Janssen Research & Development

Statistical Analysis Plan
Amendment 1

A Phase 3, 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 Ulcerative Colitis

ASTRO

Protocol CNTO1959UCO3004; Phase III

CNTO1959 (guselkumab)

Status:	Approved
Date:	22 April 2024
Prepared by:	Janssen Research & Development
Document No.:	EDMS-RIM-775632, 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

Changes to Version 1 are mainly based on two rationales:

1. FDA feedback dated 9th June 2023. This concerns specific subgroup analyses and risk difference summaries for AEs leading to discontinuation, SAEs and AESIs;

2. Internal consistency a) to CNTO1959UCO3001 (QUASAR) and b) to internal harmonization of safety outputs.

Table 1: SAP Version History Summary

SAP Version	Approval Date	Change	Rationale
1	28 September 2022	Not Applicable	Initial release
2	22 April 2024	1. Updated meaningful improvement criteria for two UC-PRO/SS scales (Section 5.5.2.3) 2. Clarified definitions for histologic-endoscopic mucosal improvement (alternative definition 1) and deep histologic-endoscopic mucosal remission (Section 5.5.1.2) 3. Added subgroup analyses for ethnicity, racial categories and age (Sections 5.3.3.4, 5.7.7 and Appendix 6.3) 4. Added risk difference summaries for AEs leading to discontinuation, SAEs and AESIs (Section 5.6.2) 5. Updated maximum toxicity grade analyses for laboratory data 6. Added Rescue Analysis Set and Week 12 Clinical Responder Analysis set 7. Added supplemental estimand excluding days of endoscopy and day before endoscopy programmatically from Mayo scoring 8. Clarified that vital signs and physical examination data were not collected on the eCRF and thus will not be summarized	1. Error in Version 1 2. Alignment with QUASAR 3. FDA feedback (9th June 2023) 4. FDA feedback (9th June 2023) 5. Internal consistency 6. Warranted by the number of rescued participants / used for internal comparison with CNTO1959UCO3001 maintenance data. 7. FDA feedback for JNJ78934804UCO2001 8. FDA feedback for CNTO1959UCO3001

1. INTRODUCTION

This Statistical Analysis Plan (SAP) contains definitions of analysis sets, derived variables, and statistical methods for the planned database locks (DBL) for protocol CNTO1959UCO3004. Only the main period through Week 24 and the portion of the extension period through Week 48 are within the scope of this document. Any analyses following the Week 48 DBL will be described in a separate SAP. Changes to the protocol-planned analyses are documented in Appendix 2.

1.1. Objectives

Primary Objective

The primary objective of this study is to evaluate the efficacy, including clinical remission, of guselkumab subcutaneous (SC) induction compared to placebo in participants with moderately to severe active ulcerative colitis (UC).

Secondary Objective

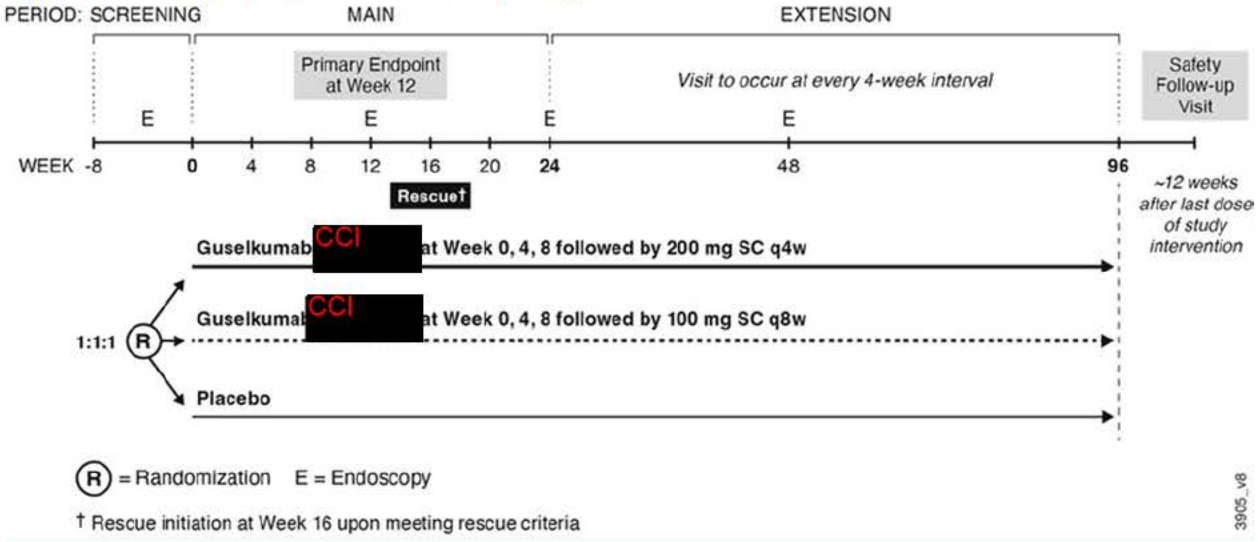
The secondary objectives of this study are

- To further evaluate the efficacy of guselkumab SC induction compared to placebo across a range of outcome measures
- To evaluate the safety of guselkumab SC induction compared to placebo

1.2. Study Design

This is a randomized, double-blind, placebo-controlled, parallel-group, multicenter study to evaluate the efficacy and safety of guselkumab SC induction therapy in adult participants with moderately to severely active UC defined as a modified Mayo score of 5 to 9, inclusive, Mayo rectal bleeding subscore ≥ 1 and a Mayo endoscopy subscore ≥ 2. Participants must have demonstrated an inadequate response to or intolerance of conventional (i.e., 6-mercaptopurine (6-MP), azathioprine (AZA), or corticosteroids) or advanced therapy (ADT; i.e., tumor necrosis factor (TNF) α antagonists, vedolizumab, ozanimod, or approved Janus kinase (JAK) inhibitors). An overview over the study design is given in Figure 1.

Figure 1: Overview of the ASTRO study design



A target of 399 participants will be enrolled in this study with 133 participants planned per intervention group. Participants who had an inadequate response to or intolerance of advanced therapy (ADT-IR) will comprise a minimum of approximately 40% and a maximum of approximately 50% of the population. Consented participants will be screened for study eligibility within 8 weeks of the Week 0 visit. Eligible participants will be randomized in a 1:1:1 ratio to the following intervention groups:

- Guselkumab CCI at Weeks 0, 4, and 8 followed by guselkumab 200 mg SC q4w through Week 24
- Guselkumab CCI at Weeks 0, 4, and 8 followed by guselkumab 100 mg SC q8w through Week 24
- Placebo SC q4w from Week 0 through Week 24

Participants will be allocated to an intervention group using permuted block randomization stratified by ADT-IR status (Yes/No) and Mayo endoscopy subscore at baseline (moderate [2] or severe [3]).

At weeks 12 and 16, all participants will be evaluated according to predefined rescue criteria as given below. All participants in the placebo group who meet rescue criteria at Week 16 will receive rescue treatment, i.e., guselkumab CCI at Weeks 16, 20, and 24 followed by guselkumab 100 mg SC q8w. Participants randomized to guselkumab who meet rescue criteria at Week 16 will continue their assigned treatment regimen and receive blinded sham rescue with matching placebo SC injections at Weeks 16, 20, and 24.

Rescue criteria are defined as:

- No improvement (i.e., no decrease) in Mayo endoscopy subscore at Week 12, when compared with baseline

  AND

- A <2 point improvement (i.e., <2 point decrease) in partial Mayo score at Weeks 12 and 16, when compared with baseline

All participants who reach the Week 24 visit and are benefiting from study intervention in the opinion of the investigator are eligible for the 72-week study extension. At Week 24, participants entering the 72-week extension period will continue the same treatment regimen they were receiving prior to Week 24 (either the treatment regimen assigned at randomization, or the rescue regimen as described above). The study will be unblinded after the last participant completes the Week 48 assessments and the Week 48 DBL and analyses occur. Upon study unblinding, placebo participants who have not been rescued with guselkumab will be discontinued from study intervention and have a safety follow-up visit. All other participants will continue on guselkumab treatment through Week 96. All participants are to complete the safety follow-up visit 12 weeks after the last dose of study intervention. Only the portion of the extension period prior to Week 48 is within the scope of this document.

Participants who are receiving oral 5-ASA compounds, oral corticosteroids, or conventional immunomodulators (AZA, 6-MP, or MTX) for the treatment of UC at baseline should maintain a stable dose through Week 48; with the exception of oral corticosteroids which require mandatory tapering at Week 12. Enrolled participants should not initiate treatment with these medications from Week 0 to Week 48.

Database locks are planned for Week 24, Week 48, and when the last participant completes the last scheduled assessment as shown in the SoA (Protocol Section 1.3). Additional DBLs may be added as necessary.

1.2.1. Safety Assessment Committee

One committee internal to the sponsor, the Safety Assessment Committee (SAC), will have access to unblinded safety data. The SAC is a multidisciplinary group of senior leaders (employees of the study sponsor) who are independent from the clinical study team. SAC members are responsible for reviewing unblinded safety data strictly related to the anticipated events process. The core responsibility of the SAC is to ensure that any anticipated event that surpasses a pre-established frequency threshold, is plausibly clinically related to the investigational product, and meets the requirements for aggregate reporting, is communicated to the competent regulatory authorities. Strict measures are in place to maintain a firewall between the SAC and the clinical teams. The plan for monitoring and analyzing the anticipated events is specified in a separate Anticipated Events Safety Monitoring Plan. The assessment of causality will be made by the sponsor's unblinded SAC.

2. STATISTICAL HYPOTHESES

Primary hypothesis

The primary hypothesis of this study is that guselkumab SC induction is superior to placebo SC in achieving clinical remission at Week 12 among participants with moderately to severely active UC.

Secondary hypotheses

The secondary hypotheses of this study are listed below:

- Guselkumab SC is superior to placebo SC in achieving symptomatic remission at Week 12
- Guselkumab SC is superior to placebo SC in achieving endoscopic improvement at Week 12
- Guselkumab SC is superior to placebo SC in achieving clinical response at Week 12
- Guselkumab SC is superior to placebo SC in achieving clinical remission at Week 24
- Guselkumab SC is superior to placebo SC in achieving symptomatic improvement at Week 24
- Guselkumab SC is superior to placebo SC in achieving endoscopic improvement at Week 24
- Guselkumab SC is superior to placebo SC in achieving clinical response at Week 24
- Guselkumab SC is superior to placebo SC in achieving histologic-endoscopic mucosal improvement at Week 12

Guselkumab groups as described in Section 1.2 will be pooled for testing of all primary and secondary hypotheses at Week 12, as they received the same induction dose regimen up to Week 12. For Week 24 secondary hypotheses, each guselkumab group will be tested separately against placebo.

All Week 48 endpoints are considered exploratory.

3. SAMPLE SIZE DETERMINATION

Sample size was determined by the power to detect significant differences in the primary endpoint of clinical remission at Week 12, and by the objective of maintaining at least 85% power across secondary endpoints at Week 12 between the combined guselkumab SC induction groups and the placebo SC group as well as for symptomatic remission at Week 24 between each guselkumab group and the placebo group, using 2-sided χ^2 tests with significance level 0.05. Combination of guselkumab SC induction groups for the primary endpoint and other Week 12 endpoints is warranted by the fact their treatment is identical through the induction period, and only differs after assessment of the primary endpoint at Week 12.

With assumed clinical remission rates of 8% for placebo and 22% for the combined guselkumab groups (based on data from Phase 2b guselkumab UC IV induction study [CNTO1959UCO3001] and Phase 3 ustekinumab UC program [CNTO1275UCO3001]), a total of 399 participants (randomized 1:1:1 for guselkumab CCI q4w [Weeks 0, 4, and 8] followed by guselkumab 200 mg SC q4w: guselkumab CCI q4w [Weeks 0, 4, and 8] followed by guselkumab 100

mg SC q8w: Placebo, yielding a 2:1 guselkumab:placebo randomization ratio for Week 12 comparisons) will ensure >95% power for the primary endpoint. This sample size also protects against a slightly lower remission rate of 20% for the combined guselkumab groups or a slightly higher remission rate of 9.5% in the placebo group (as observed in CNTO1959UCO3001), yielding a power of 90% in both cases.

Individual power values (ie, without consideration of an adjustment for multiplicity) achieved for secondary endpoints at Week 12 with a total of 399 participants (2:1 guselkumab:placebo) are described in Table 2.

Table 2: Power for Secondary Endpoints at Week 12

Endpoints	Proportion achieving endpoint in Placebo group	Proportion achieving endpoint in combined Guselkumab group	Power
Symptomatic remission	20%	45%	>99%
Endoscopic improvement	13%	28%	94%
Clinical response	30%	60%	>99%
Histologic-endoscopic mucosal improvement	8%	19%	85%
Assumed rates were based on data from the Phase 2b guselkumab UC IV induction study and the Phase 3 ustekinumab UC program			

For the secondary endpoints at Week 24, power is determined by the pairwise comparison of individual guselkumab groups to placebo (see Table 3).

Table 3: Power for Secondary Endpoints at Week 24 with a Total of 399 Participants (133 per Treatment Group)

Endpoints	Proportion achieving endpoint in Placebo group	Proportion achieving endpoint in an individual Guselkumab group	Power
Clinical remission	10%	27%	95%
Symptomatic remission	30%	60%	>99%
Endoscopic improvement	13%	28%	86%
Clinical response	35%	70%	99%

4. POPULATIONS (ANALYSIS SETS) FOR ANALYSIS

The following sets will be used for analyses.

Table 4: Analysis sets

Analysis Sets	Description
Enrolled	All participants who signed the ICF
Randomized	All participants who were randomized in the study
Full Analysis Set (FAS)	All randomized participants who received at least 1 (partial or complete) dose of study intervention
Rescue Analysis Set	All participants in the FAS who met rescue criteria (as determined by the IWRS) and received at least one (partial or complete) dose of study intervention at or after Week 16. This includes both participants randomized to placebo and participants randomized to one of the guselkumab groups who were assigned sham rescue treatment.

Analysis Sets	Description
Week 12 Clinical Responder Analysis Set	All participants in the FAS who met Clinical Response Criteria at Week 12 as defined in Section 5.4.1.1 and received at least 1 (partial or complete) dose of study intervention after Week 12
Safety	All randomized participants who received at least 1 (partial or complete) dose of study intervention
Pharmacokinetics (PK) Analysis Set	All participants who received at least 1 (partial or complete) dose of guselkumab and have at least 1 valid blood sample drawn for PK analysis after their first dose of guselkumab
Immunogenicity Analysis Set	All participants who received at least 1 (partial or complete) dose of guselkumab and have appropriate serum samples for detection of antibodies to guselkumab after their first dose of guselkumab

5. STATISTICAL ANALYSES

5.1. General Considerations

Descriptive statistics (e.g., 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 data. This applies to demographic and baseline characteristics given in Appendix 3.

Analyses suitable for categorical data (e.g., chi-square tests, CMH chi-square tests, or logistic regression, as appropriate) will be used to compare the proportions of participants achieving selected endpoints (e.g., clinical response). In cases of rare events, the Fisher's exact test will be used for treatment comparisons. Continuous response parameters measured at more than one postbaseline visit will be compared using a MMRM model (unless otherwise specified). If the normality assumption is in question, an appropriate transformation may be implemented before fitting the MMRM model. Continuous response parameters measured at only one post-baseline visit will be compared using an ANOVA or ANCOVA, unless otherwise specified. In cases of small sample size, t-test will be used for treatment comparisons.

The overall Type I error rate over primary and secondary endpoints will be controlled at the significance level of 0.05 (2-sided).

Unless otherwise specified, the analysis of endpoints related to the endoscopy subscore, including the modified Mayo score, will be based on the final reported Mayo endoscopy subscore (Section 5.3.1).

Baseline assessment

Baseline assessment is defined as the last available assessment collected prior to or on the day of the first administration of study intervention, unless otherwise specified.

Study day

Study day 1 (*Day 1*) is the date of the first administration of study intervention in the study.

The *study day* for a post-baseline scheduled or unscheduled visit is defined as:

Study day = (date of visit) – (date of first administration of study intervention) + 1.

The study day for a scheduled or unscheduled visit before Baseline is defined as:

Study day = (date of visit) – (date of first administration of study intervention).

There is no "Study Day 0".

Unscheduled visits

All data collected at unscheduled visits will not be used in "by visit" tabulation or graphs for safety analyses but will be included for analyses based on all post-baseline assessments, such as safety narratives and summary of maximum increase/decrease from Baseline for laboratory data. Unscheduled visits for efficacy assessments will be slotted to scheduled visits (see Section 5.1.1). All data collected at scheduled and unscheduled visits will be included in listings.

Combined guselkumab groups

Both guselkumab groups will receive identical study intervention for Week 0, 4, and 8. For analyses of timepoints up to Week 12, these groups will therefore be combined, and the combined group will be referred to as combined guselkumab. Note that for analyses of courses over time which include a period beyond Week 12, the two guselkumab groups may be analyzed separately.

5.1.1. Visit Windows

Unless otherwise specified, actual scheduled visits will be used for over time summaries and listings with no visit windows applied. Early Discontinuation (ED) and unscheduled efficacy visits will be slotted to scheduled visits according to the following mapping rules:

1. Assign a visit number to ED or unscheduled visit based on the visit day according to Table 6.1 and Table 6.2.
2. If the ED or unscheduled visit falls in the window of a scheduled visit for which there is no data recorded, assign the scheduled visit number to the ED or unscheduled visit. If there is data recorded for the scheduled visit, those data will be used in lieu of the ED or unscheduled visit data.

Table 5.1: Visit windows for analysis of Mayo score, endoscopy score and histology assessment

Parameter	Time interval (label on output)	Time interval (Day)*	Target time point (Day)
Partial Mayo Score	Week 4	16-43	29
	Week 8	44-71	57
	Week 12	72-99	85
	Week 16	100-127	113
	Week 20	128-155	141
	Week 24	156-183	169
	Week 28	184-211	197
	Week 32	212-239	225
	Week 36	240-267	253
	Week 40	268-295	281

	Week 44	296-323	309
	Week 48	324-351	337
Endoscopy subscore, full Mayo score including modified Mayo score	Week 12	57-113	85
	Week 24	141-197	169
	Week 48	309-365	337
Histology assessment	Week 12	57-113	85
	Week 48	309-365	337

* Relative to Study Day 1

Table 6.2: Visit windows for analysis of IBDQ, PROMIS-29, UC-PRO/SS, fecal calprotectin, CRP

Parameter	Time interval (label on output)	Time interval (Day)*	Target time point (Day)
IBDQ, PROMIS-29	Week 12	57-113	85
	Week 24	141-197	169
	Week 48	309-365	337
UC-PRO/SS	Week 4	16-43	29
	Week 8	44-71	57
	Week 12	72-99	85
	Week 16	100-127	113
	Week 20	128-155	141
	Week 24	156-183	169
	Week 28	184-211	197
	Week 32	212-239	225
	Week 36	240-267	253
	Week 40	268-295	281
	Week 44	296-323	309
	Week 48	324-351	337
Fecal calprotectin	Week 4	16-56	29
	Week 12	57-113	85
	Week 24	141-197	169
	Week 48	309-365	337
CRP	Week 4	16-43	29
	Week 8	44-71	57
	Week 12	72-99	85
	Week 16	100-127	113
	Week 20	128-155	141
	Week 24	156-196	169
	Week 32	197-253	225
	Week 48	309-365	337

* Relative to Study Day 1

5.2. Participant Dispositions

The number of participants in the following disposition categories through Week 12, Week 24, and Week 48 will be summarized by treatment group and overall based on the **Full Analysis Set, the Rescue Analysis Set and the Week 12 Clinical Responder Analysis Set**:

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

Listings of participants will be provided for the following categories:

- All randomized participants
- Participants who discontinued study intervention
- Participants who terminated study prematurely
- Listing of participants with major protocol deviations as given in Appendix 4.

Listings will identify the visits completed and when the study intervention was discontinued or the study was terminated. Further listings regarding intervention compliance are described in Appendix 7.

In addition, the number and percentage of participants who met rescue criteria (according to the IWRS) will be summarized for the full analysis set.

5.3. Primary Endpoint Analysis

5.3.1. Definition of Endpoint

The primary endpoint is clinical remission at Week 12.

Clinical remission is defined as a stool frequency subscore of 0 or 1 and not increased from baseline, a rectal bleeding subscore of 0, and an endoscopy subscore of 0 or 1 with no friability present on the endoscopy.

The above definition is based on Mayo score components as described below.

CCI [redacted]

CCI [redacted]

[redacted]

- [redacted]
- [redacted]
- [redacted]
- [redacted]

[redacted].

CCI [redacted]

CCI [redacted]

CCI [redacted]

CCI [redacted]

CCI

[redacted]

5.3.2. Estimands

An estimand is a precise definition of the primary targeted intervention 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.

Primary Trial Objective: To evaluate the efficacy of guselkumab SC at Week 12 compared with placebo SC.

Estimand Scientific Question of Interest: What difference in proportions of patients considered to experience benefit from treatment by Week 12 (according to the responder criteria for clinical remission used in the estimand variable definition) can be expected under the assignment to either guselkumab or placebo?

5.3.2.1. Primary Estimand (Estimand 1)

The primary estimand (i.e., a precise definition of the primary targeted treatment effect) is defined by the following 5 attributes:

i) Treatment by Week 12:

- Experimental: Combined guselkumab CCI induction dose group (i.e., both guselkumab groups who received CCI at Weeks 0, 4, and 8; see Section 1.2)
- Control: Placebo SC q4w (Weeks 0, 4, 8)

ii) Population: Participants with moderately to severely active ulcerative colitis as reflected in the inclusion/exclusion criteria.

iii) Variable (Endpoint): A binary response variable (response/nonresponse) where response is defined as achieving a stool frequency subscore of 0 or 1 and not increased from baseline, a rectal bleeding subscore of 0, and an endoscopy subscore of 0 or 1 with no friability present on the endoscopy at Week 12 and none of the ICEs in categories 1 to 4 and 6 outlined in Table 6 has occurred prior to the Week 12 visit.

iv) Intercurrent Events and Corresponding Strategies: ICEs and the corresponding analysis strategies are defined in Table 6.

Table 6: Intercurrent Events and Respective Analysis Strategies

ICE	Analysis strategy for ICE
1. An ostomy or colectomy (partial or total)	**Composite strategy:** Occurrence of these ICEs will be treated as an unfavorable outcome
2. A prohibited change in medications for UC (as defined in Appendix 14)	
3. Discontinuation from study intervention due to lack of efficacy or an adverse event (AE) of UC worsening	
4. Meeting rescue criteria (as determined by the IWRS) (**Note:** Since rescue criteria are assessed at Week 16 for all participants, this ICE is not applicable for the primary endpoint or any of the key secondary Week 12 endpoints)	
5. Discontinuation of study intervention due to COVID-19 related reasons (excluding COVID-19 infection) or regional crisis	**Treatment policy strategy**: Observed values will be used if available
6. Discontinuation of study intervention due to reasons other than ICEs 3 or 5 as described above	**Composite strategy:** Occurrence of these ICEs will be treated as an unfavorable outcome

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

v) Population-level summary: Difference in proportions of participants who achieved the binary response at Week 12 as defined in the variable attribute above between the combined guselkumab group and the placebo group.

5.3.2.2. Supplementary Estimands for the Primary Endpoint

The supplementary estimands are defined to support the primary endpoint.

5.3.2.2.1. Treatment policy for discontinuations (Estimand 2)

The supplementary estimand (Estimand 2) handles all treatment discontinuation ICEs, i.e. ICEs 3, 5 and 6 in Table 6, by the treatment policy strategy. ICEs 1 2 and 4 continue to be handled by the composite strategy. This leads to the following variable definition:

iii) Variable (Endpoint): A binary response variable (response/nonresponse) where response is defined as achieving a stool frequency subscore of 0 or 1 and not increased from baseline, a rectal bleeding subscore of 0, and an endoscopy subscore of 0 or 1 with no friability present on the endoscopy at Week 12 and none of the ICEs in categories 1, 2 and 4 outlined in Table 6 has occurred prior to the Week 12 visit.

5.3.2.2.2. Alternative Mayo Score Calculation (Estimand 3)

This supplementary estimand (Estimand 3) follows the primary estimand in all attributes. The eligible days to be used for the Mayo scoring for this estimand are all available days among the seven days prior to each visit, excluding programmatically the following two days:

1. The day(s) of a procedure (including colonoscopy or sigmoidoscopy)
2. The day preceding such a procedure.

5.3.2.2.3. Participants with Non-missing Stool Frequency subscore at Baseline (Estimand 4)

This estimands shares all properties of the primary estimand as given in Section 5.3.2.1 with the exception of the population, which is defined as follows for this estimand:

ii) Population: Participants with moderately to severely active ulcerative colitis as reflected in the inclusion/exclusion criteria and with a stool frequency subscore recorded at baseline.

5.3.3. Analysis Methods

5.3.3.1. Main Estimator (Analysis) for the Primary Estimand

The analysis of the primary estimand will be based on the FAS, which includes all randomized participants who received at least 1 dose (partial or complete) 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 will be ignored, and any data observed after the associated ICE will be used for analysis. After accounting for the ICE strategies, participants who are missing any or all of the three subscores that comprise the primary endpoint at Week 12 will be considered not to be in clinical remission at Week 12. Participants who are missing the stool frequency subscore at baseline will be considered not to be in clinical remission at Week 12.

Summaries of the proportion of participants in clinical remission at Week 12 by treatment group along with the adjusted treatment difference between the combined guselkumab SC induction group and the placebo group, as well as the associated 95% confidence interval will be presented. Comparison against placebo will be performed in terms of the common (overall) risk difference for multiway 2x2 tables at a two-sided significance level of 0.05. Specifically, the adjusted treatment difference, 95% confidence interval, and p-value will be in terms of the common risk difference by use of Mantel-Haenszel stratum weights and the Sato variance estimator;

stratification is by ADT-IR status (Yes/No), and Mayo endoscopy subscore at baseline (moderate [2] or severe [3]).

5.3.3.2. Multiple Testing Procedure

The testing procedure for across the primary and secondary endpoints is fully described in Section 5.4.2.2. Generally, a fixed sequence approach is applied, where the primary hypothesis is first in sequence and no secondary hypothesis will be tested in a confirmatory manner if the test of the primary hypothesis was not significant.

5.3.3.3. Sensitivity Analyses

Two types of tipping point analyses and an analysis using multiple imputation of missing data under missing at random (MAR) assumption will be performed as sensitivity analyses for the primary estimand using the full analysis set and will be based on varying assumptions for missing data.

5.3.3.3.1. Tipping Point Analysis

Tipping point sensitivity analyses will be conducted to explore the potential impact of missing data and to show the robustness of the conclusion based on the primary analysis. The analysis finds a (tipping) point in the spectrum of assumptions, at which point the result changes from statistically significant to statistically non-significant.

The following method will be utilized to vary the imputation of endpoint status for missing data for the primary endpoint, after the intercurrent event strategies have been applied. For participants with missing data for the primary endpoint, the clinical remission status (responder/non-responder) will be imputed in an increasing manner by participant level for both the combined guselkumab SC induction group and the placebo group. Specifically, for each participant with missing data, a responder/non-responder status will be imputed starting with the scenario where all participants are non-responders (i.e., not in clinical remission) up to the scenario where all participants are responders (i.e., in clinical remission). This would include all possible scenarios of responder status for all missing data allowing assumptions about the missing outcomes to vary independently, including scenarios where participants randomized to guselkumab SC have worse outcomes than participants randomized to placebo SC. For each scenario, the difference in the proportion of participants achieving the primary endpoint between the combined guselkumab SC induction group and the placebo group and the corresponding p-values based on the Chi-square test will be provided.

5.3.3.3.2. Tipping Point 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 after the intercurrent event rules have been applied, when the number of missing participants (after accounting for the ICE strategies) is ≥10 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, to impute the response status (Yes/No) for participants with a missing

response based on a Bernoulli distribution. This will be repeated 500 times to generate 500 multiple imputations.

2. Each of the resulting data sets will be analyzed based on the common risk difference test proposed for the primary analysis.
3. The results from the imputed data sets will then be combined to produce inferential results based on Rubin's rules.

The analysis will be repeated for a range of values for *p* (for example, 0% to 30% in increments of 3% and 30% to 50% in increments of 5% independently, for both the placebo and the guselkumab groups).

5.3.3.3.3. Missing at Random Multiple Imputation

Missing data for the primary estimand of the primary endpoint of clinical remission at Week 12 defined in Section 5.3.1 after the intercurrent event strategies have been applied will be summarized by intervention group. Data that is missing consists of data after discontinuation of study intervention for participants with ICEs in category 5 who do not have observed data at Week 12 and of missing data not related to ICEs.

This sensitivity analysis will evaluate the primary endpoint when all missing data as defined above are multiply imputed under the missing at random assumption. Imputation will be performed separately for all components of the primary endpoint. The following steps will be performed:

1. Any missing Mayo components pertaining to the primary endpoint at Week 12 will be imputed 500 times to generate 500 complete data sets using the fully conditional specification method, assuming missing at random. The following variables will be included in the imputation model: Mayo subscores at baseline and Week 12, partial Mayo subscores at Week 4 and Week 8, ADT-IR status and treatment group.
2. Each of the 500 resulting data sets will be analyzed using the common risk difference with Mantel-Haenszel weights, using baseline Mayo endoscopy subscore and ADT-IR status and as factors.
3. The results from the 500 data sets will be combined to produce inferential results based on Rubin's rules[5)].

5.3.3.4. Subgroup Analyses

To evaluate the consistency of the primary analysis, subgroup analyses based on demographics (e.g., age, sex, ethnicity, race, weight, region), baseline UC disease characteristics (e.g., UC disease duration, baseline endoscopy subscore) and baseline concomitant UC medications (e.g., baseline oral corticosteroid use, immunomodulators, and oral 5-ASA compounds), and history of UC-related medications (e.g., ADT-IR status) will be performed if sufficient participant data are available in the subgroup.

The subgroup analyses for the primary endpoint will be performed using the FAS. The following subgroup analyses will be performed:

- The consistency of intervention effect will be evaluated for the subgroups defined in Section 5.7.7, using the primary estimand and handling missing data by the rules specified in Section 5.3.3.1.
- For each of the subgroups identified in Section 5.7.7, the difference of the combined guselkumab SC groups vs. the placebo SC group relative to the primary endpoint at Week 12 will be presented using a forest plot and the associated p-values and 95% confidence intervals will be provided. The main estimator for the primary endpoint will be used. The main estimator for the primary estimand is based on the common (overall) risk difference by use of Mantel-Haenszel stratum weights, and the associated 95% confidence interval will use the Sato variance estimator. A two-sided Mantel-Haenszel test (at a significance level of 0.05) will be used to compute p-values; stratification is by ADT-IR status (Yes/No) and Mayo endoscopy subscore at baseline (moderate [2] or severe [3]) (where appropriate).

Subgroup analyses will only be performed if warranted by the number of participants in each subgroup.

Additionally, an analysis of the primary endpoint will be performed for each region, country and investigator site. This analysis will be descriptive and statistical testing will not be applied.

5.3.3.5. Supplementary Analysis

As a supplementary analysis, the primary analysis (Section 5.3.3.1) will be repeated for the randomized analysis set, which includes all randomized participants.

5.3.3.6. Estimator (Analysis) for the Supplementary Estimands of the Primary Endpoint

The same estimators used for the primary estimand will be used for the supplemental estimands.

5.4. Secondary Endpoints Analysis

5.4.1. Secondary Endpoints

5.4.1.1. Definition of Endpoints

The following are the secondary endpoints:

- Symptomatic remission at Week 12
- Endoscopic improvement at Week 12
- Clinical response at Week 12
- Clinical remission at Week 24
- Symptomatic remission at Week 24
- Endoscopic improvement at Week 24
- Clinical response at Week 24
- Histologic-endoscopic mucosal improvement at Week 12

They are defined as follows:

Symptomatic remission: Stool frequency subscore of 0 or 1 and not increased from baseline, and rectal bleeding subscore of 0.

Endoscopic improvement: Endoscopy subscore of 0 or 1 with no friability present on the endoscopy.

Clinical response: Decrease from baseline in modified Mayo score by ≥30% and ≥2 points, with either a ≥1-point decrease from baseline in the rectal bleeding subscore or a rectal bleeding subscore of 0 or 1. This corresponds to the main definition given in Table 11 in Section 5.5.1.1.

Histologic-endoscopic mucosal improvement: Combination of histologic improvement (neutrophil infiltration in <5% of crypts, no crypt destruction, and no erosions, ulcerations or granulation tissue according to the Geboes[2)] grading system) and endoscopic improvement as defined above.

5.4.1.2. Main Estimands for Secondary Endpoints

Week 12 secondary endpoints

For secondary endpoints at Week 12, the same estimand that is specified for the primary endpoint will be used (except for the "variable attribute"). The variable attributes are given in Table 7 below:

Table 7: Variable Attributes of Secondary Endpoints at Week 12

Endpoint	Variable attribute
Symptomatic remission Week 12	A binary response variable (response/nonresponse) where response is defined as achieving a stool frequency subscore of 0 or 1 and not increased from baseline, and a rectal bleeding subscore of 0 at Week 12 and none of the ICEs in categories 1 to 4 and 6 outlined in Table 6 has occurred prior to the Week 12 visit
Endoscopic improvement Week 12	A binary response variable (response/nonresponse) where response is defined as achieving an endoscopy subscore of 0 or 1 with no friability present on the endoscopy at Week 12, where none of the ICEs in categories 1 to 4 and 6 outlined in Table 6 has occurred prior to the Week 12 visit
Clinical response Week 12	A binary response variable (response/nonresponse) where response is defined as achieving a decrease from induction baseline in the modified Mayo score by ≥30% and ≥2 points, with either a ≥1-point decrease from baseline in the rectal bleeding subscore or a rectal bleeding subscore of 0 or 1 at Week 12, where none of the ICEs in categories 1 to 4 and 6 outlined in Table 6 has occurred prior to the Week 12 visit
Histologic-endoscopic mucosal improvement Week 12	A binary response variable (response/nonresponse) where response is defined as a combination of histologic improvement [neutrophil infiltration in <5% of crypts, no crypt destruction, and no erosions, ulcerations or granulation tissue according to the Geboes grading system (Appendix 11)] and endoscopic improvement [endoscopy subscore of 0 or 1 with no friability present] at Week 12, where none of the ICEs in categories 1 to 4 and 6 outlined in Table 6 has occurred prior to the Week 12 visit

Week 24 Secondary endpoint

The main estimand for Week 24 secondary endpoints is defined as follows:

i) Treatment by Week 24:

- Experimental 1: Guselkumab CCI q4w [Weeks 0, 4, and 8] followed by guselkumab 200 mg SC q4w

- Experimental 2: Guselkumab CCI q4w [Weeks 0, 4, and 8] followed by guselkumab 100 mg SC q8w
- Control: Placebo SC q4w (Weeks 0, 4, 8, 12, 16, 20)

ii) Population: Participants with moderately to severely active ulcerative colitis as defined by the inclusion/exclusion criteria.

iii) Variable (Endpoint): The variable attributes for secondary endpoints at Week 24 are given in Table 8 below

Table 8: Intercurrent Events and Respective Analysis Strategies for Week 24

Endpoint	Variable attribute
Clinical remission at Week 24	A binary response variable (response/nonresponse) where response is defined as achieving a stool frequency subscore of 0 or 1 and not increased from baseline, a rectal bleeding subscore of 0, and an endoscopy subscore of 0 or 1 with no friability present on the endoscopy at Week 24, where none of the ICEs in categories 1 to 4 and 6 outlined in Table 6 has occurred prior to the Week 24 visit.
Symptomatic remission at Week 24	A binary response variable (response/nonresponse) where response is defined as achieving a stool frequency subscore of 0 or 1 and not increased from baseline, and a rectal bleeding subscore of 0 at Week 24, and none of the ICEs in categories 1 to 4 and 6 outlined in Table 6 has occurred prior to the Week 24 visit.
Endoscopic improvement at Week 24	A binary response variable (response/nonresponse) where response is defined as achieving an endoscopy subscore of 0 or 1 with no friability present on the endoscopy at Week 24, where none of the ICEs in categories 1 to 4 and 6 outlined in Table 6 has occurred prior to the Week 24 visit.
Clinical response at Week 24	A binary response variable (response/nonresponse) where response is defined as achieving a decrease from induction baseline in the modified Mayo score by ≥30% and ≥2 points, with either a ≥1-point decrease from baseline in the rectal bleeding subscore or a rectal bleeding subscore of 0 or 1 at Week 24, where none of the ICEs in categories 1 to 4 and 6 outlined in Table 6 has occurred prior to the Week 24 visit

iv) Intercurrent Events and Corresponding Strategies: ICEs and the corresponding analysis strategies are defined in Table 6.

v) Population-level summary: Difference in proportions of participants who achieved the binary response at Week 24 as defined in the variable attribute above between each of the guselkumab groups and the placebo group.

5.4.1.3. Supplementary Estimands for Secondary Endpoints

5.4.1.3.1. Treatment policy for discontinuations (Estimand 5)

Supplementary estimands (Estimand 5 for secondary endpoints) handle all treatment discontinuation ICEs, i.e. ICEs 3, 5 and 6 in Table 6, by the treatment policy strategy. ICEs 1, 2 and 4 in Table 6 continue to be handled by the composite strategy. This leads to the variable definitions given in Table 9.

Table 9:Variable Attributes of Secondary Endpoints for supplementary estimand

Endpoint	Variable attribute
Symptomatic remission Week 12	A binary response variable (response/nonresponse) where response is defined as achieving a stool frequency subscore of 0 or 1 and not increased from baseline, and a rectal bleeding subscore of 0 at Week 12 and none of the ICEs in categories 1, 2 and 4 outlined in Table 6 has occurred prior to the Week 12 visit
Endoscopic improvement Week 12	A binary response variable (response/nonresponse) where response is defined as achieving an endoscopy subscore of 0 or 1 with no friability present on the endoscopy at Week 12, where none of the ICEs in categories 1, 2 and 4 outlined in Table 6 has occurred prior to the Week 12 visit
Clinical response Week 12	A binary response variable (response/nonresponse) where response is defined as achieving a decrease from induction baseline in the modified Mayo score by ≥30% and ≥2 points, with either a ≥1-point decrease from baseline in the rectal bleeding subscore or a rectal bleeding subscore of 0 or 1 at Week 12, where none of the ICEs in categories 1, 2 and 4 outlined in Table 6 has occurred prior to the Week 12 visit
Histologic-endoscopic mucosal improvement Week 12	A binary response variable (response/nonresponse) where response is defined as a combination of histologic improvement [neutrophil infiltration in <5% of crypts, no crypt destruction, and no erosions, ulcerations or granulation tissue according to the Geboes grading system (Appendix 11)] and endoscopic improvement [endoscopy subscore of 0 or 1 with no friability present] at Week 12, where none of the ICEs in categories 1, 2 and 4 outlined in Table 6 has occurred prior to the Week 12 visit
Clinical remission Week 24	A binary response variable (response/nonresponse) where response is defined as achieving a stool frequency subscore of 0 or 1 and not increased from baseline, a rectal bleeding subscore of 0, and an endoscopy subscore of 0 or 1 with no friability present on the endoscopy at Week 24, where none of the ICEs in categories 1, 2 and 4 outlined in Table 6 has occurred prior to the Week 24 visit.
Symptomatic remission Week 24	A binary response variable (response/nonresponse) where response is defined as achieving a stool frequency subscore of 0 or 1 and not increased from baseline, and a rectal bleeding subscore of 0 at Week 24, and none of the ICEs in categories 1, 2 and 4 outlined in Table 6 has occurred prior to the Week 24 visit.
Endoscopic improvement Week 24	A binary response variable (response/nonresponse) where response is defined as achieving an endoscopy subscore of 0 or 1 with no friability present on the endoscopy at Week 24, where none of the ICEs in categories 1, 2 and 4 outlined in Table 6 has occurred prior to the Week 24 visit.
Clinical response Week 24	A binary response variable (response/nonresponse) where response is defined as achieving a decrease from induction baseline in the modified Mayo score by ≥30% and ≥2 points, with either a ≥1-point decrease from baseline in the rectal bleeding subscore or a rectal bleeding subscore of 0 or 1 at Week 24, where none of the ICEs in categories 1, 2 and 4 outlined in Table 6 has occurred prior to the Week 24 visit.

5.4.1.3.2. Alternative Mayo Score Calculation (Estimand 6)

This supplementary estimand (Estimand 6) follows the main estimand for the secondary endpoints in all attributes. The eligible days to be used for the Mayo scoring for this estimand are all available days among the seven days prior to each visit, excluding programmatically the following two days:

3. The day(s) of a procedure (including colonoscopy or sigmoidoscopy)

4. The day preceding such a procedure.

This estimand does not apply to endoscopic improvement at Week 12 / Week 24 nor to histologic-endoscopic mucosal improvement at Week 12.

5.4.2. Analysis Methods

5.4.2.1. Main Estimators (Analyses) for the Main Estimands for the Secondary Estimands

The analysis of the secondary endpoints will be based on the Full Analysis Set. Participants will be analyzed according to the treatment group to which they were randomized regardless of the treatment they received. After accounting for the ICE strategies, any missing data for the secondary endpoints will be handled by nonresponder imputation. In particular the following rules will be applied:

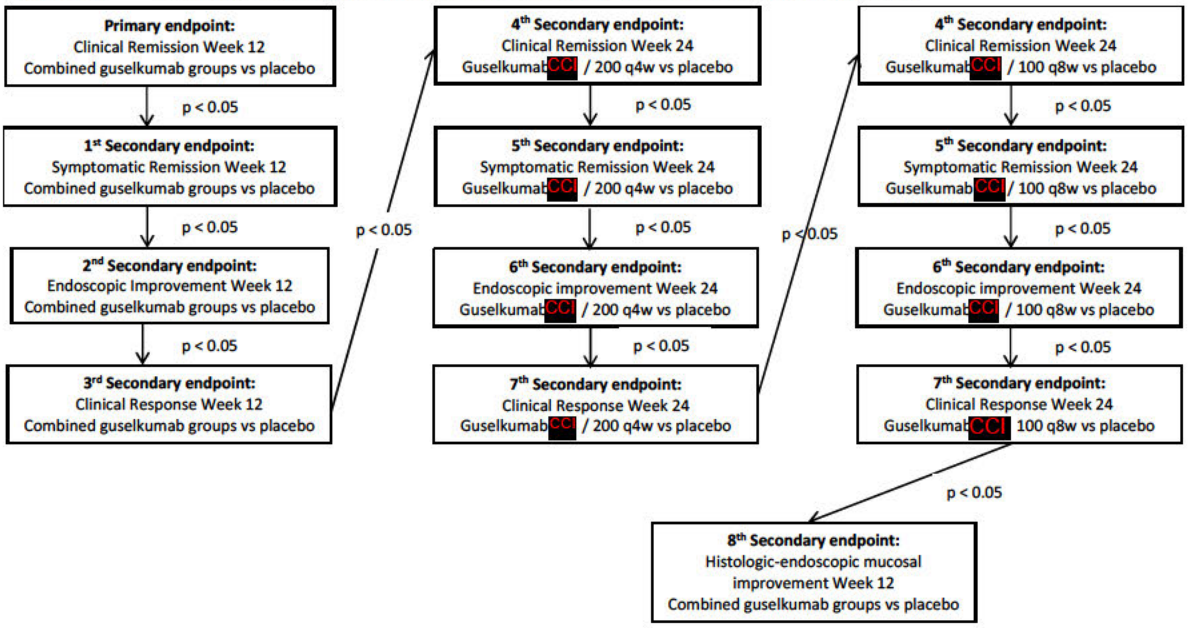
- Participants who are missing baseline stool frequency at baseline will be considered as not having achieved clinical remission or symptomatic remission at any timepoint
- Participants who are missing any of the modified Mayo subscores at baseline will be considered as not having achieved clinical response at any timepoint
- Participants who have missing endoscopy subscore at Week 12 will be considered as not having achieved endoscopic improvement or histologic-endoscopic mucosal improvement at Week 12;
- Participants who are missing either the stool frequency or the rectal bleeding subscore at Week 12 or at Week 24 will be considered as not having achieved symptomatic remission at the respective timepoint;
- Participants who are missing any of the modified Mayo subscores at Week 12 or at Week 24 will be considered as not having achieved clinical response at the respective timepoint;
- Participants who are missing any of the modified Mayo subscores at Week 24 will be considered as not having achieved clinical remission at Week 24;
- Participants who are missing any or all of the components in the Geboes grading system pertaining to the histologic improvement at Week 12 will be considered as not having achieved histologic-endoscopic mucosal improvement at Week 12.

Summaries of the proportion of participants achieving each secondary endpoint by treatment group along with the adjusted treatment difference between the combined guselkumab SC induction group (Week 12 endpoints) / individual guselkumab SC induction groups (Week 24 endpoints) and the placebo group, as well as the associated 95% confidence interval will be presented for each secondary endpoint. Each comparison against placebo will be performed in terms of the common (overall) risk difference for multiway 2x2 tables at a two-sided significance level of 0.05. Specifically, the adjusted treatment differences, 95% confidence intervals, and p-values will be in terms of the common risk difference by use of Mantel-Haenszel stratum weights and the Sato variance estimator; stratification is by ADT-IR status (Yes/No), and Mayo endoscopy subscore at baseline (moderate [2] or severe [3]).

5.4.2.2. Multiple Testing Procedure

A multiple testing procedure is planned to control the overall Type 1 error rate in the study at the 2-sided 0.05 significance level. The planned procedure is depicted in Figure 2. It follows a fixed sequence approach, where the primary endpoint and the Week 12 secondary endpoints (with the exception of histologic-endoscopic mucosal improvement) for the combined guselkumab groups vs. placebo are tested prior to the secondary endpoints at Week 24. For the Week 24 endpoints, the testing sequence continues with testing all four endpoints in the guselkumab CCI [redacted] q4w [Weeks 0, 4, and 8] followed by guselkumab 200 mg SC q4w group against placebo and then testing all four endpoints in the guselkumab CCI [redacted] q4w [Weeks 0, 4, and 8] followed by guselkumab 100 mg SC q8w group against placebo. Histologic-endoscopic mucosal improvement at Week 12 is the last element of the testing sequence and is tested for the combined guselkumab groups against placebo.

Figure 2: Multiple Testing Procedure for Primary and Secondary Endpoints



Note: All p-values are two-sided

5.4.2.3. Estimator (Analysis) for the Supplementary Estimands of the Secondary Endpoints

The same estimators used for the main estimands for the secondary endpoints will be used for the supplemental estimands.

5.4.2.4. Subgroup Analyses

To evaluate the consistency of the secondary endpoints analyses, subgroup analyses for subgroups defined by the stratification variables (ADT-IR yes/no, baseline endoscopic subscore moderate [2]/severe [3]) will be performed using the FAS. The adjusted treatment difference of the combined guselkumab SC group vs. the placebo SC group and the associated p-values and 95%

confidence intervals will be provided. The treatment differences, p-values and confidence intervals will be obtained based on the common risk difference (using Mantel-Haenszel stratum weights and the Sato variance estimator), stratified by the remaining stratification factor.

5.5. Exploratory Endpoints Analysis

The exploratory clinical endpoints, and exploratory endpoints related to Health-Related Quality of Life outcomes, inflammatory biomarkers and health economics will be analyzed. These timepoints will be summarized for each applicable timepoint through Week 48 using the FAS.

5.5.1. Exploratory Clinical Endpoints

5.5.1.1. Mayo Score Based Endpoints

In addition to the primary endpoint analysis described in Section 5.3.1, clinical remission with definitions given in Table 10 will be analyzed at Week 12, Week 24 and Week 48.

Table 10: Clinical Remission Definitions for Week 12, Week 24 and Week 48

Identifier	Definition
Main definition (Week 24 and Week 48 only)	Stool frequency subscore of 0 or 1 and not increased from baseline, a rectal bleeding subscore of 0, and an endoscopy subscore of 0 or 1 with no friability present on the endoscopy (as given in Section 5.3.1)
Legacy definition	Defined as a full Mayo score ≤2 points, with no individual subscore >1 point
Alternative definition	Stool frequency subscore of 0, a rectal bleeding subscore of 0, and an endoscopy subscore of 0 or 1 with no friability present on the endoscopy

In addition to clinical response analyses described in Section 5.4.1.1, analyses according to definitions in Table 11 will be performed for Week 12, Week 24 and Week 48.

Table 11: Clinical Response Definitions for Week 12, Week 24 and Week 48

Identifier	Definition
Main definition (Week 24 and Week 48 only)	Decrease from baseline in the modified Mayo score by ≥30% and ≥2 points, with either a ≥1-point decrease from baseline in the rectal bleeding subscore or a rectal bleeding subscore of 0 or 1 (as given in Section 5.4.1.1)
Legacy definition	Decrease from baseline in the full Mayo score by ≥30% and ≥3 points, with either a ≥1-point decrease from baseline in the rectal bleeding subscore or a rectal bleeding subscore of 0 or 1

Symptomatic remission as defined in Section 5.4.1.1 will be evaluated at all visits, overall and by ADT-IR stratification factor. In addition, the following clinical endpoints will be evaluated:

- Deep symptomatic remission (stool frequency subscore of 0 and rectal bleeding subscore of 0) over time
- Change from baseline in absolute stool number at each visit through Week 24.
- Mayo stool frequency subscore of 0 or 1 over time

- Mayo rectal bleeding subscore of 0 over time
- Time to symptomatic remission: Time from randomization to the date of first time meeting symptomatic remission definition through Week 24 / Week 48

In addition to evaluations done at individual timepoints, the following evaluations combining timepoints will be performed (all analyses based on primary/main definitions):

- Sustained clinical remission: Clinical remission at Week 12 and Week 24 (Week 24 DBL) and at Week 12, Week 24 and Week 48 (Week 48 DBL)
- Maintenance of clinical remission at Week 24 and at Week 48 among participants achieving clinical remission at Week 12
- Maintenance of clinical remission through Week 48 (at both Week 24 and Week 48) among participants achieving clinical remission at Week 12
- Sustained clinical response: Clinical response at Week 12 and Week 24 (Week 24 DBL), Week 12, Week 24 and Week 48 (Week 48 DBL)
- Symptomatic remission over time among participants achieving symptomatic remission at Week 12 (Week 48 DBL only)
- Durable symptomatic remission through Week 48: Defined as achieving symptomatic remission for >80% of all visits between Week 0 (excluding) and Week 48 (including)

The changes from baseline over time will be evaluated for full Mayo score, modified Mayo score, partial Mayo score, and Mayo stool frequency and rectal bleeding and subscores. The full Mayo score and the modified Mayo scores require an endoscopy and will therefore only be evaluated at Week 12, Week 24 and Week 48. The partial Mayo score will be evaluated at all visits.

The following evaluations for completeness/missingness of the Mayo score will be done at Week 12, Week 24 and Week 48:

- Summary of completeness of modified Mayo score and the full Mayo score
- Number of participants who experienced intercurrent events prior to Week 12, prior to Week 24 and prior to Week 48
- Summary of Mayo subscores through Week 12, through Week 24, and through Week 48
- Summary of missing modified Mayo score and the full Mayo score components after accounting for ICE strategies and by intercurrent event occurrence

5.5.1.2. Histology and Endoscopy Based Endpoints

The following endpoints will be analyzed at Week 12, Week 24 (endoscopy based endpoints only) and Week 48:

- Endoscopic normalization: Defined as a Mayo endoscopy subscore of 0.
- Endoscopic improvement as defined in Section 5.4.1.1

- Histologic improvement: Defined as neutrophil infiltration in <5% of crypts, no crypt destruction, and no erosions, ulcerations or granulation tissue according to the Geboes grading system
- Histologic improvement as defined above in participants not meeting histologic improvement criteria at baseline
- Histologic remission: Defined as absence of neutrophils from the mucosa (both lamina propria and epithelium), no crypt destruction, and no erosions, ulcerations or granulation tissue according to the Geboes grading system
- Histologic-endoscopic mucosal improvement as defined in Section 5.4.1.1
- Summary of change from baseline in Robarts Histopathology Index[7)] (see Appendix 13)
- Summary of change from baseline in Geboes total score, where the continuous histology score is derived as the sum of all Geboes Grades and may take on values from 0 to 22
- Summary of change from baseline in Geboes high activity subscore, where the continuous histology score is derived as the sum of Geboes Grades 3, 4, and 5 that define histologic healing and may take on values from 0 to 10
- Summary of change from baseline in Geboes low activity subscore, where the continuous histology score is derived as the sum of Geboes Grades 0, 1, 2A and 2B and may take on values from 0 to 12
- Summary of change from baseline in Nancy Histological Index
- Histologic remission based on the Nancy Histological Index: Defined as Nancy Histological Index[5)] score ≤1 (see Appendix 12)
- Histologic-endoscopic mucosal improvement (alternative definition 1): Defined as a combination of histologic remission (defined as absence of neutrophils from the mucosa [both lamina propria and epithelium], no crypt destruction, and no erosions, ulcerations or granulation tissue according to the Geboes grading system) and endoscopic improvement (endoscopy subscore of 0 or 1 with no friability), as defined above
- Deep histologic-endoscopic mucosal remission: Defined as combination of histologic remission (defined as absence of neutrophils from the mucosa [both lamina propria and epithelium], no crypt destruction, and no erosions, ulcerations or granulation tissue according to the Geboes grading system) and endoscopic normalization (endoscopy subscore of 0), as defined above
- Combination of clinical remission and histologic remission (i.e., symptomatic remission, endoscopic improvement and histologic remission) at Week 48
- Combination of histologic-endoscopic mucosal improvement (alternative definition 1, as defined above) and symptomatic remission and fecal calprotectin normalization (defined as fecal calprotectin concentration ≤ 250 mg/kg) at Week 48
- Combination of deep histologic-endoscopic mucosal remission (as defined above) and symptomatic remission and fecal calprotectin normalization (defined as fecal calprotectin concentration ≤ 250 mg/kg) at Week 48

- Sustained endoscopic improvement: Endoscopic improvement at Week 12, Week 24 and Week 48
- Maintenance of endoscopic improvement at Week 48 among participants achieving endoscopic improvement at Week 12
- Maintenance of endoscopic improvement through Week 48 (at both Week 24 and Week 48) among participants achieving endoscopic improvement at Week 12

5.5.1.3. Corticosteroid Endpoints

Corticosteroid-free related endpoints will be evaluated for all participants in the FAS, and for those receiving corticosteroids at baseline. Clinical remission is defined as given in Section 5.3.1. Other corticosteroid related endpoints will be evaluated at all applicable timepoints.

- Corticosteroid-free clinical remission at Week 24: Defined as being in clinical remission at Week 24 and not receiving corticosteroids at Week 24
- Corticosteroid-free clinical remission at Week 24 among participants who were receiving concomitant corticosteroids at baseline
- Corticosteroid-free clinical remission at Week 48: Defined as being in clinical remission at Week 48 and not receiving corticosteroids at Week 48
- Corticosteroid-free clinical remission at Week 48 among participants who were receiving concomitant corticosteroids at baseline
- 8-week corticosteroid-free clinical remission at Week 48: Defined as being in clinical remission at Week 48 and not having received corticosteroids between Week 40 and Week 48
- 8-week corticosteroid-free clinical remission at Week 48: Defined as being in clinical remission at Week 48 and not having received corticosteroids between Week 40 and Week 48 among participants who were receiving concomitant corticosteroids at baseline
- 12-week corticosteroid-free clinical remission at Week 48: Defined as being in clinical remission at Week 48 and not having received corticosteroids between Week 36 and Week 48
- 12-week corticosteroid-free clinical remission at Week 48: Defined as being in clinical remission at Week 48 and not having received corticosteroids between Week 36 and Week 48 among participants receiving who were receiving concomitant corticosteroids at baseline
- Elimination of concomitant corticosteroids at Week 48 among participants who were receiving concomitant corticosteroids at baseline
- Elimination of concomitant corticosteroids for at least 8 weeks prior to Week 48 among participants who were receiving concomitant corticosteroids at baseline
- Elimination of concomitant corticosteroids for at least 12 weeks prior to Week 48 among participants who were receiving concomitant corticosteroids at baseline
- Average daily prednisone-equivalent corticosteroid dose (excluding budesonide and beclomethasone dipropionate) at each visit, and change from baseline, through Week 48 among participants who were receiving concomitant corticosteroids for the treatment of UC at baseline

- Time to being corticosteroid-free prior to Week 48 among participants who were receiving concomitant corticosteroids at baseline

5.5.2. Endpoints Related to Health-Related Quality of Life

Health-related quality of life endpoints are based on the Inflammatory Bowel Disease Questionnaire (IBDQ), Patient-Reported Outcomes Measurement Information System (PROMIS-29), and UC Patient-reported outcomes signs and symptoms (UC-PRO/SS).

5.5.2.1. Inflammatory Bowel Disease Questionnaire (IBDQ) Related Endpoints

The IBDQ[4)] is a validated, 32-item, self-reported questionnaire for participants with IBD that will be used to evaluate the disease specific HRQoL, across 4 dimension scores: Bowel symptoms (loose stools, abdominal pain), systemic functions (fatigue, altered sleep pattern), social function (work attendance, need to cancel social events), and emotional function (anger, depression, irritability). Scores range from 32 to 224, with higher scores indicating better outcomes.

The individual IBDQ dimensions will be calculated when no more than 1 item is missing in the dimension. If an item is missing, it will be estimated using the average value across the non-missing items. If any of the 4 dimensions of the IBDQ cannot be calculated, then the total IBDQ score cannot be calculated and will be missing for that visit.

The following will be evaluated at all applicable timepoints:

- The change from baseline over time in the total score of the IBDQ
- The change from baseline over time in each of the 4 dimensions of the IBDQ
- A ≥ 16-point improvement from baseline in the IBDQ score
- A > 20- point improvement from baseline in the IBDQ score
- IBDQ remission, defined as total IBDQ score ≥170 (Irvine et al, 1994[4)]; Higgins et al, 2005[3)])
- Sustained improvement: A ≥ 16-point improvement from baseline in the IBDQ score at both Week 12 and Week 48
- Sustained IBDQ remission: Total IBDQ score ≥170 at both Week 12 and Week 48

5.5.2.2. PROMIS-29 Related Endpoints

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). Norm-based scores have been calculated for each domain on the PROMIS measures, with a score of 50 representing the mean or average for the reference population. On symptom-oriented domains of PROMIS-29 (anxiety, depression, fatigue, pain interference, and sleep disturbance), higher scores represent worse symptomatology. On the function-oriented domains (physical functioning and social role), higher scores represent better functioning.

The following will be evaluated at all applicable timepoints:

- The change from baseline over time in each of the 7 domains T-scores of PROMIS-29 and the pain intensity
- A ≥ 5-point improvement in each of the 7 domains T-scores of PROMIS-29 over time
- A ≥ 7-point improvement in each of the 7 domains T-scores of PROMIS-29 over time
- A ≥ 9-point improvement in each of the 7 domains T-scores of PROMIs-29 over time
- A ≥ 3-point improvement in the PROMIS-29 pain intensity over time
- A ≥ 3-point improvement from baseline in the PROMIS-29 pain intensity over time among the participants with pain intensity ≥ 3 at baseline

5.5.2.3. UC-PRO/SS Related Endpoints

The UC-PRO/SS measure was developed to standardize the quantification of gastrointestinal signs and symptoms of UC through direct report from patient ratings. It is a 9-item daily diary designed to quantify the effects of treatment on patient-reported signs and symptoms of UC (see Appendix 15). The UC-PRO/SS includes two scales – bowel signs and symptoms (6 items) and functional symptoms (3 items). Each scale is scored separately; there is no total score for the UC-PRO/SS. Only eligible days as defined for the Mayo score (see Section 5.3.1) will be used for analyses of the UC-PRO/SS.

The following will be evaluated at all applicable timepoints:

- The change from baseline in the weekly average score for each of the 9 items assessed by the UC-PRO/SS
- The change from baseline in the bowel signs and symptoms scale of UC assessed by UC-PRO/SS
- No bowel urgency, defined as a rounded weekly average score of 0 in question 7 of the UC-PRO/SS, over time
- Resolution of bowel urgency, defined as a rounded weekly average score of 0 in question 7 of the UC-PRO/SS among those with a rounded weekly average bowel urgency score ≥1 at baseline
- No bowel urgency (alternative definition), defined as a score of '0' in question 7 of the UC-PRO/SS on the three days closest to each visit, over time
- Resolution of bowel urgency (alternative definition), defined as a score of '0' in question 7 of the UC-PRO/SS on the three days closest to each visit among those with a bowel urgency score ≥1 for at least one of the three days closest to baseline
- No abdominal pain, defined as a rounded weekly average score of 0 in question 8 of the UC-PRO/SS, over time
- Resolution of abdominal pain, defined as a rounded weekly average score of 0 in question 8 of the UC-PRO/SS among those with a rounded weekly average abdominal pain score ≥1 at baseline

- No abdominal pain (alternative definition), defined as a score of '0' in question 8 of the UC-PRO/SS on the three days closest to each visit, over time
- Resolution of abdominal pain (alternative definition), defined as a score of '0' in question 8 of the UC-PRO/SS on the three days closest to each visit among those with an abdominal pain score ≥1 for at least one of the three days closest to baseline
- No incontinence, defined as a rounded weekly average score of '0' in question 5 of the UC-PRO/SS, over time
- Resolution of incontinence, defined as a rounded weekly average score of '0' in question 5 of the UC-PRO/SS among those with a rounded weekly average incontinence score ≥1 at baseline
- no incontinence (alternative definition), defined as a score of '0' in question 5 of the UC-PRO/SS on the three days closest to each visit, over time
- Resolution of incontinence (alternative definition), defined as a score of '0' in question 5 of the UC-PRO/SS on the three days closest to each visit among those with an incontinence score ≥1 for at least one of the three days closest to baseline
- The change from baseline in the functional symptoms scale of UC assessed by UC-PRO/SS
- A clinically meaningful change (defined as an improvement from baseline of ≥ 5 points) in the bowel signs and symptoms scale assessed by UC-PRO/SS
- A clinically meaningful change (defined as an improvement from baseline of ≥ 1.5 points) in the functional symptoms scale assessed by UC-PRO/SS
- A clinically meaningful change in the urgency assessed by UC-PRO/SS, defined as follows:
  - Improvement from baseline ≥ 1 point for participants with a baseline weekly average score for item #7 of at least 1 point
  - Improvement from baseline ≥ 2 points for participants with a baseline weekly average score for item #7 of at least 2 points

5.5.3. Endpoints Related to Inflammatory Biomarkers

For C-reactive protein (CRP) and fecal calprotectin, the following will be evaluated over all applicable timepoints:

- Change from baseline in CRP concentration [mg/L] over time
- Change from baseline in CRP concentration [mg/L] over time for participants with baseline CRP >3 mg/L
- CRP normalization over time for participants with baseline CRP >3 mg/L, where CRP normalization is defined as CRP ≤3 mg/L
- Decrease in CRP concentration of at least 50% relative to baseline among those with a baseline CRP >3 mg/L
- Decrease in CRP concentration of at least 75% relative to baseline among those with a baseline CRP >3 mg/L

- Change from baseline in fecal calprotectin concentration [mg/kg] over time
- Decrease in fecal calprotectin concentration of at least 50% relative to baseline among those with a baseline fecal calprotectin >250 mg/kg
- Decrease in fecal calprotectin concentration of at least 75% relative to baseline among those with a baseline fecal calprotectin >250 mg/kg
- Change from baseline in fecal calprotectin concentration [mg/kg] over time for participants with baseline fecal calprotectin >250 mg/kg
- Fecal calprotectin ≤150 mg/kg over time for participants with baseline fecal calprotectin >150 mg/kg
- Fecal calprotectin ≤250 mg/kg over time for participants with baseline fecal calprotectin >250 mg/kg

5.5.4. Analysis Methods

Unless otherwise specified, the exploratory endpoints listed and defined in Sections 5.5.1 to 5.5.3 will be analyzed based on the FAS according to the randomized intervention group, regardless of the intervention actually received.

No confirmatory tests will be performed on exploratory endpoints, regardless of the significance for primary and secondary endpoints. Nominal p-values will be presented without adjustments for multiple comparisons.

Descriptive statistics (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.

Binary endpoints

The main estimand approach and the estimator used for the secondary endpoints (Section 5.4.1.2) will be used to summarize and compare binary endpoints. Intercurrent events will be handled as specified in Table 6.

The handling of ICEs specified in Section 5.4.1.2 will be applied to each of the above binary endpoints. Participants with ICEs in categories 1-4 and 6 will be considered not to have achieved the binary endpoint, while observed values will be used for participants experiencing an ICE in category 5.

Participants with any missing data for an endpoint after application of ICE handling rules will be imputed as not having achieved the associated binary endpoint. Binary endpoints will be summarized with the number and percentage of participants who achieve the endpoint by intervention group. To compare the proportion of participants achieving the endpoints, MH tests stratified by the ADT-IR status (Yes/No), and Mayo endoscopy subscore at baseline (moderate [2] or severe [3]) will be provided. In case of rare events, Fisher's exact test will be used for intervention comparisons.

Continuous endpoints

For continuous endpoints, handling rules for ICEs as defined in Table 6:

- Composite strategy (categories 1-4 and 6): Baseline values will be assigned from the point of ICE onward
- Treatment policy strategy (category 5): Observed values will be used, if available

To account for missing data (after applying ICE handling strategies) for continuous endpoints that are collected at multiple post-baseline time points through Week 24 / Week 48, an MMRM will be used, under the MAR assumption, to compute nominal p-values between both guselkumab SC groups and placebo SC. The model will include all available data from all intervention groups through Week 24 / Week 48 for assessing efficacy. Missing data will thus not be imputed but rather accounted for through correlation of repeated measures in the model. Additionally, if the MMRM normality assumption is in question, an appropriate transformation may be implemented before fitting the MMRM.

The MMRM will include intervention group, respective baseline score, visit and interaction between visit and intervention group as categorical variables as well as ADT-IR status (Yes/No), and Mayo endoscopy subscore at baseline (moderate [2] or severe [3]). An unstructured covariance matrix for repeated measures within a participant will be used. The F-test will use Kenward-Roger's approximation for the 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.

Endpoints that are measured at only one of the visits at Week 12 or Week 24 (e.g., histology based endpoints as described in Section 5.5.1.2) will be compared between the guselkumab SC group(s) and the placebo group using an analysis of covariance (ANCOVA) with treatment group, ADT-IR status (Yes/No), Mayo endoscopy score at baseline (moderate [2] or severe [3]), and corresponding baseline value as covariates.

For analyses using ANCOVA, unless otherwise specified, multiple imputation (same as defined in Section 5.3.3.3.3) will be used for missing data (after applying ICE strategies), under the assumption that the data are missing at random. The analysis will impute the missing score at the respective Week using the corresponding value at baseline and related component score at intermediate visits if applicable, ADT-IR status, Mayo endoscopy score at baseline, and treatment group.

The intervention difference between each guselkumab SC group and placebo SC will be estimated by the difference in the least squares means (LSmeans). The 95% 2-sided CI for the differences in LSmeans and p-values will be calculated based on the MMRM / ANCOVA, as applicable.

Ordinal endpoints

For ordinal endpoints, handling rules for ICEs as defined in Table 6:

- Composite strategy (categories 1-4 and 6): Baseline values will be assigned from the point of ICE onward
- Treatment policy strategy (category 5): Observed values will be used, if available

The intervention comparison between the guselkumab SC groups with placebo SC will be performed using a CMH χ^2 test (row mean scores differ as a hypothesis test for the CMH statistics) stratified by ADT-IR status (Yes/No), and Mayo endoscopy subscore at baseline (moderate [2] or severe [3]). In case of rare events, Fisher's exact test will be used for intervention comparisons. Missing data (after applying the ICE handling strategies) will not be imputed for ordinal endpoints.

Time to event endpoints

The time to event is defined as the time from Week 0 study intervention administration to the date of the first event that occurred through Week 24 / Week 48. Participants who had not had an event by that timepoint or who had terminated study participation prior to having an event will be censored at the last measurement prior to or at that timepoint or their termination date, as applicable.

Participant data will be censored at last measurement prior to or at the time of any ICE in categories 1-4 and 6. Kaplan-Meier (KM) curves will be generated by study intervention group for time to event endpoints. KM-estimates and corresponding p-values will be provided.

5.5.4.1. Subgroup Analyses

Subgroup analyses for subgroups defined by the stratification variables (ADT-IR yes/no, baseline endoscopic subscore moderate [2]/severe [3]) will be performed using the FAS for the following exploratory endpoints (all as defined in Section 5.5.1.2):

- Histologic improvement
- Histologic remission
- Combination of histologic remission and endoscopic improvement

The following analysis approach will be applied:

- The consistency of intervention effect will be evaluated, using the main estimand and handling missing data by the rules specified in Section 5.4.2.1.
- The adjusted treatment difference of the combined guselkumab SC group vs. the placebo SC group and the associated p-values and 95% confidence intervals will be provided. The treatment differences, p-values and confidence intervals will be obtained from a MH test stratified by the remaining stratification factor.

In addition, descriptive analyses will be performed for the rescue analysis set for the following endpoints:

- Clinical remission (primary definition as given in Table 10 only) at Week 48
- Symptomatic remission over time
- Change from baseline in partial Mayo score over time
- 8-week corticosteroid-free clinical remission at Week 48: Defined as being in clinical remission at Week 48 and not having received corticosteroids between Week 40 and Week 48
- Change from baseline in CRP concentration [mg/L] over time
- Change from baseline in fecal calprotectin concentration [mg/kg] over time

Descriptive analyses will be performed for the Week 12 Clinical Responder Analysis Set for the following endpoints:

- Clinical remission at Week 24 and at Week 48
- Symptomatic remission over time
- Deep symptomatic remission at Week 24 and at Week 48
- Mayo stool frequency subscore of 0 or 1 over time
- Mayo rectal bleeding subscore of 0 over time
- Clinical response at Week 24 and at Week 48
- Clinical response through Week 48 (at both Week 24 and Week 48)
- Endoscopic normalization at Week 24 and at Week 48
- Endoscopic improvement at Week 24 and at Week 48
- Change from baseline in CRP concentration [mg/L] over time
- Change from baseline in fecal calprotectin concentration [mg/kg] over time
- Deep symptomatic remission (Mayo stool frequency subscore of 0 and Mayo at rectal bleeding subscore of 0) over time

5.6. Safety Analyses

Safety data, including but not limited to, AEs and changes in laboratory assessments, will be summarized. Since vital signs and physical examination data were not collected on the eCRF, no summaries will be provided for vital signs and physical examination findings.

All safety analyses will be based on the Safety Analysis Set. 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; participants assigned to guselkumab who received only placebo during the study will be analyzed in the placebo group.

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 analyses are planned.

5.6.1. Extent of Exposure

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

The cumulative dose and the total number of administrations of study intervention received through Week 12 / Week 24 / Week 48 will be summarized descriptively by intervention group (combined guselkumab SC group, individual guselkumab SC groups, and placebo SC).

In addition, the distribution of participants by study intervention lot through Week 12 / Week 24 / Week 48 based on the Safety Analysis Set will be provided.

5.6.2. Adverse Events

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

Summary tables will be provided for treatment-emergent adverse events through Week 24 / Week 48:

- Overall summary of AEs, presenting frequency and type of AEs
- AEs by system organ class (SOC) and preferred term (PT)
- Most frequent AEs by PT
- Serious AEs (SAEs) by SOC and PT
- Related AEs as assessed by the investigator by SOC and PT
- AEs leading to discontinuation of study intervention by SOC and PT
- Any reaction at an SC study intervention injection site that was as recorded as an injection-site reaction by the investigator on the electronic case report form (eCRF)
- AEs of interest (including special interest)
- Frequency and type of AEs associated with venous thromboembolism (VTE)

The following tables will also be provided through Week 12:

- Overall summary of AEs, presenting frequency and type of AEs
- AEs by system organ class (SOC) and preferred term (PT)
- Serious AEs (SAEs) by SOC and PT
- Related AEs as assessed by the investigator by SOC and PT
- AEs leading to discontinuation of study intervention by SOC and PT
- Any reaction at an SC study intervention injection site that was as recorded as an injection-site reaction by the investigator on the eCRF

Summaries will be provided by study intervention (combined guselkumab SC group, individual guselkumab SC groups and placebo). In addition to the summary tables, listings will be provided for participants who:

- Had SAEs
- Had AEs leading to discontinuation of study intervention
- Had SAEs of infection (SOC infections and infestations)
- Had AEs of Opportunistic Infections (Standardized MedDRA Query OPPORTUNISTIC INFECTIONS [SMQ] narrow scope search)
- Had AEs of Tuberculosis (Tuberculous Infections MedDRA High Level Term)
- Had AEs of Malignancy (Malignant Tumours [SMQ] narrow scope search)
- Had AEs associated with major adverse cardiovascular events (MACE), which includes death related to cardiovascular event, nonfatal myocardial infarction, and nonfatal stroke
- Had AEs of drug-induced liver injury (Standardized MedDRA Query DRUG RELATED HEPATIC DISORDERS – COMPREHENSIVE SEARCH [SMQ] narrow scope search)
- Had AEs of VTE
- Had AEs of Anaphylaxis or Serum sickness
- Had an injection-site reaction (any reaction at an SC study intervention injection site that was as recorded as an injection-site reaction by the investigator on the eCRF)
- Had AEs of Suicidal Ideation and Behaviour (Standardized MedDRA Query SUICIDE/SELF-INJURY [SMQ] narrow scope search)
- Deaths

In addition to the tables and listings enumerated above, adverse events will also be summarized as events per 100 subject years of follow-up. This would account for the potential for differences in follow-up times and is applicable to tables of AEs through Week 24 and through Week 48.

Risk difference tables including 95% confidence intervals between placebo and combined guselkumab groups at Week 12 and placebo and individual guselkumab groups at Week 48 will be provided for the following:

- AEs leading to discontinuation

- Serious AEs
- AEs of special interest as defined above

These tables will be provided based on raw cumulative incidence proportions (Weeks 12, 24 and 48), and based on events per 100 subject years of follow-up (Week 48 only).

5.6.3. Additional Safety Assessments

5.6.3.1. Clinical Laboratory Tests

Clinical laboratory tests will be displayed for the participants included in the safety analysis set. The laboratory tests given in Table 12 will be performed, and summaries will be provided.

Table 12: Protocol-required safety laboratory assessment

Laboratory Assessments	Parameters	
Hematology	Platelet count Hemoglobin Hematocrit	White Blood Cell (WBC) count with Differential: Neutrophils Lymphocytes Monocytes Eosinophils Basophils
Clinical Chemistry	Sodium Potassium Chloride Blood urea nitrogen (BUN) Creatinine Aspartate aminotransferase (AST)/Serum glutamic-oxaloacetic transaminase Alanine aminotransferase (ALT)/Serum glutamic-pyruvic transaminase	Total and direct bilirubin Alkaline phosphatase Calcium Phosphate Albumin Total protein
	Biochemical Hy's Law case (ALT or AST ≥3 x ULN and Total bilirubin ≥2 x ULN) reporting requirements are defined in Section 8.3.1 of the protocol.	

Descriptive statistics and graphical displays will be presented for all chemistry and hematology, laboratory tests at scheduled time points.

Clinical laboratory tests are to be graded based on modified National Cancer Institute Common Terminology Criteria for Adverse Events (NCI-CTCAE) version 5 (see Appendix 9) except ALT, AST, total bilirubin and alkaline phosphatase, which will be graded using the predefined upper limit of normal thresholds. Laboratory tests not included in the NCI-CTCAE (Table 15) or the predefined ULN thresholds of liver tests (Table 16) will not be presented in the corresponding tables or listings.

Summary of clinical laboratory tests by study intervention group over time will be provided for participants in the Safety Analysis Set:

- Change from baseline will be summarized for chemistry and hematology tests and displayed by intervention group (for analyses through Week 24: combined and individual guselkumab

SC groups and placebo, for analyses through Week 48: individual guselkumab SC groups and placebo).

- Line graphs for liver function tests
- Summary of maximum/worst toxicity grade for post-baseline laboratory values through Week 12 / Week 24 / Week 48 for the predefined hematology and chemistry laboratory parameters except for ALT, AST, alkaline phosphatase, and total bilirubin
- Shift tables for maximum/worst toxicity grade from baseline through Week 12 / Week 24 / Week 48 will be summarized for the predefined hematology and chemistry laboratory parameters except for ALT, AST, alkaline phosphatase, and total bilirubin
- Summary of maximum post-baseline measurement through Week 12 / Week 24 / Week 48 for ALT, AST, alkaline phosphatase, and total bilirubin relative to ULN threshold as given in Table 16.
- Summary of biochemical Hy's law cases (ALT or AST ≥3 x ULN and total bilirubin ≥2 x ULN) through Week 12 / Week 24 / Week 48

Listings of participants with any abnormal post-baseline laboratory values of maximum/worst toxicity grade ≥ 3 except ALT/AST, alkaline phosphatase, and bilirubin, participants with maximum post-baseline elevated liver function tests (AST or ALT ≥5 x ULN, or total bilirubin or alkaline phosphatase ≥2 x ULN) and participants with biochemical Hy's Law case ((ALT or AST ≥3 x ULN) and (total bilirubin ≥2 x ULN or INR > 1.5) at the same visit), will also be provided by study intervention group, participant, and visit.

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 postbaseline visit minus the assessment at baseline.

There will be no imputation for missing laboratory values.

Any laboratory values given as <X.X in the database will be imputed with the value of the number without the sign for the descriptive statistics and the calculation of changes from baseline, e.g. a value of <2.2 will be imputed as 2.2 for the calculations.

5.6.3.2. Other Safety Parameters

5.6.3.2.1. The Columbia Suicide Severity Rating Scale

The Columbia-Suicide Severity Rating Scale (C-SSRS) [8),9)] 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 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 or AE of suicidal ideation and behavior. Participants with positive (i.e., score >0) postbaseline ideation and behavior through Week 24 / Week 48 will be presented in a data listing.

5.7. Other Analyses

5.7.1. Pharmacokinetics

PK analyses will be performed on the PK analysis set as defined in Table 4.

5.7.1.1. Serum Guselkumab Concentrations

Blood samples for determining serum guselkumab concentrations will be drawn from all participants according to the Schedule of Activities in the protocol.

Descriptive statistics (N, mean, SD, median, range, and IQ range) will be used to summarize serum guselkumab concentrations at each sampling time point. PK data may be displayed graphically, such as mean +/- SD or median serum guselkumab concentrations over time by intervention group.

The following analyses will be performed (for Week 24 analyses: combined and individual guselkumab SC groups; for Week 48 analyses: individual guselkumab SC groups):

- Summary of serum guselkumab concentrations at each visit by study intervention through Week 24 / Week 48
- Proportion of participants without detectable serum guselkumab concentration (below the lower limit of quantification) at each visit by intervention group through Week 24 / Week 48
- Summary of serum guselkumab concentrations at each visit by study intervention through Week 24 / Week 48 on subgroups defined by body weight quartiles.

Unless otherwise specified, the following data handling rules will be applied to PK analyses:

- Participants will be analyzed according to the intervention that they actually received
- All serum concentration summaries for a particular time point will include data obtained from treated participants at the timepoint of interest without imputing any missing data
- A concentration not quantifiable (below the lower limit of quantification) will be treated as 0 in the summary statistics and shown as the lower limit of quantification (< LLOQ) in the data listings
- The data from a participant who meets any of the following dosing deviation criteria will be excluded from the by-visit data analyses from that point onwards:
  - Discontinued study intervention
  - Skipped a study intervention administration
  - Received an incomplete / incorrect dose
  - Received an incorrect study intervention
  - Received an additional dose
  - Received commercial guselkumab

- Additional samples excluded from the analysis are defined as follows:
  - Pre-administration samples if concentration prior to administration of first dose > mean + 3 x standard deviation at a given visit
  - Samples collected after the study intervention administration at the time of visit
  - Samples collected before the study intervention administration at Week 0, 4, 8 and 12 that were outside the visit window +/- 4 days
  - Samples collected before the study intervention administration at Week 16, 20 and 24 that were outside the visit window +/- 7 days
  - Samples collected before the study intervention administration at Week 32, 40, 48 that were outside the visit window +/- 10 days

5.7.1.2. Relationship Between Guselkumab Concentration and Efficacy

To explore the relationship between serum guselkumab concentrations and primary and secondary efficacy endpoints, the following will be explored:

- Relationship between serum guselkumab concentrations (quartiles) and clinical remission at Week 12 / Week 24 / Week 48
- Relationship between serum guselkumab concentrations (quartiles) and symptomatic remission at Week 12 / Week 24 / Week 48
- Relationship between serum guselkumab concentrations (quartiles) and endoscopic improvement at Week 12 / Week 24 / Week 48
- Relationship between serum guselkumab concentrations (quartiles) and histologic-endoscopic mucosal improvement at Week 12 / Week 48
- Relationship between serum guselkumab concentrations (quartiles) and clinical response at Week 12 / Week 48
- Relationship between serum guselkumab concentrations (quartiles) and modified Mayo score at Week 12 / Week 48
- Relationship between serum guselkumab concentrations (quartiles) and CRP concentration [mg/L] at Week 12 / Week 24 / Week 48
- Relationship between serum guselkumab concentrations (quartiles) and fecal calprotectin concentration [mg/kg] at Week 12 / Week 24 / Week 48

The following analyses will be explored graphically:

- Plot of median serum guselkumab concentration [µg/mL] over time through Week 24 / Week 48 by clinical remission status at Week 12 / Week 48

5.7.1.3. Relationship Between Guselkumab Concentration and Safety

To explore the relationship between serum guselkumab concentrations and safety endpoints, the following will be explored:

- Relationship between serum guselkumab concentrations (quartiles) and albumin [g/L] concentration at Week 12 / Week 24 / Week 48

5.7.1.4. Population PK Analysis

When appropriate, population PK analysis will be performed using a nonlinear mixed-effects modeling (NONMEM) approach. Details will be provided in a separate technical report.

5.7.2. Immunogenicity

5.7.2.1. Immunogenicity Analysis

Unless otherwise mentioned, summaries of immunogenicity to guselkumab will be provided based on the Immunogenicity Analysis Set.

Serum samples will be screened for antibodies binding to guselkumab and the titer of confirmed positive samples will be reported.

The incidence and titers of antibodies to guselkumab will be summarized through Week 12 / Week 24/ Week 48 for all participants who received at least one dose of guselkumab and have appropriate samples for detection of antibodies to guselkumab (i.e., participants with at least 1 sample obtained after their first dose of study intervention). The peak titers of antibodies to guselkumab will be provided for participants who are positive for antibodies to guselkumab.

A listing for participants who are positive for antibodies to guselkumab will be provided. The sample antibodies status, the titer, and the neutralizing antibodies status to guselkumab will be listed by visit. This listing will also provide information regarding immunomodulator status at baseline (Yes/No), dose administered, injection-site reactions, serum guselkumab concentration, and modified Mayo score (at applicable visits) for all visits.

5.7.2.2. Neutralizing Antibodies to Guselkumab

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

5.7.2.3. Antibodies vs Pharmacokinetics / Efficacy / Safety

To explore the relationship between antibodies to guselkumab status and serum guselkumab concentrations, efficacy and safety, the following analyses may be performed, if sufficient numbers of participants are positive for antibodies. These analyses will be performed for Week 12 / Week 24 / Week 48 as applicable.

- Summary of serum guselkumab concentrations by antibodies to guselkumab status
- Plots of median trough serum guselkumab concentrations over time by antibodies to guselkumab status

- Summary of clinical remission, symptomatic remission, endoscopic improvement, clinical response, histologic-endoscopic mucosal improvement and change in modified Mayo score by antibodies to guselkumab status
- Summary of injection-site reactions by antibodies to guselkumab status

5.7.3. Pharmacodynamics

5.7.4. Pharmacokinetic/Pharmacodynamic Relationships

The relationship between serum guselkumab concentrations and efficacy measures may be analyzed graphically. 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 an analysis plan and results of the exposure-response analysis will be presented in a separate technical report.

5.7.5. Biomarkers

Changes in serum protein analytes, whole blood RNA, and colonic biopsy RNA obtained over time (where local regulations permit) will be summarized by treatment group. Associations between baseline levels and changes from baseline in selected biomarkers and response to treatment will be explored.

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

Analysis of inflammatory biomarkers (CRP, fecal calprotectin) are described in Section 5.5.3.

5.7.6. Health Economics

The following economic analyses will be conducted based on the FAS.

- Proportion of participants having any UC-related surgery, through Week 24 / Week 48.
- Time to first UC-related surgery through Week 24 / Week 48 if number of participants with events is sufficiently large.

5.7.7. Definition of Subgroups

Demographics:

- Sex (male, female)
- Ethnicity (Hispanic or Latino, not Hispanic or Latino)
- Race (Caucasian, non-Caucasian)
- Race (Asian, American Indian or Alaska Native, Black or African American, Native Hawaiian or Pacific Islander, White)

- Baseline age (≤median age, >median age)
- Baseline age (<65 yrs old, ≥65 yrs old)
- Baseline body weight (≤ 1st quartile, >1st quartile and ≤ 2nd quartile, >2nd quartile and ≤ 3rd quartile, >3rd quartile)
- Region (Asia, Eastern Europe, Rest of the World)

  Asia: China, India, Japan, Taiwan, South Korea, Malaysia

  Eastern Europe: Bulgaria, Czech Republic, Georgia, Hungary, Latvia, Poland, Serbia and Slovakia

  Rest of the World: Belgium, France, Germany, Israel, Italy, Portugal, Spain, Sweden, United Kingdom, Canada, United States, New Zealand, Australia, Argentina, Brazil, Mexico, Jordan and Turkey

Baseline UC clinical disease characteristics

- UC disease duration (≤ 2 years, > 2 years to ≤ 5 years, > 5 years to ≤ 10 years, > 10 years)
- Extent of disease (limited, extensive)
- Severity of UC disease (moderate: 6 ≤ Mayo score ≤ 10, severe: Mayo score > 10)
- Severity of UC disease by modified Mayo score (moderate: 5 ≤ Modified Mayo score ≤ 6, severe: 7 ≤ Modified Mayo score ≤ 9)
- Baseline Mayo endoscopy subscore (moderate: subscore of 2, severe: subscore of 3)
- Albumin (≤ 1st quartile, > 1st quartile and ≤ 2nd quartile, > 2nd quartile and ≤ 3rd quartile, > 3rd quartile)
- Albumin (≤ median, > median)
- CRP (≤ 3 mg/L, > 3 mg/L)
- CRP (≤ median, > median)
- CRP (≤ 1st quartile, > 1st quartile and ≤ 2nd quartile, > 2nd quartile and ≤ 3rd quartile, > 3rd quartile)
- Fecal calprotectin (≤ 250 mg/kg, > 250 mg/kg)
- Fecal calprotectin (≤ 1st quartile, > 1st quartile and ≤ 2nd quartile, > 2nd quartile and ≤ 3rd quartile, > 3rd quartile)

Baseline UC related concomitant medications

- Oral 5-ASA compounds (receiving, not receiving)
- Oral corticosteroids, including budesonide and beclomethasone diproprionate (receiving, not receiving)
- 6-MP/AZA/MTX (receiving, not receiving)
- Oral corticosteroids and 6-MP/AZA/MTX (receiving, not receiving)

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

UC related medication history

- ADT-IR status: Inadequate response to or intolerance of advanced therapy (ADT; i.e., TNFα antagonists, vedolizumab, ozanimod, or approved JAK inhibitors) (yes/no)
- Participants with ADT-IR status
  - Primary nonresponse, secondary nonresponse, or intolerance to
    - One ADT class
      - Anti-TNF only
      - Vedolizumab only
      - Ozanimod only
      - JAK inihibitor only
    - Two ADT classes
    - Three ADT classes
      - Anti-TNF, vedolizumab and JAK inhibitor
    - Two or more ADT classes
    - Three or more ADT classes
    - ADT modalities (biologics [vedolizumab, anti-TNF] vs oral agents [ozanimod, JAK inhibitors])
      - Biologic(s) only
      - Oral agents only
    - ADT classes:
      - At least one anti-TNF (regardless of other ADTs)
      - Vedolizumab (regardless of other ADTs)
      - Ozanimod (regardless of other ADTs)
      - At least one JAK inhibitor (regardless of other ADTs)
- Participants without biologic (i.e., tumor necrosis factor alpha [TNFα] antagonists and vedolizumab) failure (naïve, bio-experienced [but not documented failure])
- Participants with biologic failure
- Participants who are not ADT-IR (naïve, ADT-experienced [but not documented failure])
- Participants who are ADT naïve
- Refractory, dependent or intolerant to corticosteroids (yes, no)
- Refractory or intolerant to 6-MP/AZA (yes, no)

- Refractory, dependent, or intolerant to corticosteroids, but not refractory or intolerant to 6-MP/AZA (yes, no)
- Refractory, dependent or intolerant to corticosteroids, and refractory or intolerant to 6-MP/AZA (yes, no)

5.8. Interim Analyses

No interim analysis is planned for this study.

6. SUPPORTING DOCUMENTATION

6.1. Appendix 1 List of Abbreviations

5-ASA	5-aminosalicylic acid
6-MP	6-mercaptopurine
ADT	advanced therapy
ADT-IR	inadequate response to or intolerance of advanced therapy
AE	adverse event
ALT/SGPT	alanine aminotransferase
ANCOVA	analysis of covariance
AST/SGOT	aspartate aminotransferase
ATC	anatomic and therapeutic class
AZA	azathioprine
BMI	body mass index
BUN	Blood urea nitrogen
CI	confidence interval
CMH	Cochran-Mantel-Haenszel
CRP	C-reactive protein
C-SSRS	Columbia-Suicide Severity Rating Scale
CTCAE	Common Terminology Criteria for Adverse Events
DBL	Database lock
ED	early discontinuation
eCRF	electronic case report form
eCoA	electronic Clinical Outcome Assessment
FAS	full analysis set
iARBM	Integrated Analytical Risk-Based Monitoring
IBD	inflammatory bowel disease
IBDQ	Inflammatory Bowel Disease Questionnaire
ICE	intercurrent event
IQ	interquartile
IV	intravenous
JAK	Janus kinase
KM	Kaplan-Meier
LLOQ	lower limit of quantification
LSmeans	least-squares means
MACE	major adverse cardiovascular event
MAR	missing at random
MedDRA	Medical Dictionary for Regulatory Activities
MH	Mantel-Haenszel
MMRM	mixed-effect model for repeated measurements
Nab	neutralizing antibodies
NCI	National Cancer Institute
NONMEM	nonlinear mixed-effects modeling
NRS	Numerical rating scale
PD	pharmacodynamic(s)
PK	pharmacokinetic(s)
PT	preferred term
PROMIS	Patient-Reported Outcomes Measurement Information System
QTL	Quality Tolerance Limit
SAC	Safety Assessment Committee
SAE	serious adverse event
SAP	Statistical Analysis Plan
SC	subcutaneous
SD	standard deviation
SFU	safety follow-up
SMQs	standardized MedDRA queries
SOC	System organ class

TEAE	treatment-emergent adverse event
TNF	tumor necrosis factor
UC	ulcerative colitis
UC-PRO/SS	UC Patient Reported Outcomes / Signs and Symptoms
ULN	upper limit of normal
VTE	venous thromboembolism
WBC	white blood cell

6.2. Appendix 2 Changes to Protocol-Planned Analyses

No changes to protocol-planned analyses at the time of this version.

6.3. Appendix 3 Demographics and Baseline Characteristics

The number of participants in each analysis set will be summarized and listed by intervention group, and overall. In addition, the number of participants by region, country and site will be presented using the FAS.

Table 13 presents a list of the demographic variables that will be summarized by intervention group and overall for the FAS analysis set, the Rescue Analysis Set and the Week 12 Clinical Responder Analysis Set.

Table 13: Demographic Variables

Continuous Variables:	**Summary Type**
Age (years)	Descriptive statistics (N, mean, standard deviation [SD], median and range [minimum and maximum], and IQ range).
Weight (kg)	
Height (cm)	
Categorical Variables	Frequency distribution with the number and percentage of participants in each category.
Sex (male, female, unknown, undifferentiated)	
Race[a] (American Indian or Alaska Native, Asian (Japanese or Other Asian), Black or African American, Native Hawaiian or other Pacific Islander, White, Multiple)	
Ethnicity (Hispanic or Latino, not Hispanic or Latino)	
Region (Asia, Eastern Europe, Rest of the World)	
Age (<65 yrs old, ≥65 yrs old)	

[a] If multiple race categories are indicated, the Race is recorded as 'Multiple'

Table 14 presents a list of the baseline disease characteristic variables that will be summarized by intervention group and overall for the FAS analysis set, the Rescue Analysis Set and the Week 12 Clinical Responder Analysis Set.

Table 14: Baseline Disease Characteristic Variables

Continuous Variables:	**Summary Type**
UC disease duration (years)	Descriptive statistics (N, mean, standard deviation [SD], median and range [minimum and maximum], and IQ range).
Full Mayo score	
Partial Mayo score	
Modified Mayo score	
CRP	
Fecal Calprotectin	
Albumin	
Categorical Variables	Frequency distribution with the number and percentage of participants in each category.
Severity of UC disease (moderate: 5 ≤ Mayo score ≤ 10, severe: Mayo score > 10)	
Severity of UC disease (moderate: 5 ≤ modified Mayo score ≤ 6, severe: 7 ≤ modified Mayo score ≤ 9)	
Mayo endoscopy score (moderate: subscore of 2; severe: subscore of 3) at baseline	
Extent of disease (extensive, limited to left side of colon)	
CRP (> 3 mg/L)	
Fecal Calprotectin (>150 mg/kg, >250 mg/kg)	
Tobacco or nicotine use status (non-user, prior user, current user)	

6.4. Appendix 4 Protocol Deviations

In general, the following list of major protocol deviations (PDs) may have the potential to impact participants' rights, safety or well-being, or the integrity and/or result of the clinical study. Participants with major protocol deviations will be identified prior to database lock.

The following categories will be considered as major PDs:

- Study intervention administration deviations
- Study entry criteria not met
- Prohibited concomitant medications deviations (as described in Appendix 5)
- Withdrawal criteria met but not withdrawn
- Other

Participants with a major protocol deviation through Week 24 / Week 48 will be summarized by category based on the FAS. A listing of participants who have major protocol deviations will be provided.

A listing of participants with protocol deviations related to COVID-19 will be provided. Should the study be affected by a regional crisis, a listing with respective protocol deviations will be added.

Quality Tolerance Limit (QTL) parameters and thresholds are defined and will be monitored in this study. QTL parameters will be summarized. More details are described in the Integrated Analytical Risk-Based Monitoring (iARBM) Plan.

6.5. Appendix 5 Prior and Concomitant Medications

Concomitant therapies received at baseline can only be discontinued or reduced in dose during the study as described below.

Week 0 Through Week 48

For those participants who are starting the study on approved concomitant medications, the following instructions apply:

- Participants who are receiving oral 5-ASA compounds at baseline (ie, Week 0) should maintain a stable dose through Week 48.
- Corticosteroids must be maintained at baseline doses through Week 12. Participants must begin tapering corticosteroids at Week 12 unless medically not feasible.
- Immunomodulators must be maintained at baseline doses through Week 48.

Enrolled participants should not initiate any of the following concomitant UC-specific medical therapies through Week 48:

- Oral or rectal 5-ASA compounds.
- Oral, parenteral, or rectal corticosteroids, including budesonide and beclomethasone dipropionate.
- 6-MP, AZA, or MTX.
- Antibiotics as a primary treatment for UC.
- Total parenteral nutrition as a treatment for UC.

Week 48 through Week 96

Concomitant therapies for UC, including 5-ASA compounds, corticosteroids, antibiotics, immunomodulators (ie, AZA, 6-MP, or MTX), and/or total parenteral nutrition, may be administered and changed at the discretion of the investigator.

Oral Corticosteroids Tapering

Participants who are receiving oral corticosteroids for the treatment of UC at baseline (Week 0) should maintain a stable dose through Week 12. The oral corticosteroid dose should not be increased above the baseline dose unless due to medical necessity.

At Week 12, all participants who were taking corticosteroids at Week 0 must begin tapering corticosteroids. This tapering is mandatory (unless medically not feasible) and should follow the recommended schedule shown in Section 6.8.1.1 of the study protocol. If participants experience worsening disease activity while tapering corticosteroids, further dose decreases may be suspended, and/or their oral corticosteroid dose may be temporarily increased up to their baseline corticosteroid dose per the discretion of the investigator. For participants whose corticosteroid taper is interrupted, investigators are encouraged to resume tapering within 4 weeks. Tapering may exceed this schedule only if warranted by medical necessity (eg, participant experiencing

corticosteroid-related side effects). After Week 48, corticosteroids may be administered, and dose changed at the discretion of the investigator.

Prior and Concomitant medications will be coded using the World Health Organization Drug Dictionary (WHO-DD). Prior medications are defined as any therapy used before the day of first dose (partial or complete) of study intervention. Concomitant medications are defined as any therapy used on or after the same day as the first dose of study intervention, including those that started before and continue on after the first dose of study intervention.

The following summaries will be provided for the full analysis set, the rescue analysis set and the Week 12 Clinical Responder Analysis Set (for the latter two, except for history of response to or intolerance of corticosteroids and immunomodulators):

- Summary of UC-specific concomitant medications at Baseline
  - The summary will include all approved concomitant medications for which a stable dose before baseline should be maintained: Oral 5-ASA compounds, conventional immunomodulators (6-MP/AZA/MTX), and oral corticosteroids
  - The proportion of participants who receive each concomitant medication will be summarized as well as the proportion of participants who receive at least 1 concomitant medication.
- Summary of history of response to or intolerance of Corticosteroids and Immunomodulators (i.e., 6-mercaptopurine [6-MP], azathioprine [AZA])
- Summary of history of response to advanced therapy
  - The summary will include TNFα antagonist therapies (infliximab, adalimumab and golimumab), vedolizumab, ozanimod, and approved JAK inhibitors as branded or as biosimilars
  - ADT-IR status: Inadequate response to or intolerance of advanced therapy (ADT; i.e., TNFα antagonists, vedolizumab, ozanimod, or approved JAK inhibitors) (yes/no)
  - The number of previous ADT classes will be tabulated
  - The number of participants with primary nonresponse, secondary nonresponse, or intolerance to:
    - At least one biologic ADT
    - At least one oral ADT
    - At least one biologic and one oral ADT

    will be tabulated.
  - The number of participants with primary nonresponse, secondary nonresponse, or intolerance to
    - At least one anti-TNF (regardless of other ADTs)
    - Vedolizumab (regardless of other ADTs)
    - Ozanimod (regardless of other ADTs)

- At least one JAK inhibitor (regardless of other ADTs)

will be tabulated

- The number of
  - Participants without biologic (i.e., tumor necrosis factor alpha [TNFα] antagonists and vedolizumab) failure (naïve, bio-experienced [but not documented failure])
  - Participants who are not ADT-IR (naïve, ADT-experienced [but not documented failure])

will be tabulated

6.6. Appendix 6 Medical History

Not applicable

6.7. Appendix 7 Intervention Compliance

Compliance will be summarized descriptively based on the **Randomized Analysis Set**.

Additional listings of participants will be provided for the following categories:

- Participants who were randomized but did not receive study intervention
- Participants who were unblinded prior to Week 24 or Week48

6.8. Appendix 8 Medications of Special Interest

Not applicable

6.9. Appendix 9 Laboratory Toxicity Grading

The grading scale used for lab assessments is based on 'Common Terminology Criteria for Adverse Events (CTCAE) v5.0'. Toxicity grades are based on the laboratory result and do not take into account the clinical component, if applicable.

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

Table 15: Toxicity Grading Scale

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 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	–	–<130-120	<120

Table 16: Liver Function Tests

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

6.10. Appendix 10 Mayo Score

Table 17 shows the Mayo scoring system[1),10)].

CCI

6.11. Appendix 11 Geboes Score

Table 18 shows the Geboes[2)] grading criteria.

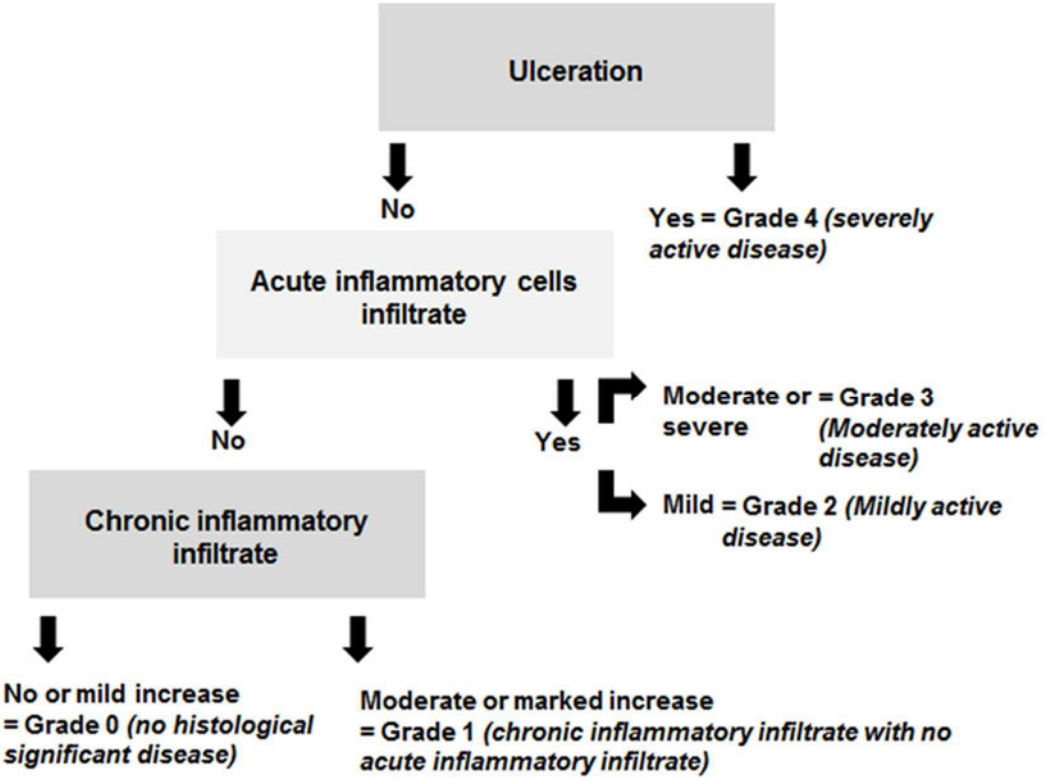
Table 18: Geboes score

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	No increase
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 distruction
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

6.12. Appendix 12 Nancy Histological Index

Figure 3 shows the grading for the Nancy histological index[5)].

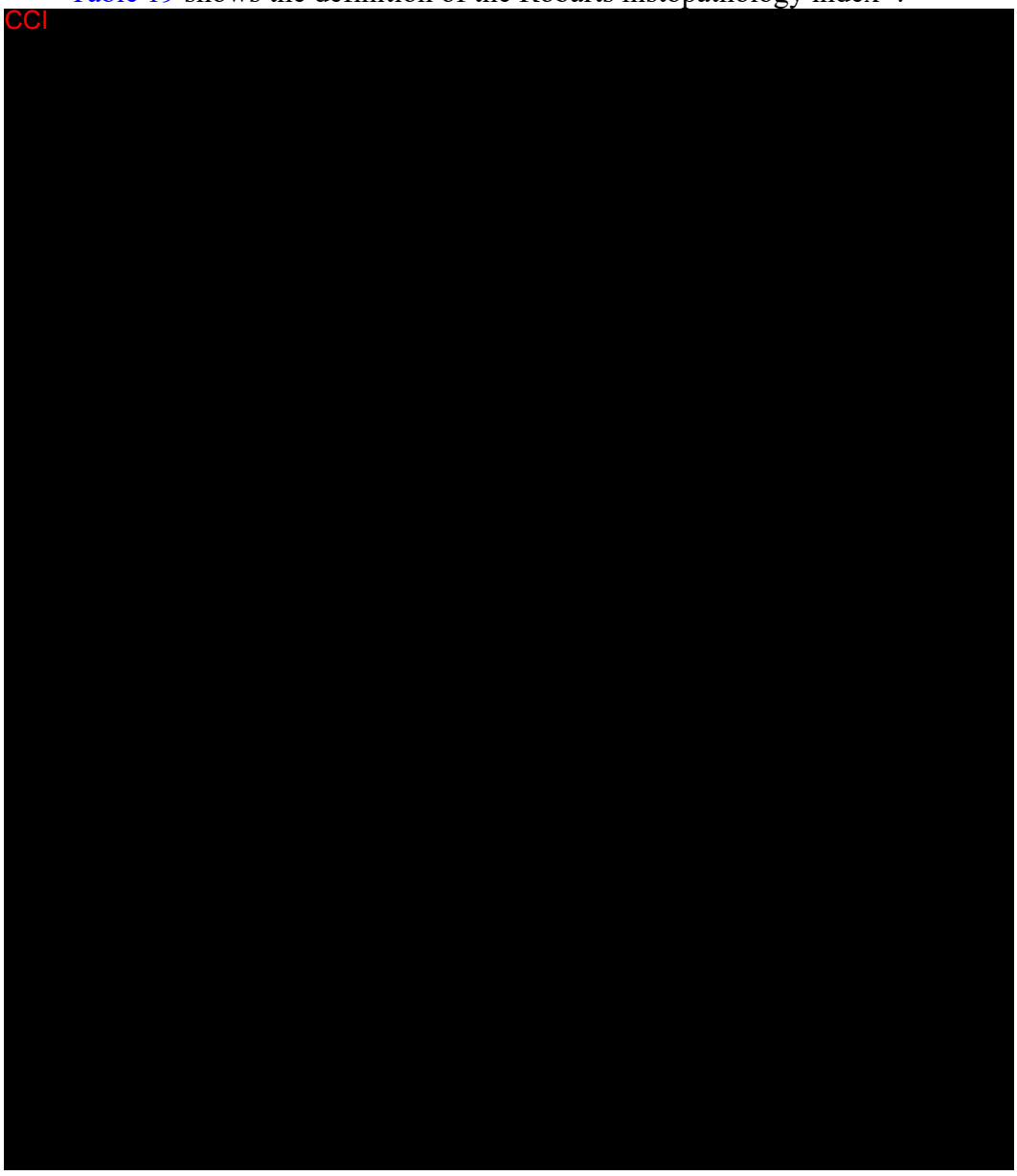
Figure 3: Nancy histological index



6.13. Appendix 13 Robarts Histopathology Index

Table 19 shows the definition of the Robarts histopathology index[7)].

CCI



6.14. Appendix 14 Detailed Prohibited Changes in UC Medications Rules

Participants had a prohibited change in UC medication described below are considered to have experienced ICE 2 as defined in Table 6:

- Initiation of restricted medications (rectal 5-ASA compounds, parenteral or rectal corticosteroids, including budesonide and beclomethasone dipropionate)
- Initiation of the following prohibited concomitant therapies
  - Immunomodulatory agents other than 6-MP, AZA, or MTX (eg, 6-thioguanine, cyclosporine, mycophenolate mofetil, tacrolimus, sirolimus, ozanimod, tofacitinib, and other JAK inhibitors).
  - Immunomodulatory biologic agents (eg, TNFα antagonists, ustekinumab, vedolizumab, abatacept, anakinra).
  - Experimental IBD medications (eg, upadacitinib, etrolizumab, brazikumab, mirikizumab, risankizumab) or other investigational medications/therapies.
  - Thalidomide or related agents
- Initiation of oral corticosteroids (including budesonide and beclomethasone dipropionate) due to worsening of disease.
- Increase in the dose of oral corticosteroids (excluding budesonide and beclomethasone dipropionate) > 5 mg/day (prednisone equivalent) above the baseline dose due to worsening of disease:
  - Any number of days with increase from Week 0 through Week 12
  - 3 days or more with increase from Week 12 through Week 24
  - 7 days or more with increase from Week 24 through Week 48
- Increase in the dose of oral budesonide > 3 mg/day above the baseline dose due to worsening of disease:
  - Any number of days with increase from Week 0 through Week 12
  - 3 days or more with increase from Week 12 through Week 24
  - 7 days or more with increase from Week 24 through Week 48
- Increase in the dose of oral beclomethasone dipropionate > 5 mg/day above the baseline dose due to worsening of disease:
  - Any number of days with increase from Week 0 through Week 12
  - 3 days or more with increase from Week 12 through Week 24
  - 7 days or more with increase from Week 24 through Week 48
- Any switch among oral budesonide, oral beclomethasone dipropionate or other oral corticosteroids (excluding prednisone equivalent changes) due to worsening of disease.
- Initiation of oral 5-ASA compounds due to worsening of disease.
- Increase above baseline in the dosage of oral 5-ASA compounds due to worsening of disease.

- Change from one oral 5-ASA compound to another 5-ASA compound due to worsening of disease.
- Initiation of 6-MP/AZA/MTX due to worsening of disease.
- Increase above baseline in the dosage of 6-MP/AZA/MTX due to worsening of disease.
- Any switch between 6-MP/AZA and MTX due to worsening of disease.

6.15. Appendix 15 UC Patient-Reported Outcomes Signs and Symptoms

CCI

[redacted]

CCI

[redacted]

CCI

[redacted]

6.15.2. UC-PRO/SS Questionnaire

CCI

[redacted]

CCI

[redacted]

CCI

[redacted]

CCI

[redacted]

7. REFERENCES

1) Baron JH, Connell AM, Lennard-Jones JE. Variation between observers in describing mucosal appearances in proctocolitis. Brit Med J. 1964;1:89-92.

2) 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.

3) Higgins PD, Schwartz M, Mapili J, Krokos I, Leung J, Zimmermann EM. Patient defined dichotomous end points for remission and clinical improvement in ulcerative colitis. Gut. 2005;54(6):782-788.

4) Irvine EJ, Feagan B, Rochon J, et al. Quality of life: a valid and reliable measure of therapeutic efficacy in the treatment of inflammatory bowel disease. Canadian Crohn's Relapse Prevention Trial Study Group. Gastroenterology. 1994;106(2):287-296.

5) Little, Roderick JA, and Donald B. Rubin. Statistical analysis with missing data. Vol. 793. John Wiley & Sons, 2019.

6) Marchal-Bressenot A, Salleron J, Boulagnon-Rombi C, Bastien C, Cahn V, Cadiot G, et al. Development and validation of the Nancy histological index for UC. Gut 2017;66(1):43-9.

7) Mosli MH, Feagan BG, Zou G, Sandborn WJ, D'Haens G, Khanna R, et al. Development and validation of a histological index for UC. Gut 2017;66(1):50-8.

8) 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.

9) 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.

10) Truelove SC, Witts LJ. Cortisone in ulcerative colitis: final report on a therapeutic trial. Brit Med J. 1955;2:1041-1048.