clinical Remission and 90-day Corticosteroid-free Clinical Remission	Clarification
	CCI [redacted]	[redacted]	[redacted]
	5.5.1.4	Removed the definitions for "abnormal" CRP and Fecal calprotectin	Categorical analyses of biomarkers will be defined by pre-specified cut points.
	5.5.1.6, 5.5.3	Rephrased definition of RHI histologic response and RHI histologic remission.	Clarifications
	5.5.2.7	Added endpoint: • Clinical remission and ≥50% reduction from baseline in CRP or fecal calprotectin at Week 12, Week 24, and Week 48 • Clinical remission and CRP concentration ≤5 mg/L or fecal calprotectin concentration ≤250 µg/g	Modified analysis plan

SAP Version	Approval Date	Change	Rationale
		over time, among participants with elevated CRP and fecal calprotectin >250 µg/g at baseline	
	CCI	[redacted]	[redacted]
	5.5.2.2, 5.5.2.5, 5.5.2.9	Deleted endpoints: • Clinical response at Week 12 and corticosteroid-free clinical remission at Week 48 • Clinical response at Week 12 and 90-day corticosteroid-free clinical remission at Week 48 • Clinical response at Week 12 and endoscopic response at Week 48 • Clinical response at Week 12 and endoscopic remission at Week 48 • Clinical response at Week 12 and endoscopic remission (Alternative definition) at Week 48 • Clinical response at Week 12 and deep remission at Week 48 • Clinical response at Week 12 and deep remission (Alternative definition) at Week 48 • Change from baseline in average daily number of BSFS types 5, 6, and 7 stools over time through Week 48. • A ≥2 reduction in average daily number of BSFS types 5, 6, and 7 stools over time through Week 48.	Modified analysis plan
	5.5.3	Added approaches in case of lack of convergence of the unstructured covariance matrix.	Clarification
	CCI	[redacted]	[redacted]
	5.6.3.1, 6.3 Appendix 3	Refer to "NCI-CTCAE" grade as a modified NCI-CTCAE grade. Updated definition of toxicity grade for hemoglobin to clarify that the ranges are absolute values above the ULN. Updates to CTCAE grading for Potassium decrease and Sodium decrease.	Clarifies that the definitions of the toxicity grade used is based on only laboratory values, and without clinical symptoms. To clarify definition of toxicity grades.
	5.7.2, 5.7.4	Provided details about additional analyses to explore the relationship between antibodies to guselkumab and efficacy outcomes and guselkumab PK, as well as related ICE strategies.	To provide additional planned analyses and clarification on ICE strategy.
	5.7.6	Added health economics endpoints. • Proportion of participants with CD-related surgeries and procedures through Week 12, Week 24, Week 48, and Week 96 • Time to first CD-related surgery through Week 24, Week 48, and Week 96	Added based on protocol

SAP Version	Approval Date	Change	Rationale
		• Time to first CD-related procedure through Week 24, Week 48, and Week 96	
	Throughout the document	Minor grammatical, formatting, and spelling changes, and changes to ensure consistency across the document were made.	Consistency changes

1. INTRODUCTION

CNTO1959CRD3004 (GRAVITI) is a randomized, double-blind, placebo-controlled, parallel-group, multicenter Phase 3 study to evaluate the efficacy and safety of guselkumab subcutaneous (SC) induction dosing in participants with moderately to severely active Crohn's disease.

This Statistical Analysis Plan (SAP) contains definitions of analysis sets, derived variables, and statistical methods for all planned analyses for GRAVITI.

1.1. Objectives

The GRAVITI objectives are as follows:

Primary Objectives

- To evaluate the efficacy, including clinical remission at Week 12 and endoscopic response at Week 12, of guselkumab SC induction

Secondary Objectives

- To evaluate the efficacy of guselkumab SC induction across a range of outcome measures (clinical response at Week 12, clinical remission at Week 24, and PRO-2 remission at Week 12)
- To evaluate the safety of guselkumab SC induction

Exploratory Objectives

- To further evaluate the efficacy of guselkumab SC induction across a range of outcome measures
- To evaluate the impact of guselkumab SC induction on biomarkers
- To evaluate the PK and immunogenicity of guselkumab SC induction
- To evaluate the impact of guselkumab SC induction on PROs

Specific primary, secondary, and other endpoints for the study are provided in Section 5.

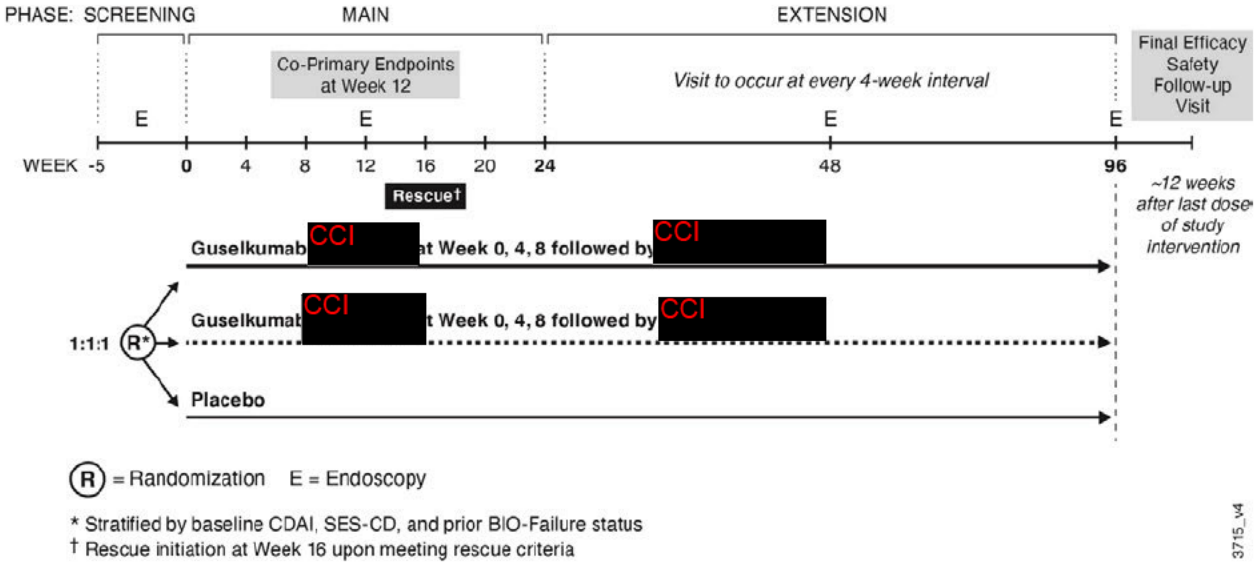
1.2. Study Design

An overview of the study design is provided in Figure 1.

Target Population

The target population is adult participants with moderately to severely active Crohn's disease (of at least 3 months duration) with colitis, ileitis, or ileocolitis previously confirmed by radiography, histology, and/or endoscopy. To be eligible for the study, participants must also have endoscopic evidence of active Crohn's disease and have demonstrated an inadequate response or failure to tolerate previous conventional or biologic therapy.

Figure 1: Schematic Overview of the Study



Active Disease Criteria

At baseline, participants must have active Crohn's disease, defined as follows:

1. **Clinically active Crohn's disease**
   a. CDAI score ≥220 but ≤450

      AND

   b. Having active symptoms defined as EITHER
      - Mean daily Stool Frequency (SF) count ≥4, based on the unweighted CDAI component of the number of liquid or very soft stools

      OR
      - Mean daily Abdominal Pain (AP) score ≥2, based on the unweighted CDAI component of AP

AND

2. **Endoscopic evidence of ileocolonic Crohn's disease**

   A SES-CD score ≥6 (or ≥4 for participants with isolated ileal disease), as assessed by central endoscopy reading at the screening endoscopy, based on the presence of ulceration in at least 1 of the 5 ileocolonic segments, resulting in the following specified ulceration component scores:

   a. a minimum score of 1 for the component of "size of ulcers"

   AND

   b. a minimum score of 1 for the component of "ulcerated surface".

Medication History Criteria

The two groups based on prior therapies comprising the target population are briefly described below.

Conventional therapy failure or intolerance (CON-Failure)

- Participants must have demonstrated an inadequate response to, or have failed to tolerate, at least 1 of the following conventional Crohn's disease therapies: oral corticosteroids (including prednisone, budesonide, and beclomethasone dipropionate) or the immunomodulators azathioprine (AZA), 6-mercaptopurine (6-MP) or methotrexate (MTX).
- Participants who have demonstrated corticosteroid dependence (ie, an inability to successfully taper corticosteroids without a return of the symptoms of Crohn's disease) are also eligible.
- Participants may either be naïve to biologic therapy (ie, infliximab, adalimumab, certolizumab pegol, vedolizumab, or approved biosimilars for these agents) or may have been exposed to biologic therapy and have not demonstrated inadequate response or intolerance.

Biologic therapy failure or intolerance (BIO-Failure)

- Participants must have demonstrated an inadequate response to, or have failed to tolerate, at least 1 or more biologic therapies (ie, infliximab, adalimumab, certolizumab pegol, vedolizumab, or approved biosimilars for these agents) at a dose that is, at minimum, a locally approved dose for the treatment of Crohn's disease.

  Inadequate response is defined as: primary nonresponse (ie, no initial response) or secondary nonresponse (ie, response initially but subsequently lost response).

- Participants with prior exposure to IL-12/23 or IL-23 agents are ineligible for this study.

Participants who had an inadequate response or failure to tolerate biologic therapy will comprise approximately 35% to 65% of the population.

At Week 0, a target of 318 total participants will be randomly assigned in a 1:1:1 ratio to one of the following SC treatments using a permuted block randomization:

- Guselkumab CCI at Weeks 0, 4, and 8 followed by guselkumab CCI q4w
- Guselkumab CCI at Weeks 0, 4, and 8 followed by guselkumab CCI q8w
- Placebo SC q4w

The randomization will be stratified by baseline Crohn's Disease Activity Index (CDAI) score (≤300 or >300), baseline Simple Endoscopic Score for Crohn's Disease (SES-CD) score (≤12 or >12), and prior BIO-Failure status (Yes or No).

All participants in the placebo group who meet at least 1 of the following rescue criteria will receive rescue medication:

- CDAI score >220 and <70-point reduction from baseline CDAI at both Week 12 and Week 16

OR

- SES-CD score increase by at least 50% from baseline at Week 12

Upon meeting at least 1 of the rescue criteria, participants in the placebo group will receive guselkumab [CCI redacted] at Weeks 16, 20, and 24 followed by guselkumab [CCI redacted] q8w.

To maintain the blind, participants randomized to guselkumab who meet at least 1 of the rescue criteria will continue their assigned treatment regimen and receive blinded sham rescue with a matching placebo SC injection. For more details, refer to Section 6.5.1 of the protocol. Each active study intervention and its matching placebo will be identical in appearance.

At Week 24, all participants will enter the extension phase and receive the same treatment regimen that they were receiving at Week 24. Database locks are planned for Week 24, Week 48, and when the last participant completes the last scheduled assessment. Additional DBLs may be added as necessary.

The co-primary endpoints of this study are clinical remission at Week 12 and endoscopic response at Week 12 (refer to Section 5.3 for endpoint definitions and analyses). Additionally, there are 3 secondary endpoints in the study (refer to Section 5.4 for definitions and analyses methods). The primary and secondary endpoints will be analyzed at the first DBL, which occurs at Week 24 DBL. Only data through Week 24 will be included in the Week 24 DBL for analyses.

The study will be unblinded to investigators and participants after the last participant completes the Week 48 evaluations and the Week 48 database lock (DBL) is completed. Upon study unblinding, placebo participants who have not been rescued with guselkumab will be discontinued from study intervention and have a final efficacy and safety (FES) visit. All other participants will continue on guselkumab treatment through Week 96.

In general, participants who are receiving oral 5-aminosalicylic acid (5-ASA) compounds, oral corticosteroids, conventional immunomodulators (AZA, 6-MP, or MTX), antibiotics, and/or enteral nutrition for the treatment of Crohn's disease at baseline should maintain a stable dose for the specified period before baseline, and through Week 48, with the exception of oral corticosteroids. Starting at Week 12, all participants who were taking corticosteroids at Week 0 must begin tapering their corticosteroid dose. This tapering is mandatory, unless not medically feasible.

From Week 0 through Week 48, enrolled participants should not initiate any of the following concomitant Crohn's disease-specific medical therapies:

- Oral or rectal 5-ASA compounds
- Immunomodulators (ie, AZA, 6-MP, or MTX)
- Oral, parenteral, or rectal corticosteroids, including budesonide and beclomethasone dipropionate
- Antibiotics as a primary treatment for Crohn's disease
- Total parenteral nutrition or enteral nutrition as a treatment for Crohn's disease

Details regarding the prohibited changes in concomitant Crohn's disease-specific medications for participants are included in Section 6.2 Appendix 2.

From Week 12 through Week 48, participants may transiently use (ie, for <4 weeks) increased doses of corticosteroids for reasons other than treatment for Crohn's disease (eg, stress doses of corticosteroids for surgery, asthma, adrenocortical insufficiency).

After Week 48 and through Week 96, concomitant therapies for Crohn's disease including 5-ASAs, corticosteroids, antibiotics, and immunomodulators (ie, AZA, 6-MP, or MTX), and/or total parenteral or enteral nutrition may be administered and changed at the discretion of the investigator.

Participants who discontinue study intervention early should return for a study intervention discontinuation (SID) visit. All randomized participants should complete the FES follow-up visit approximately 12 weeks after the last dose of study intervention.

Randomization and Blinding

Randomization will be used to minimize bias in the assignment of participants to treatment groups, to increase the likelihood that known and unknown participant attributes (e.g., demographic and baseline characteristics) are evenly balanced across treatment groups, and to enhance the validity of statistical comparisons across treatment groups. Blinded treatment will be used to reduce potential bias during data collection and evaluation of clinical endpoints.

Central randomization will be implemented in this study. It is based on a computer-generated randomization schedule prepared before the study by or under the supervision of the sponsor. Allocation to treatment groups will be performed using an interactive web response system (IWRS). The IWRS will assign a unique treatment code, which will dictate the treatment assignment and matching study intervention kit(s) for the participant.

At Week 0, a target of 318 total participants will be randomly assigned in a 1:1:1 ratio to one of the treatment groups using a permuted block randomization. The randomization will be stratified by baseline Crohn's Disease Activity Index (CDAI) score (≤300 or >300), baseline Simple Endoscopic Score for Crohn's Disease (SES-CD) score (≤12 or >12), and prior BIO-Failure status (Yes or No).

To maintain the study blind, the study intervention container will have a label containing the study name, study intervention number, and reference number. The study intervention number will be entered in the eCRF when the study intervention is dispensed. Each active study intervention and its matching placebo will be identical in appearance and will be packaged in identical containers. All participants will receive the same device(s), which could be either active or matching placebo at 4-week intervals (double-dummy) in order to maintain treatment blinding.

The investigator will not be provided with randomization codes. The codes will be maintained within the IWRS, which has the functionality to allow the investigator to break the blind for an individual participant. In a case of unblinding, the date and reason for the unblinding must be

documented in the appropriate section of the eCRF and/or in the source document. The documentation received from the IWRS indicating the code break must be retained with the participant's source documents in a secure manner.

Data that may potentially unblind the intervention assignment (ie, study intervention serum concentrations, anti-guselkumab antibodies) will be handled with special care to ensure that the integrity of the blind is maintained and the potential for bias is minimized. This can include making special provisions, such as segregating the data in question from view by the investigators, clinical team, or others as appropriate until the time of unblinding.

Additionally, a given participant's treatment assignment may be unblinded to the sponsor, the Independent Ethics Committee/Institutional Review Board (IEC/IRB), and site personnel to fulfill regulatory reporting requirements for suspected unexpected serious adverse reactions (SUSARs).

Participants who have had their intervention assignment unblinded by the investigator will not be eligible to receive further study intervention.

The Sponsor will remain blinded to participant-level treatment assignment through Week 48 with the exception of the Week 24 DBL, when a limited number of Sponsor personnel will be unblinded and have access to the treatment assignment for analysis. Identification of sponsor personnel who will have access to the unblinded participant-level data will be documented in an unblinding plan before the Week 24 DBL. A separate study team will be put in place to manage the conduct of the study after the Week 24 DBL. This study team will remain blinded to treatment assignment until the Week 48 DBL.

Treatment assignment will remain blinded to the study sites including site monitors and participants until after the last participant completes the Week 48 evaluations and the Week 48 DBL is completed.

2. STATISTICAL HYPOTHESES

The co-primary hypotheses of this study are that guselkumab is superior to placebo in inducing:

- clinical remission (CDAI score <150) at Week 12; and
- endoscopic response (≥50% improvement from baseline in the SES-CD score) at Week 12

The secondary hypotheses of the study are that guselkumab is superior to placebo in achieving:

- clinical remission at Week 24
- PRO-2 remission at Week 12 (defined as an AP mean daily score at or below 1 and stool frequency (SF) mean daily score at or below 3, ie, AP ≤1 and SF ≤3, and no worsening of AP or SF from baseline)
- clinical response (decrease from baseline in CDAI ≥100 points or clinical remission) at Week 12

The co-primary endpoints and secondary endpoints will be analyzed based on the estimands defined in sections 5.3.2.1 and 5.4.2.1, considering treatment groups, population, variable (endpoint), intercurrent event (ICE) strategies, and population-level summary.

Statistical testing will be performed at a significance level of 0.05 (2-sided). The Type I error rate will be controlled over the co-primary endpoints, secondary endpoints, and two Week 48 exploratory endpoints (clinical remission at Week 48, endoscopic response at Week 48). See more details in Section 5. Note the exploratory endpoints are referred to as tertiary endpoints in the protocol.

The study will be considered successful if the tests for both co-primary endpoints are positive.

The co-primary and secondary endpoints (as well as all endpoints through Week 24) will be evaluated at the Week 24 DBL. The evaluation of the Week 48 endpoints, including the Week 48 exploratory endpoints that are included in the testing procedure, will be evaluated at the subsequent Week 48 DBL.

3. SAMPLE SIZE DETERMINATION

Sample sizes were determined by the power to detect a significant difference in clinical remission at Week 12 and in endoscopic response at Week 12 (co-primary endpoints) between the combined guselkumab group (CCI q4w x 3 → CCI q4w group and CCI q4w x 3 → CCI CCI q8w group combined) and the placebo group, using a 2-sided chi-square test with 0.05 significance level. The assumed rates are 50% versus 15% (guselkumab versus placebo) for clinical remission at Week 12 and 30% versus 13% for endoscopic response at Week 12. The study is sized such that the guselkumab therapy achieves >90% power for the co-primary endpoints compared with placebo. This sample size also provides >90% power for all secondary endpoints.

Clinical Remission at Week 12

Assumptions for clinical remission at Week 12 were based on the Phase 2 data from the CNTO1959CRD3001 Phase 2/3 program. In the Phase 2 study (GALAXI 1), the proportions of participants who met the eligibility requirement for the GRAVITI study and who were in clinical remission at Week 12, were 12.0%, 59.2%, 55.6% and 42.9% for placebo, guselkumab CCI C CI, guselkumab CCI and guselkumab CCI, respectively, for a treatment difference of approximately 31% to 47%. Of note, in GALAXI 1, approximately 50% of the participants were BIO-Failure.

Note that the eligibility criterion on SES-CD for this study is relatively new and not many historical studies have this requirement, hence the actual placebo rate could vary. Based on these considerations and the data above, the clinical remission rates at Week 12 are assumed to be 15% for placebo and 50% for guselkumab CCI (treatment difference of 35%), assuming about 50% of the participants will be BIO-Failure in this study.

Endoscopic Response at Week 12

Assumptions for endoscopic response at Week 12 were also based on GALAXI 1. The proportions of participants who met the eligibility requirement for this current study and who were in

endoscopic response at Week 12, were 12.0%, 34.7%, 35.6% and 22.4% for placebo, guselkumab CCI, guselkumab CCI and guselkumab CCI respectively, for a treatment difference of approximately 10% to 24%.

Based on these data, the endoscopic response rates are assumed to be 13% for placebo and 30% for guselkumab CCI (treatment difference of 17%), assuming about 50% of the participants will be BIO-Failure in this study.

In the GRAVITI study, 212 participants in the combined guselkumab group and 106 participants in the placebo group will provide at least 90% power for achieving the co-primary endpoints of clinical remission at Week 12 and endoscopic response at Week 12 (Table 1), based on the assumptions above.

Table 1: Power to Detect a Treatment Effect of Guselkumab Based on Proportion of Participants Achieving Clinical Remission at Week 12 or Endoscopic Response at Week 12

Co-primary Endpoint of Clinical Remission at Week 12		
Placebo (N=106)	**Guselkumab SC (N=212)**	**Power**
15%	45%	>99%
15%	**50%**	**>99%**
15%	55%	>99%
Co-primary Endpoint of Endoscopic Response at Week 12		
10%	25%	91%
13%	**30%**	**94%**
13%	35%	>99%
Abbreviations: SC=subcutaneous		

4. ANALYSIS SETS

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

Population	Description
Randomized Analysis Set	All randomized participants
Full Analysis Set	All randomized participants who received at least 1 dose of study intervention
Safety Analysis Set	All randomized participants who received at least 1 dose of study intervention
PK Analysis Set	All randomized participants who received at least 1 dose of study intervention and have at least 1 valid blood sample available post baseline for PK analysis
Immunogenicity Analysis Set	All randomized participants who received at least 1 dose of study intervention and have appropriate samples for anti-drug antibody detection

Note that the 3 randomized subjects at site W30-PL10064 were excluded from all analysis sets because the subject identifications could not be verified, the source data could not be verified, and study intervention administration could not be confirmed.

Unless otherwise specified, all efficacy analyses will be based on the Full Analysis Set, all safety analyses will be based on the Safety Analysis Set, all PK analyses will be based on the PK Analysis Set, and all immunogenicity analyses will be based on the Immunogenicity Analysis Set.

Efficacy data will be analyzed according to randomized treatment group. In general, safety data will also be summarized according to their assigned treatment. However, participants assigned to placebo who incorrectly received guselkumab at any time will have their safety data analyzed in the guselkumab group from the time they received guselkumab.

5. STATISTICAL ANALYSES

5.1. General Considerations

Statistical testing will be performed at a significance level of 0.05 (2-sided). The Type I error rate will be controlled over the co-primary endpoints, secondary endpoints, and two Week 48 exploratory endpoints (clinical remission at Week 48, endoscopic response at Week 48). For endpoints that are not multiplicity-controlled, nominal p-values will be presented.

- Unless otherwise specified, data from all investigational centers/sites will be pooled for analysis.
- Study Day 1 refers to the day of the first study intervention administration. All efficacy and safety assessments at all visits will be assigned a day relative to this day.

  Study day for a visit is defined as:

  - Visit date - (date of Study Day 1) +1, if visit date is ≥ date of Study Day 1
  - Visit date - date of Study Day 1, if visit date < date of Study Day 1

  There is no 'Study Day 0'.
- Baseline is defined as the last observation prior to or on the day of the first study intervention, unless otherwise specified.

5.1.1. Visit Windows

Unless otherwise specified, nominal scheduled visits (described in SoA of the protocol) will be used for over time summaries and listings with no visit windows applied. However, the Unscheduled Visit (USV), Study Intervention Discontinuation (SID) and Final Efficacy and Safety (FES) visits will be slotted to scheduled visits according to the following mapping rules except for SES-CD endpoint:

1. Assign a visit number to USV, SID or FES visit based on the Study Day as illustrated in Table 2
2. If USV, SID or FES visit falls in the window of a scheduled visit and there are not already data for the scheduled visit, assign the scheduled visit number to the USV, SID or FES visit. If there are already data for a same scheduled visit, those data will be used in lieu of the USV, SID or FES visit data.

Table 2: SID/FES/Unscheduled Visit Slotting

Week 4	Day 2-42
Week 8	Day 43-70
Week 12	Day 71-98
Week 16	Day 99-126
Week 20	Day 127-154
Week 24	Day 155-182
Week 28	Day 183-210
Week 32	Day 211-238
Week 36	Day 239-266
Week 40	Day 267-294
Week 44	Day 295-322
Week 48	Day 323-350
Week 52	Day 351-378
…	
etc.	
Week 96	Day 659-686

More frequent measurements are taken for other endpoints compared to endoscopic assessments taken only at Weeks 0, 12, 48, and 96.
For SES-CD endpoint using a wider window: +/- 4 weeks
Week 12: Days 57 to 113
Week 48: Days 309 to 365
Week 96: Days 645 to 701

5.2. Participant Disposition

The number of participants in each analysis set, including the Randomized Analysis Set, will be summarized by treatment group, combined guselkumab treatment group and overall. In addition, the distribution of participants by region, country and site ID will be presented.

The number and frequency of participants in the following disposition categories will be summarized by treatment group, combined guselkumab group, and overall based on the FAS.

Note that "major disruption-related discontinuations" and "major disruption-related terminations" are study intervention discontinuations or study terminations that were due to unforeseen major disruptions such as a regional crisis.

- Participants who received study intervention by visit
- Participants discontinuing **study intervention** prior to Week 12 and reasons for discontinuation, including those due to COVID-19 related events and other major disruption-related discontinuations.
- Participants discontinuing **study intervention** prior to Week 24 and reasons for discontinuation, including those due to COVID-19 related events and other major disruption-related discontinuations.
- Participants discontinuing **study intervention** prior to Week 48 and reasons for discontinuation, including those due to COVID-19 related events and other major disruption-related discontinuations.

- Participants discontinuing **study intervention** prior to Week 92 and reasons for discontinuation, including those due to COVID-19 related events and other major disruption-related discontinuations.
- Participants who terminated **study participation** prior to Week 12 and reasons for termination, including those due to COVID-19 related events and other major disruption-related terminations.
- Participants who terminated **study participation** prior to Week 24 and reasons for termination, including those due to COVID-19 related events and other major disruption-related terminations.
- Participants who terminated **study participation** prior to Week 48 and reasons for termination, including those due to COVID-19 related events and other major disruption-related terminations.
- Participants who terminated **study participation** prior to the end of study and reasons for termination, including those due to COVID-19 related events and other major disruption-related terminations.
- All participants who met the rescue criteria.

Listings of participants will be provided for the following categories:

- Participants who discontinued study intervention
- Participants who terminated study participation
- Participants who were randomized yet did not receive study intervention
- Participants who were unblinded prior to study unblinding

Study Assessment Compliance will be summarized and listed by randomized treatment group. This will include the number of participants who missed at least one scheduled CDAI assessment, at least one scheduled endoscopy assessment, intercurrent events, or at least one study intervention administration due to any reason, including COVID-19 related events and other major disruptions.

5.3. Co-Primary Endpoints Analysis

5.3.1. Definition of Endpoints

The co-primary endpoints are:

- Clinical remission at Week 12
- Endoscopic response at Week 12

Clinical remission is defined as a CDAI score <150.

Endoscopic response is defined as ≥50% improvement from baseline in Simple Endoscopic Score for Crohn's Disease [SES-CD].

CDAI score

The CDAI is a validated multi-item measure of severity of illness derived as a weighted sum of 8 different Crohn's disease-related variables. These 8 variables are:

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

Calculation of the CDAI score:

Two approaches to calculating the CDAI score are provided below. The main approach will be used for the analysis of all endpoints that include the CDAI; of particular note is that this main approach is used for the co-primary and secondary endpoints that involve the CDAI. Additional analyses of co-primary and secondary endpoints that involve the CDAI will also be provided using the alternative approach.

Main approach

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

The CDAI score calculation algorithm is presented in Section 10.8 Appendix 8 of the study protocol. At each timepoint, the CDAI score will only be calculated if ≥ 4 of the 8 components are available. If the CDAI score cannot be calculated (ie, <4 components available), the CDAI score will be considered missing. When at least 4 of the 8 components are available, any missing components will be imputed by carrying forward the last non-missing value for that component.

Alternative approach

For components scored over 7 days (i.e., total number of liquid or very soft stools, abdominal pain/cramps, general well-being), a participant's value for that component will be considered missing if the participant recorded < 4 measurements for that week. If the participant recorded measurements on only 4, 5 or 6 days, the sum of the non-missing daily values will be multiplied, respectively, by a factor of 7/4, 7/5 or 7/6 to obtain a score representative of all 7 days.

At each timepoint, the CDAI score will only be calculated if ≥ 5 of the 8 components are available AND none of the components of the diary data (ie, total number of liquid or very soft stools,

abdominal pain/cramps, and general well-being) are missing. In other words, CDAI will be calculated only if ALL the following conditions are met:

(1) All 3 components of the diary data are available for at least 4 of 7 days.
(2) At least 2 other components are available.
(3) If there are any missing components at a visit, the last non-missing values for the missing items may be carried forward as long as it was collected no earlier than the nominal visit 8 weeks prior to the visit. For example, if a component at the Week 24 visit is missing, the last non-missing value for the missing component may be carried forward as long as it was collected no earlier than the Week 16 visit.

SES-CD score

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

Table 3: Sample score sheet and scoring definitions for the Simple Endoscopic Score for Crohn's Disease (SES-CD)

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

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

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

Calculation of the SES-CD score:

Two approaches to calculating the SES-CD score are provided below. The main approach will be used for the analysis of all endpoints that include the SES-CD; of particular note is that this main

approach will be used for the co-primary endpoint of endoscopic response at Week 12. A supportive analysis of the endpoint of endoscopic response at Week 12 will also be provided using the alternative approach.

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

The total SES-CD score at a visit will be calculated based on all segments scored at the visit. If the total SES-CD score cannot be calculated (ie, none of the segments are scored) at a visit, the total SES-CD score will be considered missing.

Alternative approach (baseline segments matched approach):

The SES-CD score at baseline is calculated as the sum of all segments scored at baseline. To calculate the SES-CD score at a post-baseline visit, the sum of the segments that were present at baseline will be used. For segments that were present at baseline but missing post-baseline, the baseline score for the missing segment(s) will be carried forward, unless all segments that were present at baseline are missing at post-baseline, in which case the post-baseline SES-CD score will be assigned as missing. In the event that a segment is missing at baseline but non-missing post-baseline, the non-missing post-baseline score is not used in the calculation of SES-CD for all post-baseline evaluations.

5.3.2. Estimand

An estimand is a precise definition of the primary targeted treatment effect defined by the following 5 attributes: Study intervention by Week 12, Population, Variable (endpoint), Intercurrent events (ICEs) and corresponding strategies, and Population-level summary.

There are 2 co-primary estimands in this study corresponding to the 2 co-primary endpoints defined in Section 5.3.1.

Primary Trial Objective: to evaluate the clinical and endoscopic efficacy of guselkumab in participants with moderately to severely active Crohn's disease.

Estimand Scientific Question of Interest: What is the proportion of participants achieving clinical remission at Week 12 and the proportion of participants achieving endoscopic response at Week 12, and thus considered to have benefited from guselkumab versus placebo?

5.3.2.1. Co-primary Estimands

Co-Primary Estimand of Clinical Remission at Week 12 (Estimand 1)

Treatment by Week 12:

Experimental: Combined guselkumab CCI induction dose group (ie, both guselkumab groups who received CCI at Weeks 0, 4, and 8)

Control: Placebo SC q4w (Weeks 0, 4, and 8)

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

Variable (Endpoint):

A binary response variable (response/nonresponse) where response is defined as achieving CDAI score <150 at Week 12 without experiencing any of the ICEs in categories 1-3 and 5 as outlined below prior to the Week 12 visit.

Intercurrent events and corresponding strategies:

ICE Identifier	ICE Description	Corresponding Strategy
1	A Crohn's disease-related surgery (with the exception of minor procedures such as drainage of a superficial abscess or seton placement, etc.)	Composite
2	A prohibited change in Crohn's disease medication (See Appendix 2)	Composite
3	Discontinuation of study intervention due to lack of efficacy or an AE of worsening of Crohn's Disease	Composite
4	Discontinuation of study intervention due to COVID-19 related reasons (excluding COVID-19 infection) or regional crisis	Treatment Policy
5	Discontinuation of study intervention due to COVID-19 infection or for reasons other than those specified in ICE categories 3 and 4.	Composite

The composite strategy assesses the treatment effects not only based on the variable measurements, but also based on ICEs. This estimand acknowledges that having an ICE in categories 1-3 and 5 is an unfavorable outcome. For participants experiencing ICE category 4, the treatment policy strategy considers the occurrence of ICE category 4 is irrelevant in defining the treatment effect.

Population-level summary:

Difference in proportions of responders (according to the **Variable** defined above) between the combined guselkumab group and the placebo group.

Co-Primary Estimand of Endoscopic Response at Week 12 (Estimand 2)

Treatment by Week 12:

Experimental: Combined guselkumab CCI induction dose group (ie, both guselkumab groups who received CCI at Weeks 0, 4, and 8)

Control: Placebo SC q4w (Weeks 0, 4, and 8)

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

Variable (Endpoint):

A binary response variable (response/nonresponse) where response is defined as achieving endoscopic response (≥50% improvement from baseline in SES-CD score) at Week 12 without experiencing any of the ICEs in categories 1 to 3 and 5 as outlined above prior to the Week 12 visit.

Intercurrent events and corresponding strategies:

The same ICEs and corresponding strategies that are specified for the co-primary endpoint of clinical remission at Week 12 will be used.

Population-level summary:

Difference in proportions of responders (according to the **Variable** defined above) between the combined guselkumab group and the placebo group.

5.3.2.2. Supplementary Estimands for the Co-Primary Endpoints

In these supplementary estimands, all intercurrent events defined above for the co-primary estimands are addressed by the **composite strategy**. The supplementary estimands for the co-primary endpoints acknowledge that having an intercurrent event is an unfavorable outcome. The components of the supplementary estimands are the same as those for the primary estimands (estimands 1 and 2) with the exception of the Variable, which is described as follows:

Supplementary Estimand of Clinical Remission at Week 12 (Estimand 3)

Variable (Endpoint): A binary response variable (response/nonresponse) where response is defined as achieving CDAI score <150 at Week 12 without experiencing any of the ICEs in categories 1-5 prior to the Week 12 visit.

Supplementary Estimand of Endoscopic Response at Week 12 (Estimand 4)

Variable (Endpoint): A binary response variable (response/nonresponse) where response is defined as achieving endoscopic response (≥50% improvement from baseline in SES-CD score) at Week 12 without experiencing any of the ICEs in categories 1-5 prior to the Week 12 visit.

5.3.3. Analysis Methods

5.3.3.1. Estimator (Analysis) of the Co-primary Estimands

The analyses of the co-primary estimands will be based on the FAS, which includes all randomized participants who received at least 1 (partial or complete) dose of study intervention. Participants will be analyzed according to the intervention group to which they were randomized regardless of the study intervention they actually received.

Regarding the ICE strategies, for the treatment policy strategy, the associated ICE event will be ignored, and any data observed after the associated ICE event will be used for the analysis. After accounting for the treatment policy strategy, any missing data will be handled by the missing data rules described below.

If the CDAI score cannot be calculated based on the main approach (Section 5.3.1) at a visit, the CDAI score will be considered missing for that visit. Participants who have a missing CDAI score at Week 12 after accounting for the ICE strategies will be considered not to have achieved clinical remission at Week 12 (i.e., Nonresponder Imputation (NRI)).

For the co-primary endpoint of endoscopic response at Week 12, the SES-CD score will be calculated based on the main approach (Section 5.3.1). If the SES-CD score is missing at Week 12 after accounting for the ICE strategies, the participant will be considered not to have achieved the endpoint of endoscopic response at Week 12 (i.e., NRI).

For testing of the co-primary endpoints, the efficacy of the induction dose of guselkumab versus the placebo group will be compared. The two guselkumab groups that are randomized to receive an identical guselkumab induction dose regimen through Week 12 will be combined for these comparisons.

Summaries of the proportion of participants that achieve each co-primary endpoint and the associated 95% confidence interval by treatment group, along with the adjusted treatment difference between guselkumab group and the placebo group, as well as the associated 95% confidence interval, will be presented for each co-primary endpoint. Guselkumab versus placebo will be compared for the co-primary endpoints using a 2-sided Mantel-Haenszel test (Common Risk Difference Test using Mantel-Haenszel stratum weights) at a significance level of 0.05. Specifically, the adjusted treatment differences will be in terms of the common risk difference using Mantel-Haenszel stratum weights and the variance will be based on the Sato variance estimator. The stratification variables used are baseline CDAI score (≤300 or >300), baseline SES-CD score (≤12 or >12), and BIO-Failure status (Yes or No).

To control the Type I error rate, a fixed sequence testing approach will be used for the co-primary endpoints. In particular, clinical remission at Week 12 will be tested first, followed by the testing of endoscopic response at Week 12.

5.3.3.2. Estimator (Analysis) for the Supplementary Estimands of the Co-primary Endpoints

The same estimators used for the co-primary estimands will be used for these supplementary estimands 3 and 4, with the exception that the rules for the treatment policy strategy will no longer apply. All ICEs for the supplementary estimands are addressed by the composite strategy. After accounting for the ICE strategies, participants whose responder status is missing (i.e., missing clinical remission/endoscopic response status at Week 12) for a co-primary endpoint will be considered to be a nonresponder for that co-primary endpoint.

5.3.4. Subgroup Analyses

For subgroup analyses, the analysis sets are based on the FAS. Subgroup analyses based on demographics, baseline disease characteristics, and baseline and previous use of medications for Crohn's disease (including BIO-Failure status) will be performed when the number of participants in each subgroup permits (See Section 5.7.7 for definition of subgroups).

For each of the subgroups, the rate (risk) difference of induction dose of guselkumab vs. placebo and the associated p-values and 95% confidence intervals will be provided using forest plots. The difference in proportions, p-values and confidence intervals will be provided based on the same model specified in Section 5.3.3.1. The subgroup analysis on BIO-Failure status, baseline CDAI score (≤300 or >300), or baseline SES-CD score (≤12 or >12) will not include the corresponding stratification variable.

Using the co-primary estimands, an analysis of the co-primary endpoints will be performed for each region, country and investigator site. This analysis will be descriptive and statistical testing will not be applied.

5.3.5. Sensitivity Analyses

The following sensitivity analyses will be based on varying assumptions of the missing data.

5.3.5.1. Tipping Point Analysis Based on Multiple Imputation with Bernoulli Draws

A sensitivity analysis will be performed using a tipping point analysis with Bernoulli draws to impute missing clinical remission status at Week 12 and missing endoscopic response status at Week 12 after the intercurrent event rules have been applied, when the number of missing participants (after accounting for the ICE strategies) is >5% in any treatment group. This tipping point analysis involves the following distinct steps:

1. Some p will be assumed for each treatment group's response rate, which could vary by treatment group. The response status (Yes/No) for participants with a missing response will be imputed based on a Bernoulli distribution of p. This will be repeated n times (e.g., 200 times) to generate n multiple imputations.

2. Guselkumab versus placebo will be analyzed for the co-primary endpoints in terms of the common (overall) risk difference using Mantel-Haenszel stratum weights, and the variance will be based on the Sato variance estimator. The associated 2-sided Mantel Haenszel test (Common Risk Difference Test) at a significance level of 0.05 will be used to compare guselkumab group to placebo while adjusting for the stratification factors.

3. The results from the imputed data sets will then be combined to produce inferential results based on Rubin's rule.

The analysis will be repeated for a range of values for p (for example, 0% to 100% in increments of 10%, for the placebo and the guselkumab group independently). This tipping point analysis will allow for assumptions about the response rates in the two arms to vary independently; furthermore, it will include scenarios where imputed missing values on guselkumab have worse outcomes than missing values on placebo.

5.3.5.2. Missing at Random (MAR) MI (Multiple Imputation)

This sensitivity analysis will evaluate the co-primary endpoints (estimands 1 and 2) when all missing data (after accounting for ICE strategies) is imputed by MAR MI rather than NRI. Data that is missing, after accounting for the ICE strategies, include data after discontinuation of study intervention for participants with ICE 4 who do not have observed data at Week 12.

Multiple imputation (MI) as described below will be used for imputing missing data (after accounting for the ICE strategies) under the assumption that the data are missing at random (MAR), and binary responses will be obtained from the imputed scores of CDAI and SES-CD.

Multiple Imputation Methods for CDAI and SES-CD at Weeks 12

Variable	MI specifications	
CDAI at Week 12	Multiple imputation with monotone regression of total scores	MIdataset1 (N=200, Seed=4362478) • Imputation variables: CDAI at Week 12 • Ancillary variables: Treatment group, Baseline CDAI, randomization stratification factors (except categorical Baseline CDAI)
SES-CD at Week 12	Multiple imputation with monotone regression of total scores	MIdataset1 (N=200, Seed=4362478) • Imputation variables: SES-CD Total Score at Week 12 • Ancillary variables: Treatment group, Baseline SES-CD, randomization stratification factors (except categorical Baseline SES-CD)

Repeat it 200 times, which results in 200 data sets. Each of the 200 resulting data sets will be analyzed in terms of the common (overall) risk difference using Mantel-Haenszel stratum weights and the Sato variance estimator. The associated 2-sided Mantel Haenszel test (Common Risk Difference Test) at a significance level of 0.05 will be used to compare guselkumab group to placebo while adjusting for the stratification factors: baseline CDAI score (≤300 or >300), baseline SES-CD score (≤12 or >12), BIO-Failure status (Yes or No). The results from the 200 data sets will be combined to produce inferential results based on Rubin's rule.

5.3.6. Supportive Analyses

5.3.6.1. Co-primary endpoint clinical remission at Week 12 based on alternative CDAI calculation approach

A supportive analysis of the endpoint of clinical remission at Week 12 will be performed, and it will use the same attributes as those for the co-primary estimand 1 defined in Section 5.3.2.1 However, the calculation of the CDAI will be based on the alternative approach specified in Section 5.3.1. If the CDAI is missing at Week 12, after accounting for the ICEs, nonresponder imputation will be used.

The co-primary endpoint of clinical remission at Week 12 will be compared between guselkumab induction dose group and placebo using the Estimand 1 and its associated estimator. The estimand will be estimated by the adjusted treatment differences between guselkumab and placebo in the

percentage of participants who achieve clinical remission at Week 12 and the associated 95% CI. Guselkumab versus placebo will be compared for the co-primary endpoint using a 2-sided Mantel-Haenszel test (Common Risk Difference Test) at a significance level of 0.05. Specifically, the adjusted treatment differences will be in terms of the common risk difference using Mantel-Haenszel stratum weights and the variance will be based on the Sato variance estimator. The stratification variables used are baseline CDAI score (≤300 or >300), baseline SES-CD score (≤12 or >12), and BIO-Failure status (Yes or No).

5.3.6.2. Co-primary endpoint endoscopic response at Week 12 based on baseline segments matched approach

A supportive analysis of the endpoint of endoscopic response at Week 12 will be performed, and it will use the same attributes as those for the co-primary estimand 2 defined in Section 5.3.2.1 However, the calculation of the SES-CD will be based on the alternative approach specified in Section 5.3.1 (ie, the baseline segments matched approach). If the SES-CD is missing at Week 12, after accounting for the ICEs, nonresponder imputation will be used.

The co-primary endpoint of endoscopic response at Week 12 will be compared between guselkumab induction dose group and placebo using the Estimand 2 and its associated estimator. The estimand will be estimated by the adjusted treatment differences between guselkumab and placebo in the percentage of participants who achieve endoscopic response at Week 12 and the associated 95% CI. Guselkumab versus placebo will be compared for the co-primary endpoint using a 2-sided Mantel Haenszel test (Common Risk Difference Test) at a significance level of 0.05. Specifically, the adjusted treatment differences will be in terms of the common risk difference using Mantel-Haenszel stratum weights and the variance will be based on the Sato variance estimator. The stratification variables used are baseline CDAI score (≤300 or >300), baseline SES-CD score (≤12 or >12), and BIO-Failure status (Yes or No).

5.4. Secondary Endpoints Analysis

5.4.1. Definition of Endpoints

The secondary endpoints are listed below.

- Clinical remission (CDAI score <150) at Week 24
- PRO-2 remission at Week 12 (defined as an AP mean daily score at or below 1 and SF mean daily score at or below 3, ie, AP ≤1 and SF ≤3, and no worsening of AP or SF from baseline)
- Clinical response (decrease from baseline in CDAI ≥100 points or clinical remission) at Week 12

5.4.2. Estimands

5.4.2.1. Estimands for the Secondary Endpoints

Secondary Estimand of Clinical Remission at Week 24 (Main Estimand 5)

The following describes the attributes of the Main Estimand 5.

Treatment by Week 24:

Experimental:

- Guselkumab CCI q4w (Weeks 0, 4, and 8) followed by guselkumab CCI q4w (Weeks 12, 16, and 20)
- Guselkumab CCI q4w (Weeks 0, 4, and 8) followed by guselkumab CCI q8w starting at Week 16

Control: Placebo SC q4w (Weeks 0, 4, 8, 12, 16, and 20)

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

Variable (Endpoint):

A binary response variable (response/nonresponse) where response is defined as achieving CDAI score <150 at Week 24 without experiencing any of the ICEs in categories 1, 2, 3, 5 and 6 as outlined below prior to the Week 24 visit.

Intercurrent events and corresponding strategies:

ICE Identifier	ICE Description	Corresponding Strategy
1	A Crohn's disease-related surgery (with the exception of minor procedures such as drainage of a superficial abscess or seton placement, etc.)	Composite
2	A prohibited change in Crohn's disease medication (See Appendix 2)	Composite
3	Discontinuation of study intervention due to lack of efficacy or an AE of worsening of Crohn's Disease	Composite
4	Discontinuation of study intervention due to COVID-19 related reasons (excluding COVID-19 infection) or regional crisis	Treatment Policy
5	Discontinuation of study intervention due to COVID-19 infection or for reasons other than those specified in ICE categories 3 and 4	Composite
6	Meets rescue criteria as determined by the IWRS	Composite

The composite strategy assesses the treatment effects not only based on the variable measurements, but also based on ICEs. This estimand acknowledges that having an ICE in categories 1-3, 5 and 6 is an unfavorable outcome. For participants experiencing ICE category 4, the treatment policy strategy considers the occurrence of ICE category 4 is irrelevant in defining the treatment effect.

Population-level summary:

Difference in proportions of responders (according to the **Variable** defined above) between each guselkumab group and the placebo group.

Secondary Estimand of PRO-2 Remission at Week 12 (Main Estimand 6)

The same estimand (except for the "**Variable**" attribute) that is specified for the co-primary endpoint analyses will be used.

Variable (Endpoint):

A binary response variable (response/nonresponse) where response is defined as achieving PRO-2 remission (defined as AP ≤1 and SF ≤3, and no worsening of AP or SF from baseline) at Week 12 without experiencing any of the ICEs in categories 1 to 3 and 5 as outlined above, prior to the Week 12 visit.

Secondary Estimand of Clinical Response at Week 12 (Main Estimand 7)

The same estimand (except for the "**Variable**" attribute) that is specified for the co-primary endpoint analyses will be used.

Variable (Endpoint):

A binary response variable (response/nonresponse) where response is defined as achieving clinical response (decrease in CDAI score ≥100 points or clinical remission) at Week 12 without experiencing any of the ICEs in categories 1 to 3 and 5 as outlined above, prior to the Week 12 visit.

5.4.2.2. Supplementary Estimands for the Secondary Endpoints

In these supplementary estimands, ALL intercurrent events defined above for the main estimands of the secondary endpoints are addressed by the **composite strategy**. The supplementary estimands for the secondary endpoints acknowledge that having an intercurrent event is an unfavorable outcome. The attributes of the supplementary estimands for the secondary endpoints are the same as those for the main estimands of the secondary endpoints with the exception of the **Variable**, which is described as follows:

Supplementary Estimand of Clinical Remission at Week 24 (Estimand 8)

Variable (Endpoint): A binary response variable (response/nonresponse) where response is defined as achieving CDAI score <150 at Week 24 without experiencing any of the ICEs in categories 1-6 prior to the Week 24 visit.

Supplementary Estimand of PRO-2 Remission at Week 12 (Estimand 9)

Variable (Endpoint): A binary response variable (response/nonresponse) where response is defined as achieving PRO-2 remission (defined as AP ≤1 and SF ≤3, and no worsening of AP or SF from baseline) at Week 12 without experiencing any of the ICEs in categories 1-5 as outlined above, prior to the Week 12 visit.

Supplementary Estimand of Clinical Response at Week 12 (Estimand 10)

Variable (Endpoint): A binary response variable (response/nonresponse) where response is defined as achieving clinical response (decrease in CDAI score ≥100 points or clinical remission) at Week 12 without experiencing any of the ICEs in categories 1-5 as outlined above, prior to the Week 12 visit.

5.4.3. Analysis Methods

5.4.3.1. Estimator (Analysis) of the Secondary Endpoints

The analyses of the secondary endpoints will be based on the FAS, which includes all randomized participants who received at least 1 (partial or complete) dose of study intervention. Participants will be analyzed according to the intervention group to which they were randomized regardless of the study intervention they actually received.

Regarding the ICE strategies, for the treatment policy strategy, the associated ICE event will be ignored, and any data observed after the associated ICE event will be used for the analysis. After accounting for the treatment policy strategy, any missing data will be handled by the missing data rules described below.

For CDAI-related endpoints, if the CDAI score cannot be calculated at a visit, the CDAI score will be considered missing for that visit. Participants who have a missing CDAI score at a visit after accounting for the ICE strategies will be considered not to be achieving the endpoint involving that visit as measured by the CDAI score (ie, Nonresponder Imputation (NRI)).

Participants who have missing AP or SF scores at a visit after accounting for the ICE strategies will be considered not to be in PRO-2 remission for that visit.

Summaries of the proportion of participants that achieve each secondary endpoint and the associated 95% confidence interval by treatment group, along with the adjusted treatment difference between guselkumab group and the placebo group, as well as the associated 95% confidence interval, will be presented for each secondary endpoint. Guselkumab dose group (s) versus placebo will be compared for the secondary endpoints using a 2-sided Mantel Haenszel test (Common Risk Difference Test) at a significance level of 0.05. Specifically, the adjusted treatment differences will be in terms of the common risk difference using Mantel-Haenszel stratum weights and the variance will be based on the Sato variance estimator. The stratification variables used are baseline CDAI score (≤300 or >300), baseline SES-CD score (≤12 or >12), and BIO-Failure status (Yes or No).

5.4.3.2. Testing procedure

The Type I error rate will be controlled at the 0.05 (2-sided) significance level over the co-primary endpoints, secondary endpoints, and two Week 48 exploratory endpoints in the multiplicity-controlled testing procedure.

The testing procedure begins with a fixed-sequence testing procedure which will be used to control the overall Type 1 error rate at the 0.05 level across the co-primary and secondary endpoints listed below. The endpoints will be tested sequentially by the following prespecified order:

1. Clinical remission at Week 12 (co-primary)
2. Endoscopic response at Week 12 (co-primary)
3. Clinical remission at Week 24 compared between the high dose group (CCI [redacted] q4w) and the placebo group (secondary)
4. Clinical remission at Week 24 compared between the low dose group CCI [redacted] q8w) and the placebo group (secondary)
5. PRO-2 remission at Week 12 (secondary)
6. Clinical response at Week 12 (secondary)

If any test in the sequence, including the tests of the co-primary endpoints, does not achieve significance at the 2-sided 0.05 level, the p-values for all of the subsequent endpoints will be considered nominal.

Two exploratory endpoints (clinical remission at Week 48, endoscopic response at Week 48) will be tested at the end of the fixed-sequence testing procedure above; the part of the testing procedure for these endpoints is described in Section 5.5.3. Other exploratory endpoints will not be multiplicity-controlled, and nominal p-values will be presented.

5.4.3.3. Estimator for the Supplementary Estimands of the Secondary Endpoints

The same estimators used for the secondary estimands will be used for the supplementary estimands, with the exception that the rules for treatment policy strategy no longer apply as all ICEs are addressed by the composite strategy. After accounting for the ICE strategies, participants whose responder status is missing for a secondary endpoint will be considered to be a nonresponder for that endpoint.

5.4.4. Subgroup Analyses

The consistency of treatment effect for the secondary endpoints will be evaluated for the subgroups of BIO-Failure, CON-Failure and BIO-Naïve using the secondary estimands defined in Section 5.4.2.1. The secondary estimator will be applied to each subgroup.

5.4.5. Supportive analyses

5.4.5.1. Secondary endpoint clinical remission at Week 24 based on alternative CDAI calculation approach

A supportive analysis of the endpoint of clinical remission at Week 24 will be performed, and it will use the same attributes as those for the secondary estimand 5 defined in Section 5.4.2.1 However, the calculation of the CDAI will be based on the alternative approach specified in Section 5.3.1. If the CDAI is missing at Week 24, after accounting for the ICE strategies, nonresponder imputation will be used.

5.4.5.2. Secondary endpoint clinical response at Week 12 based on alternative CDAI calculation approach

A supportive analysis of the endpoint of clinical response at Week 12 will be performed, and it will use the same attributes as those for the secondary estimand 7 defined in Section 5.4.2.1 However, the calculation of the CDAI will be based on the alternative approach specified in Section 5.3.1. If the CDAI is missing at Week 12, after accounting for the ICE strategies, nonresponder imputation will be used.

5.5. Exploratory Endpoints Analysis

Two exploratory endpoints (clinical remission at Week 48, endoscopic response at Week 48) will be controlled for multiplicity. These two endpoints will be tested after the testing of the secondary endpoints. See the description of the testing procedure later in Section 5.5.3. The testing of other exploratory endpoints will not be controlled for multiplicity, and nominal p-values will be provided.

5.5.1. Definition of Endpoints and Description of Scoring Systems

All definitions and explanation of scoring systems related to the exploratory endpoints are provided below.

5.5.1.1. Clinical

Abdominal Pain Numerical Rating Scale

The Numerical Rating Scale (NRS) for pain is a unidimensional measure of pain intensity in adults. An 11-point (0-10) NRS is used to evaluate abdominal pain. The score of 0 represents “no pain” and the score of 10 represents “pain as bad as you can imagine”, with greater scores indicating greater pain severity and intensity. Participants will select only one number that best reflects their pain at its worst in the past 24 hours. The abdominal pain NRS will be assessed daily for 7 days prior to a scheduled assessment visit (and every day for 14 days prior to final efficacy & safety follow-up). The average over the 7 days prior to a visit will be used to determine the NRS score at that visit. Daily average abdominal pain NRS at a scheduled visit will not be calculated if total days of assessment is less than 4.

AP Daily Score

Average daily abdominal pain score based on the CDAI is defined as the sum of abdominal pain/cramps ratings in the previous 7 days in a diary card divided by the total number of days assessments were performed. Average daily abdominal pain score at a scheduled visit will not be calculated if total days of assessment is less than 5.

SF Daily Score

Average daily stool frequency score based on the CDAI is defined as the sum of number of liquid or very soft stools in the previous 7 days in a diary card divided by the total number of days assessments were performed. Average daily SF score at a scheduled visit will not be calculated if total days of assessment is less than 5.

Bristol Stool Form Scale

The Bristol Stool Form Scale (BSFS) is a medical aid to classify the form (or consistency) of human feces into 7 categories as illustrated in Table 4. It has been used as a research tool to evaluate the effectiveness of treatments for various diseases of the bowel (eg, irritable bowel syndrome). Types 1 and 2 indicate constipation, with 3 and 4 being the ideal stools as they are easy to defecate while not containing excess liquid, 5 indicates lack of dietary fiber, and 6 and 7 indicate diarrhoea.

Table 4: Bristol Stool Form Scale

Type 1: Separate hard lumps, like nuts (difficult to pass)
Type 2: Sausage-shaped, but lumpy
Type 3: Like a sausage but with cracks on its surface
Type 4: Like a sausage or snake, smooth and soft (average stool)
Type 5: Soft blobs with clear cut edges
Type 6: Fluffy pieces with ragged edges, a mushy stool (diarrhoea)
Type 7: Watery, no solid pieces, entirely liquid (diarrhoea)

Average number of BSFS types 6 and 7 stools per day is defined as the sum of number of BSFS types 6 and 7 stools in previous 7 days in a dairy card divided by the total number of days assessments were performed. Similarly, **average number of BSFS types 5, 6 and 7 stools per day** is defined as the sum of number of BSFS types 5, 6 and 7 stools in previous 7 days in a dairy card divided by the total number of days assessments were performed.

Average number of BSFS stools per day at a scheduled visit will not be calculated if total days of assessment is less than 5 within the previous 7 days prior to a scheduled visit.

PRO-2 Remission

PRO-2 remission is defined as AP mean daily score at or below 1 (AP $\leq$1) AND an SF mean daily score at or below 3 (SF $\leq$3), and no worsening of AP or SF from baseline.

Corticosteroid-free Clinical Remission

Corticosteroid-free clinical remission is defined as CDAI score <150 and not receiving corticosteroids (at a specified visit).

90-day Corticosteroid-free Clinical Remission

90-day corticosteroid-free clinical remission is defined as achieving clinical remission and not receiving corticosteroids for 90 days prior to the specified visit e.g. 90-day corticosteroid-free clinical remission at Week 48 is defined as CDAI score <150 at Week 48 and not receiving corticosteroids for 90 days prior to Week 48.

5.5.1.2. Endoscopic

Endoscopic Remission

The definition of endoscopic remission is SES-CD ≤4 and at least a 2-point reduction from baseline and no subscore greater than 1 in any individual component.

NOTE: All 3 of the following conditions must be met to be in endoscopic remission:

- Total SES-CD score is ≤4
- None of the 20 individual cells making up the Total SES-CD can be greater than 1
- Change from Baseline in Total SES-CD score is at least a 2-point decrease

Endoscopic Remission (Alternative Definition)

The *alternative* definition of endoscopic remission is an SES-CD Score ≤2.

Endoscopic Healing

Endoscopic healing is defined as the complete absence of mucosal ulcerations in any ileocolonic segment.

Endoscopic Response (alternative definition)

The alternative definition of endoscopic response is >50% improvement from baseline in SES-CD or SES-CD score ≤2.

Deep Remission

Deep remission is defined as achieving both clinical remission and endoscopic remission.

Deep Remission (Alternative Definition)

Deep remission (Alternative definition) is defined as achieving both clinical remission and endoscopic remission (Alternative definition).

5.5.1.3. Fistula

Fistula Response

Fistula Response is defined as ≥ 50% reduction from baseline in number of open or draining fistulas, among participants with 1 or more fistulas at baseline.

Complete Fistula Response

Complete Fistula Response is defined as 0 open or draining fistulas, among participants with 1 or more fistulas at baseline.

5.5.1.4. Health Related Quality of Life (HRQOL) Measures

IBDQ

The Inflammatory Bowel Disease Questionnaire (Irvine et al, 1994) is a 32-item questionnaire specifically designed for participants with IBD. The range of the total IBDQ score is 32 to 224. Higher scores indicate better quality of life. The total IBDQ score has 4 dimension scores (bowel, systemic, social, and emotional). Each of the individual IBDQ dimensions will be calculated when ≤ 1 item is missing in the dimension. The missing item will be estimated using the average value across the nonmissing items. If any one of the dimensions within the total IBDQ score cannot be calculated, then the total IBDQ score cannot be calculated.

PROMIS-29

The PROMIS-29 is a validated general health profile instrument that is not disease-specific. It is a collection of short forms containing 4 items for each of 7 domains (depression, anxiety, physical function, pain interference, fatigue, sleep disturbance, and ability to participate in social roles and activities). PROMIS-29 also includes an overall average pain intensity 0-10 numeric rating scale (NRS). They assess all domains over the past seven days except for Physical Function which has no timeframe specified.

The raw score of each domain is converted into a standardized score with a mean of 50 and a standard deviation (SD) of 10 (T-Score). The standardized T-score is reported as the final score for each participant. Pain Intensity is presented as raw responses (0-10). For PROMIS domains of Depression, Anxiety, Physical Function, Pain Interference, and Fatigue, a score of 50 is the average for the United States general population with a standard deviation of 10, because testing was performed on a large sample of the general population. However, the other two domains (Ability to Participate in Social Roles and Activities and Sleep Disturbance) were not centered in a national sample. For these two domains, a score of 50 represents the average of the calibration sample which was generally more enriched for chronic illness, and a score of 50 likely represents somewhat sicker people than the general population. For symptom-oriented domains of PROMIS-29 (anxiety, depression, fatigue, pain interference, and sleep disturbance), higher scores represent worse symptoms and a T-score of 60 is one SD worse than average. For the function-oriented domains (physical functioning and social role), higher scores represent better functioning and a T-score of 60 is one SD better than average. Additionally**,** the physical component summary score (PCS) and mental component summary score (MCS) will each be derived from all 7 domain scores of PROMIS-29 (reference) as measures for general health related quality (HRQOL). Higher PCS and MCS scores indicate better HRQOL.

5.5.1.5. Histologic assessment

Histologic assessment will be performed using biopsy samples collected during endoscopy. Two biopsy samples from each of the 5 segments will be collected at screening, Week 12, Week 48, and Week 96 from each of 5 predefined anatomic locations: ileum, right colon, transverse colon, left colon, and rectum.

Biopsies will be collected from representative areas that are consistent with the mucosal appearance visually observed during endoscopy. In a segment with ulcers, all biopsies will be taken from the edge of the ulcer. In a segment without ulcers but with macroscopically abnormal areas, biopsies will be taken from the most inflamed area. In a segment that is endoscopically normal, random biopsies will be taken.

Histologic assessment will be conducted by a group of central expert pathologist readers who are blinded to treatment groups, sites and visits. The modified Global Histology Activity Score (GHAS), Geboes grading scale and Robarts Histologic Index (RHI) will be used to evaluate histologic improvement.

Modified Global Histology Activity Score

The Global Histology Activity Score was first described in 1998, and has been subsequently utilized in a number of studies resulting in peer-reviewed publications (D'Haens et al 1998; D'Haens et al 1999; Mojtahed et al 2014; Baert et al 1999; Laharie et al 2011; Mantzaris et al 2009).

All biopsies for each region will be scored in a blinded manner using the GHAS, with minor adaptations for the circumstances of this study:

- At each time point, all biopsies (up to 2 biopsies per region) obtained from each of the predefined anatomic regions will be scored separately for each of the 8 histological features; feature 8 has been modified from the original GHAS:
    1. epithelial damage (scored 0-2)
    2. architectural changes (scored 0-2)
    3. infiltration of mononuclear cells into the lamina propria (scored 0-2)
    4. polymorphonuclear cells in the lamina propria (scored 0-2)
    5. polymorphonuclear cells in the epithelium (scored 0-3)
    6. the presence of erosions/ulcers (1 for presence and 0 for absence)
    7. presence of granulomas (1 for presence and 0 for absence)
    8. percentage of the tissue affected on the slide image(s) (0 for none, 1 for >0 and <33%, 2 for 33-66%, and 3 for >66%); blinded readers will assess the percentage of tissue involved, considering both biopsies for a given segment (note that in the original GHAS, this component was measured as "number of biopsies affected")
- For items 1-7, the single highest scoring feature from each of the biopsies will be used as the score for that feature.
- The sum of the scores from the 8 histologic features within a region (terminal ileum, right colon, transverse colon, left colon, rectum) will be used as the Total score for that region (range 0-16). Furthermore,

Total Ileum:

- The Ileal score will be equal to the Total Ileum Score.
- If any of the items 1-8 is missing, the Total Ileal score is missing. However, if items 1, 4, 5, and 6 are non-missing, then histologic ileal remission (defined in the Table below) can still be determined.

Total Colonic score:

- The worst available scores for each feature across the right colon, transverse colon, left colon, and rectum will be summed to obtain a Total Colonic score.
- If the worst score across the right colon, transverse colon, left colon, and rectum for any of the items 1-8 is missing, then the Total Colonic score is missing. However, if items 1, 4, 5, and 6 are non-missing, then histologic colonic remission (defined in the Table below) can still be determined.

Total Subject score:

- The worst available scores for each feature across the 5 regions will be summed to obtain a Subject Score
- If the worst score across the ileum, right colon, transverse colon, left colon, and rectum for any of the items 1-8 is missing, then the Total Subject score is missing. However,
  - If items 1, 4, 5, or 6 corresponding to the subject are non-missing, then histologic remission can be determined.
  - If items 1, 4, 5, and 6 corresponding to the subject are all non-missing and all equal to 0 (Q1=Q4=Q5=Q6=0), then the participant is considered to be in histologic response even in the presence of missing scores for items 2, 3, 7, and 8; otherwise, histologic response cannot be determined if at least one of items 2, 3, 7 or 8 is missing.

If both biopsies within a region (ileum, right colon, transverse colon, left colon and rectum) are missing or unevaluable, the GHAS score will be missing for that region.

The definitions of baseline disease activity and histologic remission using the GHAS System apply to a segment (region), Ileal, Colonic, and Subject level; the range of possible scores is 0 – 16 irrespective of extent of the area being analyzed. The binary (yes or no) endpoints histologic response, histo-endoscopic response, and histo-endoscopic remission will be evaluated on a subject level only.

Term	Scope	Definition
Baseline Active Disease (GHAS)	Subject-level Ileal Colonic	GHAS with a score > 0 for infiltration of polymorphonuclear cells in the lamina propria, polymorphonuclear cells in epithelium, or presence of erosions and/or ulcers [Q1 > 0 or Q4 >0 or Q5 > 0 or Q6 > 0]
Histologic Response (GHAS)	Subject-level	≥ 50% reduction in subject-level GHAS score from baseline, or absence of mucosal neutrophils (epithelium and lamina propria), epithelial damage, erosions and ulceration [Q1 = 0, Q4 = 0, Q5 = 0 and Q6 = 0]

Histologic Remission (GHAS)	Subject-level Ileal Colonic	Absence of mucosal neutrophils (epithelium and lamina propria), epithelial damage, erosions and ulceration [Q1 = 0, Q4 = 0, Q5 = 0 and Q6 = 0]
Histo-Endoscopic Response (GHAS)	Subject-level	Combination of histologic response and endoscopic response (≥50% improvement from baseline in SES-CD score)
Histo-Endoscopic Remission (GHAS)	Subject-level	Combination of histologic remission and endoscopic remission (SES-CD ≤4 and at least a 2-point reduction from baseline and no subscore greater than 1 in any individual component)
Histo-Endoscopic Remission (GHAS, Alternative Definition)	Subject-level	Combination of histologic remission and endoscopic remission (SES-CD Score ≤2).

Geboes Grading System

Histological disease activity will be scored in a blinded manner by an experienced pathologist using the Geboes Scoring system (GS). The GS is a 7-item scale (with 4 levels of severity for each item) that categorizes inflammation as grade 0 (architectural change only), grade 1 (chronic inflammation), grade 2 (2a, lamina propria eosinophils and 2b, lamina propria neutrophils), grade 3 (neutrophils in the epithelium), grade 4 (crypt destruction), or grade 5 (erosion or ulceration), as illustrated in Table 5.

Table 5: Grading Criteria for the histological evaluation of Disease Activity in CD (Geboes grading system)

Grade 0	Structural (architectural change)
Subgrades	
0.0	No abnormality
0.1	Mild abnormality
0.2	Mild or moderate diffuse or multifocal abnormalities
0.3	Severe diffuse or multifocal abnormalities
Grade 1	Chronic inflammatory infiltrate
Subgrades	
1.0	No increase
1.1	Mild but unequivocal increase
1.2	Moderate increase
1.3	Marked increase
Grade 2	Lamina propria neutrophils and eosinophils
2A Eosinophils	
2A. 0	No increase
2A.1	Mild but unequivocal increase
2A.2	Moderate increase
2A.3	Marked increase
2B Neutrophils	
2B. 0	None
2B.1	Mild but unequivocal increase
2B.2	Moderate increase
2B.3	Marked increase
Grade 3	Neutrophils in epithelium
3.0	None
3.1	< 5% crypts involved
3.2	< 50% crypts involved
3.3	> 50% crypts involved
Grade 4	Crypt destruction
4.0	None
4.1	Probable—local excess of neutrophils in part of crypt
4.2	Probable—marked attenuation
4.3	Unequivocal crypt destruction
Grade 5	Erosion or ulceration
5.0	No erosion, ulceration, or granulation tissue
5.1	Recovering epithelium+adjacent inflammation
5.2	Probable erosion—focally stripped
5.3	Unequivocal erosion
5.4	Ulcer or granulation tissue

- At each time point, all biopsies (up to 2 biopsies per region) obtained from each of the predefined anatomic regions will be scored separately for each of the 7 items:
- The single highest scoring item from each of the biopsy will be used as the score for that item.
- The Ileal score is equal to the highest grade with a score > 0 (eg, a score of 3 for Grade 5 will be a 5.3 regardless of the scores for the other grades). All items within the ileum must be non-missing to obtain an Ileal score.
- The Colonic score is equal to the highest grade with a score > 0 across the right colon, transverse colon, left colon, and rectum.
- The Subject score is equal to the highest grade with a score > 0 across the ileum, right colon, transverse colon, left colon, and rectum.

If both biopsies within a region are missing or unevaluable, the GS will be missing for that region. The definitions of baseline disease activity and histologic remission using the Geboes Scoring

System apply to a segment (region), Ileal, Colon, and Subject level. Histologic response, histo-endoscopic response, and histo-endoscopic remission will be evaluated on a subject level only.

Term	Scope	Definition
Active Disease (GS)	Subject-level Ileal Colonic	>2B.0 [Grade 2B ≠ 0 OR Grade 3 ≠ 0 OR Grade 4 ≠ 0 OR Grade 5 ≠ 0]
Histologic Response (GS)	Subject-level	≤ 3.1, indicating neutrophil infiltration in <5% of crypts, no crypt destruction and no erosions, ulcerations or granulation tissue)
Histologic Remission (GS)	Subject-level Ileal Colonic	≤ **2B.0** (i.e., no increase in neutrophils in lamina propria [GS 2B.0], no neutrophils in epithelium [GS 3.0], no crypt destruction [GS 4.0], and no erosions, ulcerations or granulation tissue [GS 5.0])
Histo-Endoscopic Response (GS)	Subject-level	Histologic response and endoscopic response (≥50% improvement from baseline in SES-CD score)
Histo-Endoscopic Remission (GS)	Subject-level	Histologic remission and endoscopic remission (SES-CD ≤4 and at least a 2-point reduction from baseline and no subscore greater than 1 in any individual component)
Histo-Endoscopic Remission (GS, Alternative Definition)	Subject-level	Histologic remission and endoscopic remission (SES-CD Score ≤2).

Robarts Histologic Index

Histological disease activity will be scored by a blinded experienced pathologist using the Robarts Histopathology Index (RHI). The RHI is a 4-item index (with 4 levels for each item) that evaluates chronic inflammation, lamina propria neutrophils, neutrophils in the epithelium, and erosion or ulceration as illustrated in Table 6. Total score ranges from 0 to 33, where higher scores denote more severe inflammation.

Table 6: Robarts Histologic Index

Component
1. Chronic inflammatory infiltrate
0=No increase
1=Mild but unequivocal increase
2=Moderate increase
3=Marked increase
2. Lamina propria neutrophils
0=None
1=Mild but unequivocal increase
2=Moderate increase
3=Marked increase

3. Neutrophils in epithelium

0=None

1=<5% crypts involved

2=<50% crypts involved

3=>50% crypts involved

4. Erosion or ulceration

0=No erosion, ulceration or granulation tissue

1=Recovering epithelium+adjacent inflammation

1=Probable erosion—focally stripped

2=Unequivocal erosion

3=Ulcer or granulation tissue

RHI = 1 x chronic inflammatory infiltrate level (4 levels)
+ 2 x lamina propria neutrophils (4 levels)
+ 3 x neutrophils in epithelium (4 levels)
+ 5 x erosion or ulceration (4 levels)

That is, RHI incorporates four components from the Geboes score (Grades 1, 2B, 3, and 5) into a weighted sum of the subgrades using the following formula,

RHI = 1*GS1 + 2*GS2B + 3*GS3 + 5*GS5

All components need to be scored with the re-parameterized grade score (Grade 5) to calculate the RHI.

If any of the components needed to calculate the RHI is missing, then the RHI will be set to missing.

All components need to be scored with the re-parameterized grade score (Grade 5) to calculate the RHI.

- At each time point, all biopsies (up to 2 biopsies per region) obtained from each of the predefined anatomic regions will be scored separately for each of the 4 items:
- The single highest scoring item from each of the biopsies will be used as the score for that item.
- The weighted sum of the 4 scores within a region will be used as the Total score for that region (range 0-33). Furthermore,
  - The Ileal score will be equal to the Total Ileum Score
  - The worst scores for each item across the right colon, transverse colon, left colon, and rectum will be summed to obtain a Total Colonic Score.
  - The worst scores for each item across the 5 regions will be summed to obtain a Subject Score.

If both biopsies within a region are missing or unevaluable, the RHI will be missing for that region.

The definitions of baseline disease activity and histologic remission using RHI apply to a segment (region), Ileum, Colon, and Subject level. Histologic response, histo-endoscopic response, and histo-endoscopic remission will be evaluated on a subject level only.

Term	Scope	Definition
Active Disease (RHI)	Subject-level Ileal Colonic	RHI with a score > 0 for any of Items 2-4 of RHI (lamina propria neutrophils, neutrophils in epithelium, or erosions or ulcerations)
Histologic Response (RHI)	Subject-level	≥ 50% reduction in RHI score from baseline or sub-scores of lamina propria neutrophils and neutrophils must be equal to 0, with no ulcers or erosions
Histologic Remission (RHI)	Subject-level Ileal Colonic	sub-scores of lamina propria neutrophils, neutrophils in the epithelium and erosions or ulcerations must be equal to 0.
Histo-Endoscopic Response (RHI)	Subject-level	Histologic response and endoscopic response (≥50% improvement from baseline in SES-CD score)
Histo-Endoscopic Remission (RHI)	Subject-level	Histologic remission and endoscopic remission (SES-CD ≤4 and at least a 2-point reduction from baseline and no subscore greater than 1 in any individual component)
Histo-Endoscopic Remission (RHI Alternative Definition)	Subject-level	Histologic remission and endoscopic remission (SES-CD Score ≤2).

5.5.2. List of Exploratory Endpoints

In addition to the primary and secondary endpoints, exploratory endpoints related to disease status, HRQOL outcomes, inflammatory biomarkers, and health economics will be analyzed.

5.5.2.1. CDAI based endpoints

- Change in CDAI score from baseline over time
- Clinical remission over time
- Clinical remission at both Week 12 and Week 24
- Clinical remission at both Week 12 and Week 48 (Sustained remission)
- Clinical response over time
- Clinical response at both Week 12 and Week 24
- Clinical response at both Week 12 and Week 48 (Sustained response)
- Durable clinical remission at Week 48 (CDAI <150 for ≥80% of all visits between Week 12 and Week 48 [ie, at least 8 of 10 visits]), which must include Week 48)
- Change from baseline in the weighted CDAI component scores over time

5.5.2.2. Corticosteroid related endpoints

- Corticosteroid-free clinical remission at Week 24
- Corticosteroid-free clinical remission at Week 48
- 90-day corticosteroid-free clinical remission at Week 48
- Average daily prednisone-equivalent (P.Eq) corticosteroid dose (excluding budesonide and beclomethasone dipropionate) over time in participants who were receiving corticosteroids other than budesonide and beclomethasone dipropionate at baseline
- Change from baseline in the average daily prednisone equivalent (P.Eq) oral corticosteroids dose (mg/day; excluding budesonide and beclomethasone dipropionate) at each visit over time in participants who were receiving oral corticosteroids other than budesonide and beclomethasone dipropionate at baseline
- Number of participants who were not receiving concomitant corticosteroids at Weeks 24, 48, 96 in participants who were receiving concomitant corticosteroids at baseline
- Number of participants who were not receiving corticosteroids for at least 90 days prior to Weeks 48, 96 in participants who were receiving concomitant corticosteroids at baseline

5.5.2.3. PRO-2 remission

- PRO-2 remission over time
- Corticosteroid-free PRO-2 remission at Week 48 (defined as AP ≤1 and SF ≤3, and no worsening of AP or SF from baseline, and not receiving corticosteroids at Week 48)
- Durable PRO-2 remission at Week 48 (defined as AP ≤1 and SF ≤3, and no worsening of AP or SF from baseline, for ≥80% of all visits between Week 12 and Week 48 [ie, at least 8 of 10 visits], which must include Week 48)

5.5.2.4. Abdominal Pain and Stool Frequency

- AP score (daily average based on the CDAI assessment) ≤1 over time, among participants with daily average AP score >1 at baseline
- Number of liquid or very soft stools (daily average based on the CDAI assessment) ≤3 over time, among participants with daily average number of liquid or very soft stools >3 at baseline
- Absence and/or resolution of abdominal pain (defined as a mean daily CDAI AP score of 0 in the week prior to the visit) among participants with mean AP>1 at baseline, over time
- SF remission (defined as number of liquid or very soft stools daily average based on CDAI assessment SF ≤1.5 in the week prior to the visit) among participants with mean SF >2.5 at baseline, over time

5.5.2.5. Endoscopy

- Change in SES-CD score from baseline at Week 12, Week 48, and Week 96
- Endoscopic response at Week 48 and Week 96
- Endoscopic response (Alternative Definition) at Week 12, Week 48, and Week 96

- Endoscopic remission at Week 12, Week 48, and Week 96
- Endoscopic remission (Alternative Definition) at Week 12, Week 48, and Week 96
- Endoscopic response at both Weeks 12 and 48 (sustained endoscopic response)
- Endoscopic response at Week 12 and endoscopic remission at Week 48
- Endoscopic response at Week 48 and CRP concentration ≤5 mg/L at Week 48 among participants with elevated CRP (>5 mg/L) at baseline
- Endoscopic response at Week 48 and fecal calprotectin concentration ≤250 µg/g at Week 48 among participants with elevated fecal calprotectin >250 µg/g at baseline.
- Clinical remission at Week 12 and endoscopic response at Week 12
- Clinical remission at Week 48 and endoscopic response at Week 48
- Deep remission at Week 48
- Endoscopic healing at Week 12, Week 48, and Week 96

5.5.2.6. Fistula

- Fistula response over time, defined as a ≥50% reduction from baseline in the number of draining fistulas, among participants with 1 or more fistulas at baseline
- Complete fistula response over time, among participants with 1 or more fistulas at baseline

5.5.2.7. Biomarkers

- Clinical response and ≥50% reduction from baseline in CRP or fecal calprotectin at Week 12, Week 24, and Week 48
- Clinical remission and ≥50% reduction from baseline in CRP or fecal calprotectin at Week 12, Week 24, and Week 48
- Change in CRP over time
- CRP concentrations ≤5 mg/L over time among participants with elevated CRP (>5 mg/L) at baseline
- Change in fecal calprotectin from baseline over time
- Fecal calprotectin concentration ≤250 µg/g over time, among participants with fecal calprotectin >250 µg/g at baseline
- Fecal calprotectin concentration ≤100 µg/g over time, among participants with fecal calprotectin >250 µg/g at baseline
- Fecal calprotectin concentration ≤50 µg/g over time, among participants with fecal calprotectin >250 µg/g at baseline
- Clinical remission and CRP concentration ≤5 mg/L at Week 24 and Week 48, among participants with elevated CRP at baseline
- Clinical remission and fecal calprotectin concentration ≤250 µg/g at Week 24 and Week 48, among participants with elevated fecal calprotectin (>250 µg/g) at baseline

- Clinical response and CRP concentration ≤5 mg/L at Week 24 and Week 48, among participants with elevated CRP at baseline
- Clinical response and fecal calprotectin concentration ≤250 µg/g at Week 24 and Week 48, among participants with elevated fecal calprotectin (>250 µg/g) at baseline
- Clinical response and CRP concentration ≤5 mg/L or fecal calprotectin concentration ≤250 µg/g over time, among participants with elevated CRP and fecal calprotectin >250 µg/g at baseline
- Clinical remission and CRP concentration ≤5 mg/L or fecal calprotectin concentration ≤250 µg/g over time, among participants with elevated CRP and fecal calprotectin >250 µg/g at baseline

5.5.2.8. IBDQ and PROMIS-29

- Change from baseline in total IBDQ score and domain scores over time
- IBDQ remission (IBDQ ≥170) over time
- IBDQ response (≥16-point improvement from baseline) over time
- Change from baseline in the 7 PROMIS-29 domains over time
- Change from baseline in the pain NRS score of PROMIS-29 over time
- Responses in the 7 PROMIS-29 domains (improvement of ≥5, ≥7 and ≥9 points from baseline) over time
- Responses in the pain NRS score of PROMIS-29 (improvement of ≥3 points from baseline) over time
- Change from baseline in the PROMIS-29 PCS and MCS over time
- Responses in the PROMIS-29 PCS and MCS (improvement of ≥5, ≥7, and ≥9 points from baseline) over time

5.5.2.9. BSFS and AP-NRS

- Change from baseline in average daily number of BSFS types 6 and 7 stools over time through Week 48.
- A ≥2 reduction in average daily number of BSFS types 6 and 7 stools over time through Week 48.
- Change in abdominal pain NRS (daily average based on a 0-10 NRS) from baseline over time through Week 48.
- A 3-point reduction in abdominal pain NRS (daily average based on a 0-10 NRS) from baseline over time through Week 48, among participants with pain NRS score ≥3 at baseline.

5.5.2.10. Histologic assessments

Histologic endpoints - GHAS

- Change from baseline in the Ileal GHAS score at Week 12, Week 48, and Week 96 (among participants with Ileal histologic disease at baseline per GHAS)
- Change from baseline in the Colonic GHAS score at Week 12, Week 48, and Week 96 (among participants with Colonic histologic disease at baseline per GHAS)

- Change from baseline in the Subject GHAS Score at Week 12, Week 48, and Week 96 (among participants with histologic disease at baseline per GHAS)
- GHAS-determined Histologic response at Week 12, Week 48, and Week 96 (among participants with histologic disease at baseline per Subject GHAS score)
- GHAS-determined Histologic ileal remission at Week 12, Week 48, and Week 96 (among participants with Ileal histologic disease at baseline per Ileal GHAS score)
- GHAS-determined Histologic colonic remission at Week 12, Week 48, and Week 96 (among participants with colonic histologic disease at baseline per Colonic GHAS score)
- GHAS-determined Histologic remission at Week 12, Week 48, and Week 96 (among participants with histologic disease at baseline per Subject GHAS score)
- GHAS-determined Histo-endoscopic response at Week 12, Week 48, and Week 96
- GHAS-determined Histo-endoscopic remission at Week 12, Week 48, and Week 96
- GHAS-determined Histo-endoscopic remission (Alternative definition) at Week 12, Week 48, and Week 96

Histologic endpoints - Geboes

- Geboes-determined Histologic response at Week 12, Week 48, and Week 96 (among participants with histologic disease at baseline per Subject Geboes score)
- Geboes-determined Histologic ileal remission at Week 12, Week 48, and Week 96 (among participants with Ileal histologic disease at baseline per Ileal Geboes score)
- Geboes-determined Histologic colonic remission at Week 12, Week 48, and Week 96 (among participants with colonic histologic disease at baseline per Colonic Geboes score)
- Geboes-determined Histologic remission at Week 12, Week 48, and Week 96 (among participants with histologic disease at baseline per Subject Geboes score)
- Geboes-determined Histo-endoscopic response at Week 12, Week 48, and Week 96
- Geboes-determined Histo-endoscopic remission at Week 12, Week 48, and Week 96
- Geboes-determined Histo-endoscopic remission (Alternative definition) at Week 12, Week 48, and Week 96

Histologic endpoints – RHI

- Change from baseline in the Ileal RHI score at Week 12, Week 48, and Week 96 (among participants with Ileal histologic disease at baseline per RHI)
- Change from baseline in the Colonic RHI score at Week 12, Week 48, and Week 96 (among participants with Colonic histologic disease at baseline per RHI)
- Change from baseline in the Subject RHI Score at Week 12, Week 48, and Week 96 (among participants with histologic disease at baseline per RHI)

- RHI-determined Histologic response at Week 12, Week 48, and Week 96 (among participants with histologic disease at baseline per Subject RHI score)
- RHI-determined Histologic ileal remission at Week 12, Week 48, and Week 96 (among participants with Ileal histologic disease at baseline per Ileal RHI score)
- RHI-determined Histologic colonic remission at Week 12, Week 48, and Week 96 (among participants with colonic histologic disease at baseline per Colonic RHI score)
- RHI-determined Histologic remission at Week 12, Week 48, and Week 96 (among participants with histologic disease at baseline per Subject RHI score)
- RHI-determined Histo-endoscopic response at Week 12, Week 48, and Week 96
- RHI-determined Histo-endoscopic remission at Week 12, Week 48, and Week 96
- RHI-determined Histo-endoscopic remission (Alternative definition) at Week 12, Week 48, and Week 96

5.5.3. Analysis Methods

Through Week 24

Exploratory endpoints through Week 24 (eg, clinical remission, clinical response) will be tested regardless of the significance of the co-primary and secondary endpoints; the testing of these exploratory endpoints through Week 24 will not be controlled for multiplicity. Nominal p-values will be presented.

Descriptive statistics (i.e., N, mean, median, SD, IQ range, minimum, and maximum) will be used to summarize continuous variables. Counts and percentages will be used to summarize categorical variables. Graphical data displays (e.g., line plots) may also be used to summarize the data.

The estimand approach and analysis methods specified in Sections 5.4.2.1 and Section 5.4.3.1 for the secondary endpoints will apply to binary endpoints. The ICEs are the same as those specified in Sections 5.4.2.1

For continuous endpoints, the attributes of the estimands will be the same as those specified for the secondary endpoints relative to **Study Intervention**, **Population**, and **ICEs**. If a participant has an ICE in categories 1-3, 5, or 6, baseline values will be assigned from the point of ICE onward. For participants experiencing ICE 4, the associated ICE event will be ignored and any data observed after the associated ICE event will be used for the analysis.

Continuous efficacy endpoints that are collected at multiple post baseline time points through a given visit will be compared using a Mixed Model Repeated Measures (MMRM). The model will include all available data from all treatment groups through that visit for assessing efficacy.

The MMRM model will be used to test the difference between the combined guselkumab group or each guselkumab group versus the placebo group. This model relies on the MAR assumption for the missing responses. Under the assumption of MAR, the missing data will be accounted for through correlation of repeated measures in the model.

The explanatory variables of the MMRM model will include treatment group, visit, applicable baseline score, BIO-Failure status (yes, no), baseline CDAI (≤ 300 or > 300), baseline SES-CD (≤ 12 or > 12), an interaction term of visit with treatment group, and an interaction term of visit with applicable baseline score. Change from baseline in CDAI endpoints will not include CDAI strata in the model, as the baseline CDAI is already included as a factor.

The estimates of the treatment difference between the combined guselkumab group or each guselkumab group versus the placebo group will be provided by the difference in the least squares means (LSmeans). The 95% confidence interval (CI) for the differences in LSmeans and p-values will be calculated based on the MMRM. An unstructured covariance matrix for repeated measure within a subject will be used. The F-test will use Kenward-Roger's approximation for degrees of freedom. In case of lack of convergence, empirical structured covariances will be used in the following order until convergence is reached: 1) Toeplitz 2) first order Autoregressive Moving Average. If the normality assumption is in question, an appropriate transformation will be used.

Continuous endpoints involving SES-CD or histologic endpoints that are measured only at Week 12 will be compared using an analysis of covariance (ANCOVA) with treatment as a fixed factor, and baseline score, CDAI (≤ 300 or > 300), baseline SES-CD (≤ 12 or > 12), and Bio-Failure status (yes or no) as covariates. Continuous endpoints involving SES-CD will not include SES-CD strata in the model. Multiple imputation (MI) will be used for imputing missing Total SES-CD data under the assumption that the data are missing at random (MAR). For continuous endpoints involving GHAS and RHI, no imputation will be performed for missing GHAS or RHI values; the missing values will remain as missing. For binary endpoints, NRI will be used to impute missing GHAS, Geboes and RHI endpoints.

After Week 24

Two exploratory endpoints (clinical remission at Week 48, endoscopic response at Week 48) will be controlled for multiplicity. The analyses of the two endpoints will be placed after the 6 sequential tests of the co-primary and secondary endpoints in the multiple testing procedure described in Section 5.4.3.1. If all 6 p-values are <0.05, the testing procedure will continue with the 4 tests of these two endpoints (clinical remission at Week 48, endoscopic response at Week 48, for both dose groups) using the Hochberg procedure (Y Hochberg, Biometrika (1988), 75, 4, pp.800-2) with a 2-sided alpha of 0.05. The Type I error will be controlled over the co-primary, secondary, and these two Week 48 exploratory endpoints in this multiplicity-controlled testing procedure.

The testing of other exploratory endpoints will not be controlled for multiplicity, and nominal p-values will be provided.

Similar methodology used for analyzing endpoints through Week 24 will be applied to the endpoints analyses after Week 24. Specifically, the estimand approach and analysis methods specified in Sections 5.4.2.1 and Section 5.4.3.1 for the secondary endpoints will apply to the two exploratory binary endpoints (clinical remission at Week 48, endoscopic response at Week 48). Unless otherwise specified, the ICEs are the same as those specified in Sections 5.4.2.1.

5.6. Safety Analyses

Safety analyses will be based on the Safety Analysis Set, unless otherwise specified. In general, participants will be analyzed according to their assigned treatment. However, participants assigned to placebo who incorrectly received guselkumab at any time will be analyzed in the guselkumab group from the time they received guselkumab.

For all continuous safety variables, descriptive statistics by intervention group will include the N, mean, standard deviation, median, minimum, and maximum. Categorical variables will be summarized by intervention group using frequency counts and percentages. No formal statistical comparisons are planned.

Safety data, including but not limited to, AEs and changes in laboratory assessments, will be summarized. Treatment-emergent AEs will be summarized by intervention group and Medical Dictionary for Regulatory Activities (MedDRA) system organ class and preferred terms.

Safety data will be summarized through Weeks 12, 24, 48, and the end of the study.

- Safety data limited to the induction period (ie, through Week 12) will be summarized by treatment groups (columns):

  1. Placebo
  2. Guselkumab CCI → CCI q8w
  3. Guselkumab CCI → CCI q4w
  4. Guselkumab Combined (columns 2-3)

- Safety data through Weeks 24, 48, and the end of the study will be summarized with the following groups (columns):

  1. Placebo (include all placebo participants excluding data after a participant was rescued with guselkumab)

  2. Placebo → Guselkumab (include placebo participants who were rescued with guselkumab; data in this group occurred after a subject crossed over to guselkumab).

  3. Guselkumab CCI → CCI q8w
  4. Guselkumab CCI → CCI q4w
  5. Guselkumab Combined (columns 3-4)
  6. All Guselkumab (columns 2-4)

5.6.1. Extent of Exposure

The overall treatment exposure (e.g., duration of treatment exposure, number of treatment administrations, and cumulative dose received) and duration of study follow-up time will be descriptively summarized by treatment group defined above based on the safety analysis set.

5.6.2. Adverse Events

The verbatim terms used in the eCRF by investigators to identify AEs will be coded using MedDRA. Treatment-emergent AEs (TEAEs) are AEs with onset date on or after the date of the first administration of study intervention. All reported AEs which are treatment-emergent will be included in the analysis. For each AE, the percentage of participants who experience at least 1 occurrence of the given event will be summarized by treatment group. Since safety should be assessed relative to exposure and follow-up, all AE summary tables will summarize the average weeks of follow-up and average exposure (number of administrations) for each treatment group.

The following analyses of AEs will be used to assess the safety of participants:

- Frequency and type of AEs by system organ class and preferred term
- Frequency and type of SAEs
- Frequency and type of AEs leading to discontinuation of study intervention
- Frequency and type of AEs associated with the Infections and Infestations SOC, including serious infections. An infection is any AE that was recorded based on the MedDRA system organ class “Infections and Infestations”.
- Frequency and type of severe AEs
- Frequency and type of injection-site reactions. A study intervention injection-site reaction is any reaction at a SC study intervention injection site that was recorded as an injection-site reaction by the investigator on the eCRF.
- Frequency and type of related AEs as assessed by the investigator.

*Adverse events may also be summarized as events per 100 subject years of follow-up, which would account for the potential for differences in follow-up times of participants.

Summary tables will provide the count and percentage of participants with 1 or more of the specified TEAEs by treatment group, system-organ class and preferred term.

In addition to the summary tables, listings of participants with treatment-emergent adverse events of special interest, such as, SAEs and TEAEs leading to discontinuation of study intervention will be provided.

AEs of interest, such as, any deaths, possible anaphylactic or serum-sickness like reactions, AEs associated with drug-related hepatic disorders, opportunistic infections, MACE events (cardiovascular death, nonfatal myocardial infarction, nonfatal stroke), tuberculosis, venous thromboembolism, suicidal ideation and behavior, or malignancies, will either be presented in a table, listing, or described in the clinical study report.

5.6.3. Additional Safety Assessments

5.6.3.1. Clinical Laboratory Tests

Blood samples for serum chemistry and hematology will be collected. The following tests will be performed by the central laboratory.

Hematology assessments will include but are not limited to the following: hemoglobin, hematocrit, platelet count, total and differential WBC count.

Blood chemistry assessments will include but are not limited to the following: chemistry panel (total and direct bilirubin, ALT, AST, alkaline phosphatase, albumin, total protein, calcium, phosphate, sodium, potassium, chloride, blood urea nitrogen /urea, and creatinine).

The baseline value for a participant is the value closest to but prior to the first dose of study agent. In addition, change from baseline is defined to be the assessment at the post-baseline visit minus the assessment at baseline. There will be no imputation for missing laboratory values.

Clinical laboratory test values are to be graded based on modified National Cancer Institute Common Terminology Criteria for Adverse Events (NCI-CTCAE) version 5.0 (Appendix 3) except ALT, AST, total bilirubin (TBili), and alkaline phosphatase which will be graded using the predefined upper limit of normal (ULN) thresholds. The laboratory tests not included in the table Laboratory Toxicity Grading in Appendix 3 or the predefined ULN thresholds of liver tests will not be presented in the tables or listings.

The following summaries of clinical laboratory tests will be provided for participants in the Safety Analysis Set:

- Laboratory parameters and change from baseline in laboratory parameters (hematology and chemistry).
- Summary of maximum NCI-CTCAE toxicity grade for post-baseline laboratory values for the predefined hematology and chemistry lab parameters except ALT, AST, total bilirubin, and alkaline phosphatase.
- Shift tables for maximum NCI-CTCAE toxicity grade from baseline will be summarized for the predefined hematology and chemistry lab parameters except ALT, AST, total bilirubin, and alkaline phosphatase.
- Summary of maximum postbaseline measurement for ALT, AST, total bilirubin, and alkaline phosphatase relative to ULN threshold.
- Line graphs will also be provided for ALT, AST, total bilirubin, and alkaline phosphatase.

Listings of participants with the following will also be provided:

- Abnormal post-baseline laboratory values of NCI-CTCAE grade ≥3 except liver tests
- Post-baseline elevated liver tests of ALT or AST ≥ 5x ULN, or total bilirubin ≥2x ULN, or alkaline phosphatase ≥ 2x ULN
- Post-baseline elevated liver test with combined ALT or AST ≥ 3x ULN and (total bilirubin ≥ 2x ULN or INR > 1.5) at the same visit.

5.6.3.2. The Columbia-Suicide Severity Rating Scale (C-SSRS)

The Columbia-Suicide Severity Rating Scale (C-SSRS) will be used as a screening tool to prospectively evaluate suicidal ideation and behavior in this study, as part of a comprehensive evaluation of safety. The C-SSRS is an investigator-administered questionnaire (Mundt et al 2013; Posner et al 2011) that defines five subtypes of suicidal ideation and 4 possible suicidal behaviors, as well as non-suicidal self-injurious behavior and completed suicide.

The baseline is defined as the most severe/maximum score at screening and Week 0. Suicidal ideation and behavior will be analyzed by the most severe/maximum post-baseline C-SSRS outcome of AE of suicidal ideation. Participants with positive (i.e., score >0) post-baseline suicidal ideation and behavior will be presented in a data listing.

5.7. Other Analyses

5.7.1. Pharmacokinetics

The analyses are based on PK Analysis Set.

Descriptive statistics of the serum guselkumab concentrations will be calculated at each sampling time point. These concentrations will be summarized over time for each treatment group.

All concentrations below the lowest quantifiable concentration or missing data will be labeled as such in the concentration database or data listings. Concentrations below the lowest quantifiable concentration will be treated as zero in the summary statistics.

A population PK analysis approach using nonlinear mixed-effects modeling may be used to evaluate guselkumab PK parameters. The influence of important covariates on the population PK parameter estimates may be evaluated. If performed, details will be given in a population PK analysis plan and the results of the population PK analysis will be presented in a separate report.

Participants' data will be excluded from the PK analysis if their data do not allow for accurate assessment of the PK (eg, incomplete administration of the study intervention; missing time of study intervention administration).

5.7.2. Immunogenicity

The analyses are based on Immunogenicity Analysis Set.

The incidence and titers of antibodies to guselkumab as well as injection-site-reactions by antibody to guselkumab will be summarized over time for all participants who receive a dose of guselkumab and have appropriate samples for detection of antibodies to guselkumab (ie, participants with at least 1 sample obtained after their first dose of guselkumab).

A listing of participants who are positive for antibodies to guselkumab will be provided. The maximum titers of antibodies to guselkumab will be provided for participants who are positive for antibodies to guselkumab.

The incidence of neutralizing antibodies to guselkumab will be summarized for participants who are positive for antibodies to guselkumab and have samples evaluable for neutralizing antibodies to guselkumab.

The relationship between antibodies to guselkumab and guselkumab PK may be explored. The relationship between antibodies to guselkumab and efficacy outcomes may be explored. For example, efficacy outcomes (clinical remission, clinical response, PRO-2 remission, endoscopic remission, and endoscopic response) at Weeks 24 and 48 may be explored by antibody to guselkumab status through Week 24 and through Week 48. For all these analyses, having an ICE in categories 1-3, and 5 is an unfavorable outcome. For participants experiencing ICE categories 4 and 6, the treatment policy strategy applies and the occurrence of ICE categories is irrelevant in defining the treatment effect.

5.7.3. Biomarkers

Planned biomarker analyses may be deferred if emerging study data show no likelihood of providing useful scientific information. Any biomarker samples received by the contract vendor or sponsor after the cutoff date will not be analyzed, and therefore, excluded from the biomarker analysis.

Changes in serum protein analytes, whole blood and biopsy tissue RNA obtained over time will be summarized by intervention group where local regulations permit. Associations between baseline levels and changes from baseline in select biomarkers and response to treatment will be explored. RNA analyses will be summarized in a separate technical report.

The biomarker analyses will characterize the effects of guselkumab to identify biomarkers relevant to treatment, and to determine if these biomarkers can predict response to guselkumab. Results of serum, whole blood analyses, and ileocolonic biopsy analyses will be reported in separate technical reports.

5.7.4. Pharmacokinetic/Pharmacodynamic Relationships

The relationship between serum guselkumab concentrations and efficacy measures will be analyzed graphically.

The relationship between guselkumab serum concentrations (quartiles) at Week 12 and change from baseline in CDAI score, CRP, or fecal calprotectin, and the proportion of participants with clinical response, clinical remission, CRP concentration ≤5 mg/L, fecal calprotectin concentration ≤250 µg/g, endoscopic remission, and endoscopic response at Week 12, will be explored. Likewise, the same relationships will be explored at Week 24, Week 48, and Week 96 where applicable. For all these analyses, having an ICE in categories 1-3, and 5 is an unfavorable outcome. For participants experiencing ICE categories 4 and 6, the treatment policy strategy applies and the occurrence of ICE categories is irrelevant in defining the treatment effect.

If feasible, a suitable exposure-response model may be developed to describe the relationship between serum guselkumab exposure and efficacy. Details will be provided in a population PK/PD

analysis plan and results of the population PK/PD analysis will be presented in a separate technical report.

5.7.5. Pharmacogenomic Analyses

Genetic (DNA) analyses will be conducted only in participants who sign the consent form to participate in the pharmacogenomic substudy. These analyses are considered exploratory and will be summarized in a separate technical report.

5.7.6. Health Economics

- Proportion of participants with CD-related surgeries and procedures through Week 12, Week 24, Week 48, and Week 96
- Time to first CD-related surgery through Week 24, Week 48, and Week 96
- Time to first CD-related procedure through Week 24, Week 48, and Week 96

5.7.7. Definition of Subgroups

When the number of participants in each subgroup permits, the consistency of treatment effect for the co-primary endpoints will be evaluated for the subgroups listed below using the co-primary estimands.

The following subgroups will be evaluated for the co-primary endpoints in this study:

1. **Demographics:**
    a. Age (≤ median age, >median age), Age (<65 years old, ≥65 years old)
    b. Sex (male, female)
    c. Race (white, non-white), Race (White, Black or African American, Asian, Other)
    d. Ethnicity (Hispanic or Latino, not Hispanic or Latino)
    e. Weight at baseline (quartiles)
    f. Region (Asia, Eastern Europe, North America, Rest of World)
2. **Baseline disease characteristics:**
    a. Crohn's disease duration (≤ 5 years, > 5 years to ≤15 years, or > 15 years)
    b. Involved gastrointestinal areas based on central reader assessment (ileum only, colon only, ileum & colon)
    c. CDAI score (≤ 300, >300)
    d. CRP (≤ 5 mg/L, > 5 mg/L)
    e. Fecal calprotectin (≤ 250 µg/g, > 250 µg/g)
    f. SES-CD (≤ 12, > 12)
    g. SF ≥ 4 only, AP ≥ 2 only, both SF ≥ 4 and AP ≥ 2

3. Concomitant medication use at baseline:

a. Oral 5-ASA compounds (receiving, not receiving)

b. Oral corticosteroids including budesonide and beclomethasone dipropionate (receiving, not receiving)

c. 6-MP/AZA/MTX (receiving, not receiving)

d. Oral corticosteroids and (6-MP/AZA/MTX) (receiving, not receiving)

e. Oral corticosteroids or (6-MP/AZA/MTX) (receiving, not receiving)

4. CD-related Medication History:

a. Refractory or intolerant to 6-MP/AZA/MTX (yes, no)

b. Refractory, dependent or intolerant to oral or IV corticosteroids (yes, no)

c. Refractory, dependent, or intolerant to oral or IV corticosteroids, but not refractory or intolerant to 6-MP/AZA/MTX (yes, no)

d. Refractory, dependent or intolerant to oral or IV corticosteroids, and refractory or intolerant to 6-MP/AZA/MTX (yes, no)

e. BIO-failure status

- BIO-Failure (Bio-failure status = yes)
- CON-Failure (Bio-failure status = no)
  i. Bio-naïve
  ii. Bio-experienced [but not documented failure]

f. Participants with biologic failure

- Primary nonresponse, secondary nonresponse, or intolerance to
  - At least one anti-TNF (yes, no)
  - Two or more anti-TNFs
  - Anti-TNF only (yes, no)
  - Vedolizumab (yes, no)
  - Vedolizumab and at least one anti-TNF (yes, no)
- For participants with biologic failure to at least one anti-TNF
  - primary nonresponse (yes, no)
  - secondary nonresponse (yes, no)
  - intolerance (yes, no)
- For participants with biologic failure to vedolizumab
  - primary nonresponse (yes, no)
  - secondary nonresponse (yes, no)
  - intolerance (yes, no)

5.8. Interim Analyses

No interim analyses are planned for this study.

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

No DMC for this study.

6. SUPPORTING DOCUMENTATION

6.1. Appendix 1 List of Abbreviations

ADA	anti-drug antibody
AE	adverse event
ALT/SGPT	alanine aminotransferase
ANCOVA	analysis of covariance
AP	abdominal pain
AST/SGOT	aspartate aminotransferase
BSFS	Bristol Stool Form Scale
CDAI	Crohn's Disease Activity Index
CI	confidence interval
CMH	Cochran-Mantel-Haenszel
CRF	case report form
CRP	C-reactive protein
CSR	Clinical Study Report
C-SSRS	Columbia-Suicide Severity Rating Scale
DBL	Database lock
DMC	Data Monitoring Committee
DPS	Data Presentation Specifications
eCRF	electronic case report form
FAS	full analysis set
FDA	Food and Drug Administration
ICE	Intercurrent Event
ICH	International Conference on Harmonization
IQ	interquartile
IBD	Inflammatory Bowel Disease
IBDQ	Inflammatory Bowel Disease Questionnaire
IWRS	interactive web response system
LOCF	last observation carried forward
LTE	long-term extension
MACE	Major Adverse Cardiovascular Events
MedDRA	Medical Dictionary for Regulatory Activities
MMRM	Mixed Model Repeated Measures
NAb	neutralizing antibodies
NCI-CTCAE	National Cancer Institute-Common Terminology Criteria for Adverse Events
NRI	Nonresponder Imputation
NRS	Numerical Rating Scale
PD	pharmacodynamics
PK	pharmacokinetic(s)
PROMIS	Patient-Reported Outcomes Measurement Information System
SAE	serious adverse event
SAP	Statistical Analysis Plan
SD	standard deviation
SES-CD	Simple Endoscopic Score for Crohn's Disease
SF	stool frequency
SoA	Schedule of Activities
TEAE	treatment-emergent adverse event
TF	Treatment Failure

6.2. Appendix 2 Prohibited Changes in CD Medications (Intercurrent Event 2)

(1) Prohibited medications

Initiation of the following prohibited medications after Week 0

a. Immunomodulatory agents other than AZA, 6-MP, or MTX (including, but not limited to, 6-thioguanine, cyclosporine, mycophenolate mofetil, tacrolimus, and sirolimus).
b. Immunomodulatory biologic agents (including, but not limited to, TNFα antagonists, natalizumab, ustekinumab, rituximab, vedolizumab).
c. Experimental Crohn's disease medications (including, but not limited to, upadacitinib, filgotinib, ozanimod, etrolizumab, brazikumab, mirikizumab, risankizumab, and andecaliximab).
d. Thalidomide or related agents.

(2) Corticosteroids

The occurrence of the following changes in corticosteroid usage during induction treatment (ie, before week 12), OR, between week 36 and week 48, including changes initiated before week 36 and continued after week 36, unless otherwise specified.

a. Initiation of oral corticosteroids (including budesonide and beclomethasone dipropionate), parenteral, or rectal corticosteroids due to worsening Crohn's disease.
b. Increase in the dose of oral corticosteroids (including budesonide and beclomethasone dipropionate), as specified below, above the baseline dose due to worsening Crohn's disease.
    i. Oral corticosteroids > 5 mg/day (prednisone equivalent)
    ii. Oral budesonide > 3 mg/day
    iii. Oral beclomethasone dipropionate > 5 mg/day
c. Initiation of oral corticosteroids (including budesonide and beclomethasone dipropionate) due to reasons other than worsening Crohn's disease for more than 7 days during induction treatment, OR, for more than 28 days during maintenance treatment.
d. Increase in the dose of oral corticosteroids (including budesonide and beclomethasone dipropionate), as specified below, above the baseline dose due to reasons other than worsening Crohn's disease for more than 7 days during induction treatment, OR, for more than 28 days during maintenance treatment.
    i. Oral corticosteroids > 5 mg/day (prednisone equivalent)
    ii. Oral budesonide > 3 mg/day
    iii. Oral beclomethasone dipropionate > 5 mg/day

(3) Immunomodulator agents

a. Initiation of oral 6-MP/AZA due to worsening Crohn's disease.
b. Initiation of oral, subcutaneous, or intramuscular MTX due to worsening Crohn's disease.
c. Increase in the dose of oral 6-MP/AZA above the baseline dose due to worsening Crohn's disease.
d. Increase in the dose of oral, subcutaneous, or intramuscular MTX above the baseline dose due to worsening Crohn's disease (within the same route).

(4) 5-ASA
Initiation or increase of oral 5-ASA compounds due to worsening of Crohn's disease.

(5) Antibiotics
Initiation or change of antibiotics due to worsening Crohn's disease.

6.3. Appendix 3 Laboratory Toxicity Grading

Hematology Tests		**Criteria**			
Test	**Direction**	**1**	**2**	**3**	**4**
Hemoglobin (g/dL)	Increase	>0 - 2 g/dL above ULN	>2 - 4 g/dL above ULN	>4 g/dL above ULN	
Hemoglobin (g/dL)	Decrease	<LLN - 10.0	<10.0 - 8.0	<8.0	
Lymphocytes (/mm3)	Increase		>4000 - 20,000	>20,000	
Lymphocytes (/mm3)	Decrease	<LLN - 800	<800 - 500	<500 - 200	<200
Neutrophils (/mm3)	Decrease	<LLN - 1500	<1500 - 1000	<1000 - 500	<500
Platelets (/mm3)	Decrease	<LLN - 75,000	<75,000 - 50,000	<50,000 - 25,000	<25,000
Total WBC count (/mm3)	Increase			>100,000	
Total WBC count (/mm3)	Decrease	<LLN - 3000	<3000 - 2000	<2000 - 1000	<1000
Chemistry Tests		**Criteria**			
Test	**Direction**	**1**	**2**	**3**	**4**
Albumin (g/L)	Decrease	≥30 - <LLN	≥20 - <30	<20	
Corrected Calcium (mmol/L)	Increase	>ULN - ≤2.9	>2.9 - ≤3.1	>3.1 - ≤3.4	>3.4
Corrected Calcium (mmol/L)	Decrease	≥2.0 - <LLN	<2.0 - ≥1.75	<1.75 - ≥1.5	<1.5
Creatinine	Increase	>ULN - ≤1.5 x ULN	>1.5 - 3.0 x baseline; >1.5 - 3.0 x ULN	>3.0 x baseline; >3.0 - 6.0 x ULN	>6.0 x ULN
Glucose (mmol/L)	Decrease	<LLN - 3.0	<3.0 - 2.2	<2.2 - 1.7	<1.7
Potassium (mmol/L)	Increase	>ULN - ≤5.5	>5.5 - 6.0	>6.0 - 7.0	>7.0
Potassium (mmol/L)	Decrease		<LLN - 3.0	<3.0 - 2.5	<2.5
Sodium (mmol/L)	Increase	>ULN - 150	>150 - 155	>155 - 160	>160
Sodium (mmol/L)	Decrease	<LLN - 130		Sodium <130-120	<120

Note: A modified NCI-CTCAE where the toxicity grades are based on the laboratory result and do not take into account the clinical component, if applicable.

Liver Function Tests	**ULN Thresholds**
ALT/AST	> 1 x to < 3 x ULN ≥ 3 x to < 5 x ULN ≥ 5 x ULN to < 8 x ULN ≥ 8 x ULN
Alkaline Phosphatase	> 1 to < 2 x ULN ≥ 2 x to < 4 x ULN ≥ 4 x ULN
Total Bilirubin	> 1 to < 2 x ULN ≥ 2 x ULN

7. REFERENCES

Baert FJ, D'Haens GR, Peeters M, et al. Tumor necrosis factor alpha antibody (infliximab) therapy profoundly down-regulates the inflammation in Crohn's ileocolitis. Gastroenterology. 1999;116(1):22-28.

Daperno M, D'Haens G, Van Assche G, et al. Development and validation of a new, simplified endoscopic activity score for Crohn's disease: the SES-CD. Gastrointest Endosc. 2004; 60(4):505-512.

D'Haens GR, Geboes K, Peeters M, Baert F, Penninckx F, Rutgeerts P. Early lesions of recurrent Crohn's disease caused by infusion of intestinal contents in excluded ileum. Gastroenterology. 1998;114(2):262-267.

D'Haens G, Van Deventer S, Van Hogezand R, et al. Endoscopic and histological healing with infliximab anti-tumor necrosis factor antibodies in Crohn's disease: A European multicenter trial. Gastroenterology. 1999;116(5):1029-1034.

Geboes K, Riddell R, Öst A, et al. A reproducible grading scale for histological assessment of inflammation in ulcerative colitis. Gut. 2000;47:404-409.

Laharie D, Reffet A, Belleannée G, et al. Mucosal healing with methotrexate in Crohn's disease: a prospective comparative study with azathioprine and infliximab. Aliment Pharmacol Ther. 2011;33(6):714-721.

Mantzaris GJ, Christidou A, Sfakianakis M, et al. Azathioprine is superior to budesonide in achieving and maintaining mucosal healing and histologic remission in steroid-dependent Crohn's disease. Inflamm Bowel Dis. 2009;15(3):375-382.

Mojtahed A, Khanna R, Sandborn WJ, et al. Assessment of histologic disease activity in Crohn's disease: a systematic review. Inflamm Bowel Dis. 2014;20(11):2092-2103.

Mundt, JC, Greist, JH, Jefferson, JW, et al. Prediction of suicidal behavior in clinical research by lifetime suicidal ideation and behavior ascertained by the electronic Columbia-Suicide Severity Rating Scale. J Clin Psych. 2013; 74(9):887-893.

Posner, K, Brown, GK, Stanley, B, et al. The Columbia-Suicide Severity Rating Scale: Initial validity and internal consistency findings from three multisite studies with adolescents and adults. Am J Psychiatry. 2011;168(12):1266-1277.

SAS Institute Inc. 2020 SAS/STAT® 15.2 User's Guide, Cary, NC: SAS Institute Inc. The FREQ Procedure: Statistical Computations: Common Risk Difference

Unblinding Plan at Week 24 DBL for Study CNTO1959CRD3004 (GRAVITI)

The purpose of this document is to provide clear instruction on the unblinding and communications of clinical trial data of Week 24 DBL for study CNTO1959CRD3004 (GRAVITI).

The Sponsor will remain blinded to participant-level treatment assignment through Week 48 DBL, with the exception of the Week 24 DBL, when a limited number of Sponsor personnel will be unblinded and have access to the treatment assignment for analysis. These individuals are identified to have access to individual treatment assignment, but it does not necessarily mean they actually accessed and analyzed data in an unblinded fashion. The following measures will be taken to minimize bias and protect the integrity of the clinical program:

Treatment assignment blinding will be maintained for investigative sites, site monitors, and subjects participating in this protocol until after the Week 48 DBL of the study.

The individuals who will be unblinded to participant-level treatment assignment at the Week 24 DBL of study CNTO1959CRD3004 (GRAVITI) will be documented. The Compound Development Team Leader determines who should have access to unblinded participant-level treatment assignment. Approval for access for these individuals will be documented via an amendment to the unblinding plan.

A separate study team will be put in place to manage the conduct of the study after the Week 24 DBL. This study team will remain blinded to treatment assignment until the Week 48 DBL.