Janssen Research & Development

Statistical Analysis Plan

A Phase 2a Randomized, Double-blind, Active-controlled, Parallel-group, Multicenter, Proof-of-concept Clinical Study to Evaluate the Efficacy and Safety of Combination Therapy with Guselkumab and Golimumab in Participants with Moderately to Severely Active Ulcerative Colitis

VEGA

Protocol CNTO1959UCO2002; Phase 2a

CNTO1959 (guselkumab); CNTO148 (golimumab)

Status:	Approved
Date:	24 February 2021
Prepared by:	Janssen Research & Development, LLC
Document No.:	EDMS-RIM-346041

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

TABLE OF CONTENTS

LIST OF TABLES

ABBREVIATIONS

5-ASA	5-aminosalicyclic acid
6-MP	6-mercaptopurine
AE	adverse event
ALT	alanine aminotransferase
ANCOVA	analysis of covariance
AST	aspartate aminotransferase
AZA	azathioprine
BSFS	bristol stool form scale
BMI	body mass index
CI	confidence interval
CMH	cochran-mantel-haenszel
CRP	c-reactive protein
CV	coefficient of variation
DBL	database lock
DMC	data monitoring committee
eCRF	electronic case report form
FAS	full analysis set
FCS	full conditional specification
GCP	good clinical practice
HRQOL	health-related quality of life
IAP	interim analysis plan
IBD	inflammatory bowel disease
IBDQ	inflammatory bowel disease questionnaire
IEC	independent ethics committee
IQ	interquartile
IRB	institutional review board
IV	intravenous
IWRS	interactive web response system
LLOQ	lower limit of quantification
LSmeans	least squares means
MedDRA	medical dictionary for regulatory activities
MTX	methotrexate
MI	multiple imputation
MMRM	mixed-effect model repeated measure
MAR	missing data random
NAbs	neutralizing antibodies
NCI-CTCAE	national cancer institute common terminology criteria for adverse events
NHI	Nancy Histologic Index
NONMEM	nonlinear mixed-effects modeling
NA	North America
PD	pharmacodynamic
PGIC	patient's global impression of change (of severity of Ulcerative Colitis)
PGA	physician's global assessment
PK	pharmacokinetic
POC	proof-of-concept
PRO	patient-reported outcomes
PROMIS	patient-reported outcomes measurement information system
q4w	every 4 weeks
q8w	every 8 weeks
RBS	rectal bleeding subscore
RHI	Roberts Histologic Index
RNA	ribonucleic acid
SAE	serious adverse event
SAP	Statistical Analysis Plan
SC	Subcutaneous
SD	standard deviation

SI	système internationale
SoA	schedule of activities
SUSAR	suspected unexpected serious adverse reaction
$t_{1/2}$	half-life
TF	treatment failure
TNF	tumor necrosis factor
UC	ulcerative colitis
UCEIS	Ulcerative Colitis endoscopic index of severity
ULN	upper limit of normal
US	United States
WBC	white blood cell

1. INTRODUCTION

The protocol CNTO1959UCO2002 is a Phase 2a clinical trial designed to evaluate the efficacy and safety of combination therapy with guselkumab and golimumab in participants with moderately to severely active ulcerative colitis (UC). This study is comprised of 2 distinct phases: a 12-week **combination comparison phase** followed by a 26-week **monotherapy phase**. This Statistical Analysis Plan (SAP) contains definitions of analysis sets, derived variables, and statistical methods for all planned analyses for protocol CNTO1959UCO2002.

1.1. Trial Objectives

Primary Objectives

Combination Comparison Phase

- To evaluate the clinical efficacy of combination therapy with guselkumab and golimumab in participants with moderately to severely active UC.
- To evaluate the safety of combination therapy with guselkumab and golimumab in participants with moderately to severely active UC.

Secondary Objectives

Combination Comparison Phase

- To evaluate the effect of combination therapy with guselkumab and golimumab on endoscopic improvement.
- To evaluate the impact of combination therapy with guselkumab and golimumab on disease-specific health-related quality of life (HRQoL), including fatigue.
- To evaluate the efficacy of combination therapy with guselkumab and golimumab by negative response signature status at baseline.
- To evaluate the pharmacokinetic (PK), immunogenicity, and pharmacodynamics (PD) of combination therapy with guselkumab and golimumab, including changes in C-reactive protein (CRP), fecal calprotectin, and other PD biomarkers.

Monotherapy Phase

- To evaluate the clinical efficacy of combination therapy followed by guselkumab monotherapy.
- To evaluate the safety of combination therapy followed by guselkumab monotherapy.
- To evaluate the effect of combination therapy followed by guselkumab monotherapy on endoscopic improvement.
- To evaluate the impact of combination therapy followed by guselkumab monotherapy on disease specific HRQoL, including fatigue.

- To evaluate the efficacy of combination therapy followed by guselkumab monotherapy by negative response signature status at baseline.
- To evaluate the PK, immunogenicity, and PD of combination therapy followed by guselkumab monotherapy, including changes in CRP, fecal calprotectin, and other PD biomarkers.

Exploratory Objectives

- To explore the effect of combination therapy on patient-reported outcome (PRO) instruments (e.g., Bristol Stool Form Scale [BSFS] and Patient's Global Impression of Change [PGIC] of Severity of UC).

1.2. Trial Design

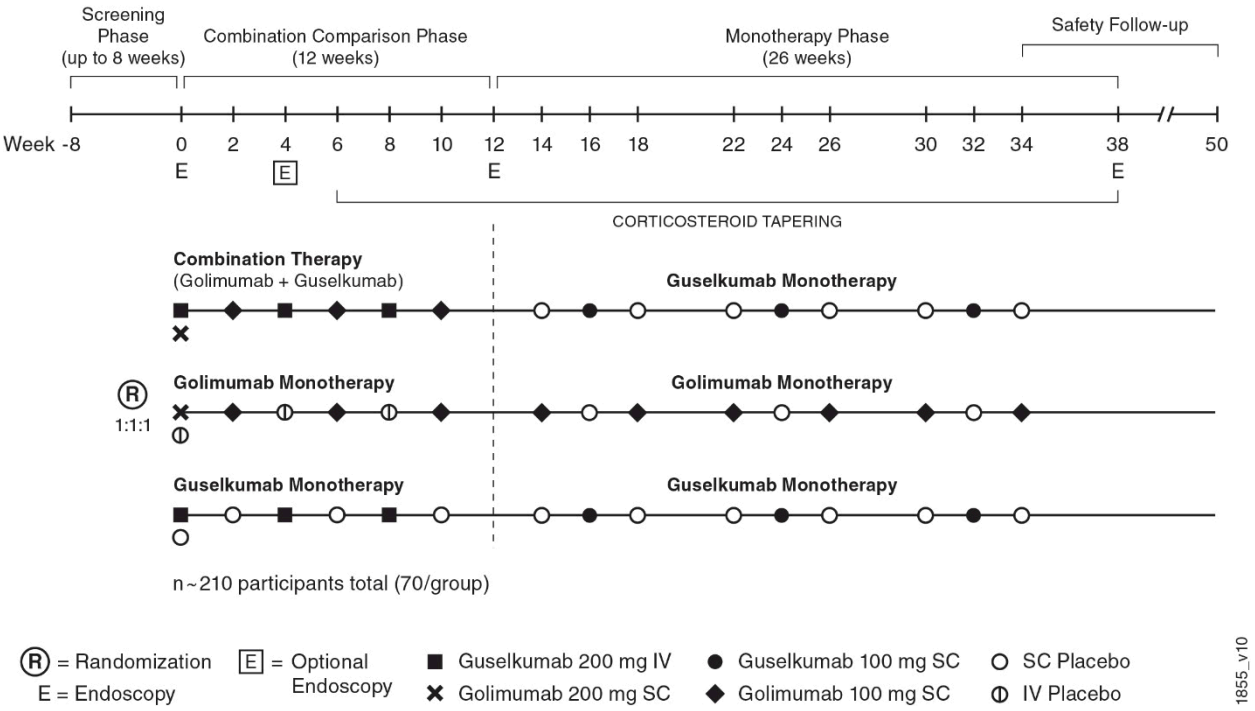
This is a Phase 2a, randomized, double-blind, active-controlled, parallel-group, multicenter, interventional proof-of-concept (POC) clinical study designed to evaluate the efficacy and safety of combination therapy with guselkumab and golimumab in adults with moderately to severely active UC. The target population is men or women 18 to 65 years old with moderately to severely active UC, as defined by a Mayo score of 6 to 12, inclusive, at baseline, including an endoscopy subscore ≥ 2 as obtained during the central review of the video endoscopy. Participants must be naïve to tumor necrosis factor (TNF) antagonists and have failed or not tolerated conventional therapy with oral or intravenous (IV) corticosteroids or immunomodulators (6-mercaptopurine [6-MP] or azathioprine [AZA]).

This study will consist of 2 distinct phases: a 12-week **combination comparison phase** followed by a 26-week **monotherapy phase**. At Week 0, a target of 210 participants will be randomized in a 1:1:1 ratio to either combination therapy with guselkumab and golimumab, guselkumab monotherapy, or golimumab monotherapy, stratified by corticosteroid use at baseline (yes, no). Participants randomized to combination therapy will receive guselkumab monotherapy after Week 12. Participants randomized to a monotherapy group will continue on their originally randomized monotherapy after Week 12. The combination therapy arm will employ the same dose regimens of guselkumab and golimumab being used in the respective monotherapy intervention groups to facilitate scientific interpretation of the results. The following is a description of the 3 intervention groups:

- **Combination therapy (N = 70)**: guselkumab 200 mg IV and golimumab 200 mg subcutaneous (SC) at Week 0; golimumab 100 mg SC at Weeks 2, 6, and 10; guselkumab 200 mg IV at Weeks 4 and 8 followed by guselkumab 100 mg SC q8w
- **Guselkumab monotherapy (N = 70)**: guselkumab 200 mg IV at Weeks 0, 4, and 8 followed by guselkumab 100 mg SC q8w
- **Golimumab monotherapy (N = 70)**: golimumab 200 mg SC injection at Week 0, followed by golimumab 100 mg at Week 2 and then golimumab 100 mg every 4 weeks (q4w)

In addition, placebo administrations (IV or SC) will be given, as appropriate, to maintain blinding throughout the duration of the study. Refer to Protocol Section 6 for additional details. A diagram of the study design is provided in Figure 1.

Figure 1: Schematic Overview of Study CNTO1959UCO2002



Immunomodulators (6-MP, AZA, and methotrexate [MTX]) must be discontinued for at least 2 weeks before the first dose of study intervention. For participants who are receiving oral corticosteroids at baseline, the investigator must begin tapering the daily dose of corticosteroids at Week 6. The use of concomitant and prohibited therapies is described in Protocol Section 6.5. In general, doses of concomitant therapies for UC should remain stable through Week 38 (except for mandatory oral corticosteroid tapering beginning at Week 6), and concomitant therapies for UC should not be initiated unless considered medically necessary by the investigator. Initiation of prohibited therapies will result in discontinuation of study intervention.

All participants will be evaluated for clinical worsening of UC throughout the study. Endoscopy with central read is planned for screening/baseline, Week 12, and Week 38. Consenting participants will have an additional endoscopy at Week 4, which will also be assessed by a central reader. Efficacy, PK and PD parameters, biomarkers, and safety will be assessed according to the Schedule of Activities (SoA) (Protocol Section 1.3). A pharmacogenomic blood sample will be collected from participants who consent to this component of the protocol (where local regulations permit). Participation in pharmacogenomic research is optional.

Overall participant duration will be up to 58 weeks total (screening: up to 8 weeks; treatment duration: 38 weeks [12 weeks for the combination comparison phase; 26 weeks for the monotherapy phase]; safety follow-up: approximately 16 weeks after the last administration of study intervention at Week 34).

An interim analysis (IA) was performed in April 2020 to inform future clinical development planning after the first 126 randomized participants (i.e., 60% of the planned sample size) who received at least one administration of study intervention (complete or partial) completed the Week 12 assessments or terminated study participation prior to the Week 12 visit. The details regarding this IA are provided in the Interim Analysis Plan (IAP).

There are 3 planned database locks (DBLs) for this study, respectively, at Week 12, at Week 38, and at end of the study. The end of the study is defined as when the last participant completes his or her final safety follow-up visit. The first DBL will occur when all participants randomized in this study have either completed the Week 12 assessments or terminated study participation prior to the Week 12 visit (referred to as Week 12 DBL hereafter). The second DBL will occur when all participants randomized in this study have either completed the Week 38 assessments or terminated study participation prior to the Week 38 visit (referred to as Week 38 DBL hereafter). The third DBL will occur at end of study when all participants randomized in this study have either completed their final safety follow-up visit or have terminated study participation [referred to as Final Safety DBL hereafter].

The primary endpoint of this study is clinical response at Week 12 (refer to Section 5.3 for endpoint definition and analyses). The major secondary endpoint is clinical remission at Week 12 (Section 5.4 – definitions and analyses methods). The primary and major secondary endpoints will be analyzed at the Week 12 DBL.

An external independent data monitoring committee (DMC) will be commissioned for this study to monitor the safety of the study in an unblinded fashion on a regular basis and whenever deemed necessary. Refer to the DMC charter for more details on the DMC.

1.3. Statistical Hypotheses for Trial Objectives

The primary endpoint of this study is clinical response at Week 12 (refer to Section 5.3 for endpoint definition and analyses).

The primary hypothesis is that guselkumab and golimumab combination therapy is superior to both monotherapy arms as assessed by the proportion of participants achieving clinical response at Week 12.

The study will be considered positive if both comparisons of guselkumab and golimumab combination therapy versus guselkumab monotherapy therapy and guselkumab and golimumab combination therapy versus golimumab monotherapy therapy achieve the statistical significance at the 2-sided significance level of 0.2 for the primary endpoint.

1.4. Sample Size Justification

A sample size of 210 participants (70 per intervention group) was determined by the power to detect a significant difference in the proportion of participants in clinical response at Week 12 (primary endpoint) between the combination therapy and both monotherapies using a 1-sided chi-square test with 0.1 significance level for each comparison. The study is sized such that the

combination therapy has approximately 80% power (based on simulations) to achieve both comparisons to monotherapy for the primary endpoint.

The assumptions for the sample size calculations were based on data from the past golimumab programs conducted by the sponsor in participants with UC who were naïve to biologic therapy. The proportion of participants in clinical response at Week 12 is expected to be 55% for golimumab compared to 51% observed at Week 6 in study C0524T17 (PURSUIT), as the PURSUIT program suggested that the response rate might increase slightly with further treatment. This was also confirmed in study CNTO148UCO2001 (PROgECT)[13]. The tested dose of guselkumab is assumed to have the same clinical response rate as golimumab at Week 12. The proportion of participants in clinical response at Week 12 is assumed to be 75% for the combination therapy, which is based on the additive effect from both monotherapies (20% improvement from each monotherapy relative to a historical placebo response of 35%). The different combinations of clinical response rate assumptions and associated power are displayed in Table 1.

Table 1: Power to Detect a Treatment Effect of the Combination Therapy Versus Both Guselkumab Monotherapy and Golimumab Monotherapy Based on the Proportion of Participants Achieving Clinical Response at Week 12

Golimumab (n=70)	Guselkumab (n=70)	Combination Therapy (n=70)	Power
50%	50%	70%	79%
50%	55%	70%	66%
55%	60%	75%	69%
55%	**55%**	**75%**	**81%**
55%	60%	80%	90%
60%	60%	80%	85%

1.5. Randomization and Blinding

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

Intervention Allocation

Central randomization will be implemented in this study. Participants will be randomly assigned to 1 of 3 intervention groups (1:1:1 ratio), based on a computer-generated randomization schedule prepared before the study by or under the supervision of the sponsor. The randomization will be balanced by using randomly permuted blocks and will be stratified by corticosteroid use at baseline (yes, no). The interactive web response system (IWRS) will assign a unique treatment code, which will dictate the treatment assignment and matching study intervention kit(s) for the participant.

The requestor must use his or her own user identification and personal identification number when contacting the IWRS and will then be given the relevant participant details to uniquely identify the participant.

Blinding

To maintain the study blind, the study intervention container will have a label containing the study name, study intervention number, and reference number. The study intervention number will be entered in the electronic case report form (eCRF) when the study intervention is dispensed. Each active study intervention and its matching placebo will be identical in appearance.

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

The post-baseline results of CRP and fecal calprotectin tests performed by the central laboratory will be blinded to the study sites. If a study site requests these data, it will be provided to them after the analyses based on the Final Safety DBL have been completed.

Treatment assignment blinding will be maintained for study sites, site monitors, and participants until the analyses based on the Final Safety DBL are completed. The full sponsor unblinding will occur after the Week 38 DBL. A limited number of sponsor personnel were unblinded to the IA data (only those randomized participants who were included in the IA; the treatment assignment for the remaining participants remained blinded) for data analyses and review. The sponsor personnel who were unblinded to the IA data were documented in the IA unblinding plan. In addition, a limited number of sponsor personnel will become unblinded at the Week 12 DBL for data analyses and review. Identification of sponsor personnel who will have access to the unblinded data and at what level (treatment group-level or participant-level) will be documented in the unblinding plan before unblinding occurs at the Week 12 DBL.

The investigator will not be provided with randomization codes. The codes will be maintained within the IWRS, which has the functionality to allow the investigator to break the blind for an individual participant.

Under normal circumstances, the investigator blind should not be broken unless specific emergency treatment/course of action would be dictated by knowing the treatment status of the participant. In such cases, the investigator may in an emergency determine the identity of the treatment via the IWRS. It is recommended that the investigator contact the sponsor or its designee, if possible, to discuss the particular situation before breaking the blind. Telephone contact with the sponsor or its designee will be available 24 hours per day, 7 days per week. If the blind is broken, the sponsor must be informed as soon as possible. The date and reason for the unblinding must be documented in the appropriate section of the eCRF and in the source document. The documentation received from the IWRS indicating the code break must be retained with the participant's source documents in a secure manner.

Additionally, a given participant's treatment assignment may be unblinded to the sponsor, the Independent Ethics Committee/Institutional Review Board (IEC/IRB), and site personnel to fulfill regulatory reporting requirements for suspected unexpected serious adverse reactions (SUSARs). If a participant is unblinded for this reason, the information must be entered in the appropriate section of the eCRF and in the participant's source documents.

Participants who have had their treatment assignment unblinded by the investigator will not be eligible to receive further study intervention but should complete evaluations specified in the appropriate SoA (Protocol Section 1.3) for participants who discontinue study intervention.

A separate code break procedure will be available for use by the J&J Global Medical Safety group to allow for unblinding of individual participants to comply with specific requests from regulatory or health authorities.

2. GENERAL ANALYSIS DEFINITIONS

2.1. Visit Windows

Unless otherwise specified, actual scheduled visits will be used for over time summaries and listings with no visit windows applied.

2.2. Analysis Sets

2.2.1. Randomized Analysis Set

The randomized analysis set includes all participants who are randomized in the study. Participants in the Randomized Analysis Set will be analyzed according to the randomized study intervention regardless of the study intervention they actually received.

2.2.2. Efficacy Analysis Set(s)

Unless otherwise specified, all efficacy analyses will be based on the Full Analysis Set (FAS), which includes all randomized participants who receive at least 1 (partial or complete) dose of study intervention. Participants in FAS will be analyzed according to the randomized study intervention regardless of the study intervention they actually received.

2.2.3. Safety Analysis Set

The safety analysis set includes all randomized participants who receive at least 1 (partial or complete) dose of study intervention. Participants in the Safety Analysis Set will be analyzed according to the study intervention they actually received.

2.2.4. Pharmacokinetics Analysis Set

The PK analysis set includes all randomized participants who receive at least 1 (complete) dose of study intervention (golimumab or guselkumab) and have at least 1 valid blood sample drawn for PK analysis after their first dose of study intervention (golimumab or guselkumab). Participants in the Pharmacokinetics Analysis Set will be analyzed according to the study intervention they actually received.

2.2.5. Immunogenicity Analysis Set

The immunogenicity analysis set is defined as all participants who receive at least 1 (partial or complete) dose of guselkumab and/or golimumab and have appropriate samples for detection of antibodies to guselkumab and/or to golimumab (ie, participants with at least 1 sample obtained after their first dose of guselkumab or golimumab, respectively). Participants in the Immunogenicity Analysis Set will be analyzed according to the study intervention they actually received.

2.3. Definition of Subgroups

The primary endpoint will be evaluated for subgroups based on demographics and baseline UC clinical disease characteristics, UC-related concomitant medication usage, and UC-related medication history.

Demographics

- Gender (male, female)
- Race (Caucasian, non-Caucasian)
- Baseline age (≤ median age, > median age)
- Baseline body weight (≤ 1st quartile, >1st quartile and ≤ 2nd quartile, >2nd quartile and ≤ 3rd quartile, >3rd quartile)
- Baseline body mass index (BMI) (≤ 1st quartile, >1st quartile and ≤ 2nd quartile, >2nd quartile and ≤ 3rd quartile, >3rd quartile)
- Tobacco use status (non-user, prior user, current user)
- Region
  - Eastern Europe: Poland, Russia, Ukraine
  - Latin America: Argentina, Mexico, Brazil
  - Rest of World: United States, Germany, Australia

Baseline UC clinical disease characteristics

- UC disease duration (≤ 5 years, > 5 years to ≤ 15 years, > 15 years)
- Extent of disease (limited, extensive)
- Severity of UC disease (moderately active: 6 ≤ Mayo score ≤ 10, severely active: Mayo score >10)
- Severity of endoscopic disease (moderate: endoscopic subscore = 2, severe: endoscopic subscore = 3)
- CRP (≤ 3 mg/L, > 3 mg/L)
- 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-aminosalicylic acid [5-ASA] compounds (receiving, not receiving)
- Oral corticosteroids including budesonide and beclomethasone dipropionate (receiving, not receiving)
- Participants taking conventional immunomodulators (6-MP/AZA/MTX) during Screening but discontinuing it prior to Week 0
- Participants taking Vedolizumab during Screening but discontinuing it prior to Week 0

UC-related medication history

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

2.4. Study Day and Relative Day

Study Day 1 refers to the date of the first study intervention administration. All efficacy and safety assessments at all visits will be assigned a day relative to this date.

Study day for a visit is defined as:

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

There is no 'Study Day 0'.

2.5. Baseline

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

2.6. Imputation Rules for Missing AE Date/Time of Onset/Resolution

Partial adverse event (AE) onset dates will be imputed as follows:

- If the onset date of an adverse event is missing day only, it will be set to:
  - First day of the month that the AE occurred, if month/year of the onset of AE is different than the month/year of the date of the first study intervention administration
  - The day of the first study intervention administration, if the month/year of the onset of AE is the same as month/year of the first study intervention administration and month/year of the AE resolution date is later
  - The day of the first study intervention administration or day of AE resolution date, whichever is earliest, if month/year of the onset of AE and month/year of the first study intervention administration date and month/year of the AE resolution date are same
- If the onset date of an adverse event is missing both day and month, it will be set to the earliest of:
  - January 1 of the year of onset, as long as the year is the same or after the year of the first study intervention administration
  - Month and day of the first study intervention administration, if this date is the same year that the AE occurred
  - Last day of the year if the year of the AE onset is prior to the year of the first study intervention administration
- Completely missing onset dates will not be imputed.

Partial AE resolution dates not marked as ongoing will be imputed as follows:

- If the resolution date of an adverse event is missing day only, it will be set to the earliest of the last day of the month of occurrence of resolution or the day of the date of death, if the death occurred in that month.
- If the resolution date of an adverse event is missing both day and month, it will be set to the earliest of December 31 of the year or the day and month of the date of death, if the death occurred in that year.
- Completely missing resolution dates will not be imputed.

AE onset/resolution dates with missing times will be imputed as follows:

- A missing time of onset of an adverse event will be set to:
  - 00:01 as long as the onset date is after the study intervention start date
  - The time of the study intervention start date if this is the same day the AE occurred.
- The missing time of resolution of an adverse event will be set to 23:59.

If a missing time is associated with a partial or missing date, the date will be imputed first prior to imputing the time.

3. INTERIM ANALYSIS AND DATA MONITORING COMMITTEE REVIEW

For the purpose of future clinical development planning, an IA was performed in April 2020 after the first 126 randomized participants (i.e., 60% of the planned sample size) completed their Week 12 visit or terminated their study participation before Week 12.

An external DMC has been established and are meeting periodically to review interim unblinded safety data to ensure the continuing safety of the participants enrolled in the study. The DMC consists of 2 physicians and a statistician. The DMC responsibilities, authorities, and procedures are documented in a separate DMC charter.

4. SUBJECT INFORMATION

The number of participants in each analysis set will be summarized by treatment group and overall for this study. In addition, the distribution of participants by region, country and site ID will be presented.

Descriptive statistics (mean, standard deviation (SD), median, interquartile (IQ) range, minimum and maximum) will be provided for continuous variables. Frequency distributions will be provided for categorical variables. No formal statistical analyses for treatment comparisons will be performed.

4.1. Demographics and Baseline Characteristics

Table 2 presents a list of the demographic and baseline characteristics variables that will be summarized by treatment group and overall for the FAS (Section 2.2.2).

Table 2: Demographic and Baseline Characteristics 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)	
BMI (kg/m^2)	
UC Disease Duration (years)	
Mayo score	
Partial Mayo score	
Modified Mayo score	
Mayo subscores (rectal bleeding subscore, stool frequency subscore)	
CRP (mg/L)	
Fecal Calprotectin (mg/kg)	
Categorical Variables	
Sex (male, female)	Frequency distribution with the number and percentage of participants in each category.
Race[a] (American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or other Pacific Islander, White, Not Reported, Multiple)	
Ethnicity (Hispanic or Latino, not Hispanic or Latino, Not Reported)	
Region[b] (Eastern Europe, Latin America, Rest of World)	
UC Disease Duration (≤ 5 years, > 5 and ≤ 15 years, > 15 years)	
Mayo Endoscopy Subscore (moderate: subscore = 2, severe: subscore = 3)	
Extent of disease (limited, extensive)	
Severity of UC disease (moderately active: 6 ≤ Mayo score ≤ 10, severely active: Mayo score > 10)	
Abnormal CRP (> 3 mg/L)	
Abnormal fecal calprotectin (> 250 mg/kg)	
Tobacco use status (non-user, prior user, current user)	
(randomization stratification factor) Corticosteroid use (yes, no)	

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

[b] Eastern Europe: Poland, Russia, Ukraine; Latin America: Argentina, Mexico, Brazil; Rest of World: United States, Germany, Australia

4.2. Disposition Information

Participants who discontinued study intervention early, along with the reasons for discontinuation of study intervention (including those due to COVID-19 related events), will be summarized by treatment group and overall based on the FAS (Section 2.2.2). In addition, participants who terminated study participation early, along with the reasons for termination of study participation (including those due to COVID-19 related events), will be summarized by treatment group and overall based on the FAS.

A list of participants who discontinued study intervention early (including those due to COVID-19 related events) and a list of participants who terminated study participation early (including those due to COVID-19 related events) will also be provided.

4.3. Treatment Compliance

The number of participants receiving each scheduled study intervention will be summarized by treatment group for the FAS (Section 2.2.2). Compliance with randomized study intervention versus actual received study intervention will be presented in a summary table for the FAS.

In addition, a listing of participants who were assigned treatment but were never treated and a listing of participants who were unblinded during the study will be provided.

4.4. Extent of Exposure

The number and percentage of participants who receive study intervention will be summarized by treatment group for the Safety Analysis Set (Section 2.2.3). The cumulative dose of guselkumab and golimumab received will be summarized by treatment group based on the Safety Analysis Set.

In addition, the number of administrations of study intervention and average duration of follow-up in weeks will be summarized by treatment group for all participants in the Safety Analysis Set as part of the AE tables.

The distribution of participants by study intervention lot will also be summarized by treatment group for the Safety Analysis Set.

4.5. Protocol Deviations

Participants with a major protocol deviation will be summarized by treatment group based on the FAS (Section 2.2.2). Major protocol deviations will be categorized as study intervention administration deviations, study entry criteria not met, prohibited concomitant medications deviations, withdrawal criteria met but not withdrawn, or other (including a sub-category for COVID-19 related deviations).

In addition, participants who did not meet study entry criteria will be further summarized by category (UC disease criteria, Medication criteria, Laboratory criteria, Medical history criteria, and other) based on the FAS. Further, participants who had study intervention administration

deviations will be further summarized by category identified prior to unblinding (e.g., received wrong study intervention administration, missed study intervention administration, received study intervention administration outside of protocol specified window).

A listing of participants who had a major protocol deviation (including COVID-19 related deviations), a more detailed listing of participants who did not meet study entry criteria (by category), and a more detailed listing of participants who had a study intervention administration deviations will be provided. A listing of participants with minor protocol deviations related to COVID-19 will also be provided.

4.6. Prior and Concomitant Medications

Summaries of UC medication history (participants who took medications for UC and their length of exposure prior to the study), UC-related non-biologic medication history (ie, history of response to or intolerance of corticosteroids and immunomodulators [ie, 6-MP/AZA]), UC-related biologic medication history (ie, vedolizumab), and prior use of JAK inhibitors (ie, tofacitinib) will be provided by treatment group based on the FAS (Section 2.2.2).

UC-specific baseline medications (ie, 5-ASA and corticosteroids) will be summarized by treatment group based on the FAS. Baseline medications are defined as any therapy used at baseline prior to the first dose of study intervention.

UC-specific concomitant medications that are permitted and taken during the trial (i.e., oral 5-ASA and oral corticosteroids) will be summarized by treatment group based on the FAS. Concomitant medications are defined as any therapy used on or after the first dose of study intervention, including those that started before and are continued after the first dose of study intervention.

A listing of concomitant medications for COVID-19 infections will be provided.

5. EFFICACY

5.1. General Method of Analysis

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

Analyses suitable for categorical data (e.g., chi-square tests, Cochran-Mantel-Haenszel [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 treatment groups. In cases of rare events, Fisher's exact test will be used for treatment comparisons. Continuous response parameters will be compared using a Mixed-Effect Model Repeated Measure (MMRM) model unless otherwise specified. If the normality assumption is in question, an appropriate transformation may be implemented before fitting the MMRM model. In cases of small sample size, t-test will be used for treatment comparisons.

5.2. Analysis Specifications

5.2.1. Level of Significance

A 2-sided significance level of 0.2 (equivalent to the 1-sided 0.1 level specified in the protocol) will be used for all hypothesis testing and no adjustments for multiple comparisons will be made. Nominal p-values will be presented.

5.3. Primary Efficacy Endpoint

The primary efficacy endpoint is clinical response at Week 12.

5.3.1. Definition

Clinical response: A decrease from baseline in the Mayo score (Section 5.3.1.1) ≥ 30% and ≥ 3 points, with either a decrease in the rectal bleeding subscore (RBS) ≥ 1 or a RBS of 0 or 1.

5.3.1.1. Mayo Score, Partial Mayo Score, and Modified Mayo Score

The **Mayo score** was developed from the criteria of Truelove and Witts[14] for mild, moderate, and severe UC and from the criteria of Baron et al[2] for grading endoscopic appearance. The Mayo score consists of the following 4 subscores:

- Stool frequency
- Rectal bleeding
- Findings of endoscopy
- Physician's global assessment (PGA)

Each subscore is rated on a scale from 0 to 3, indicating normal to severe activity, as defined in Attachment 1.

The **Mayo score** is calculated as the sum of the 4 subscores (stool frequency, rectal bleeding, PGA, and endoscopy findings) and ranges from 0 to 12 points. A score of 3 to 5 points indicates mildly active disease, a score of 6 to 10 points indicates moderately active disease, and a score of 11 to 12 points indicates severely active disease.

The **partial Mayo score,** which is the Mayo score without taking into account the findings of endoscopy, is calculated as the sum of stool frequency, rectal bleeding, and PGA subscores, and may take on values from 0 to 9.

The **modified Mayo score**, which is the Mayo score without the PGA subscore, is calculated as the sum of the stool frequency, rectal bleeding, and endoscopy subscores, and may take on values from 0 to 9.

Due to the requirement for endoscopy findings for the evaluation of the Mayo score, it is not feasible to evaluate the Mayo score at each scheduled visit in the study. Therefore, the Mayo score will be evaluated at Weeks 0, 12 and 38, and the partial Mayo score will be evaluated at the other

study visits. The modified Mayo score can be derived from the Mayo score and is not a separate evaluation.

Mayo Rectal Bleeding and Stool Frequency Subscores

The eCRFs capture seven days of rectal bleeding data and the number of stools per day prior to each visit at which the Mayo score or partial Mayo score is collected. Data from 3 of these 7 days are used to calculate the Mayo rectal bleeding and stool frequency subscores; sites are instructed to check the boxes next to the 3 days which are used (see below for information on what days the sites are instructed to use).

The Mayo rectal bleeding subscore is calculated as the average rectal bleeding number for the three days based on the criteria in Attachment 1.

The Mayo stool frequency subscore is calculated as follows: The absolute stool number is the average of the daily stool number over the three days. At the screening visit, each person indicates the number of stools he/she passed in a 24-hour period when in remission or before his/her UC diagnosis. The stool frequency subscore will be calculated based on the criteria in Attachment 1 by subtracting the number of stools when in remission or prior to UC from the absolute stool number.

Instructions on which 3 days to use in the calculation of the Mayo rectal bleeding and stool frequency subscores: Sites are directed to use the most recent 3 consecutive days within the 7 days prior to the visit and are directed to exclude the following:

- The day medications were taken for constipation, diarrhea or irregularity
- The day of a procedure or preparation for procedure (e.g. enema, other laxatives, or clear liquid diet) that would affect stool frequency and/or blood content of the stool
- The 48 hours after the use of antimotility agents (e.g. diphenoxylate hydrochloride with atropine sulfate or loperamide)
- The 48 hours immediately following a colonoscopy

If 3 consecutive days are not available, the sites are instructed to choose 2 consecutive days and the closest nonconsecutive day. If 2 consecutive days are not available, then 3 nonconsecutive days closest to the visit should be chosen. If 3 days (within the 7 days prior to the indicated visit) that meet the criteria defined above are not available, then the absolute stool number, stool frequency subscore, and rectal bleeding subscore cannot be calculated and will be missing in the eCRF.

Mayo Endoscopy Subscore

The endoscopic findings will be based on the criteria of the Mayo endoscopy subscore described in Attachment 1. The endoscopic findings will be assessed by the investigator (ie, local endoscopist) during the endoscopy procedure and by the central reader reviewing a video of the endoscopy. The endoscopy may be either a colonoscopy or sigmoidoscopy. A full colonoscopy will replace a sigmoidoscopy if screening for polyps or dysplasia is required. The central reader

will also perform a friability assessment (yes, no) for each endoscopy, except for those with an assessed endoscopy subscore of 0, which requires that no friability is present.

Participant eligibility at baseline will be based on the **final reported endoscopic subscore** as determined by the following process:

- If the local endoscopist and the central reader agree on the endoscopic subscore, the agreed score will be the **final reported endoscopic subscore**.
- If there is a discrepancy between the local endoscopist and the central reader scores, the video endoscopy will be submitted to a second central reader (designated for adjudication) who is blinded to the scores of the local and the first central reader. From the scores of the 3 readers (i.e., local reader, central reader, second central reader designated for adjudication), the score with which 2 readers agree will be reported as the **final reported endoscopic subscore**. If the scores of the 3 readers (ie, local reader, central reader, second central reader designated for adjudication) are all different, then the median score of the 3 scores will be the **final reported endoscopic subscore**.

Further details are provided in the imaging charter.

Unless otherwise specified, the analysis of endpoints that include the Mayo endoscopy subscore will be based on the final reported endoscopic subscore. If the final reported endoscopic score is not available, the corresponding central endoscopy score (central read 1) will be used, if available. If the central endoscopy score (central read 1) is also missing, then the local endoscopy score will be used, if available. If the local endoscopy score is not available, then the endoscopy subscore for the analysis will be left missing.

Mayo Physician's Global Assessment Subscore

The physician's global assessment acknowledges the 3 other Mayo subscores, the patient's recall of abdominal discomfort and general sense of well-being, and other observations, such as physical findings and the patient's performance status.

5.3.2. Primary Estimand (Estimand 1)

The primary estimand (Estimand 1), i.e. a precise definition of the primary targeted treatment effect, is defined by the following 5 attributes:

Treatment by Week 12:

Experimental:

- **Combination Therapy**: Guselkumab 200 mg IV q4w (Weeks 0, 4, 8) **AND** Golimumab 200 mg SC (Week 0) and 100 mg SC q4w (Weeks 2, 6, 10)

Controls:

- **Golimumab Monotherapy**: Golimumab 200 mg SC (Week 0) and 100 mg SC q4w (Weeks 2, 6, 10)

- **Guselkumab Monotherapy**: Guselkumab 200 mg IV q4w (Weeks 0, 4, 8)

Population:

Participants 18 -65 years old with moderately to severely active UC, as defined by a Mayo score of 6 to 12, inclusive, at baseline, including an endoscopy subscore ≥ 2 as obtained during the central review of the video endoscopy.

Variable (Endpoint):

Clinical response at Week 12 (Section 5.3.1). Participants who have intercurrent events in categories 1-3 (defined below) prior to Week 12 visit will be considered not to have achieved clinical response at Week 12.

Intercurrent Events and Corresponding Strategies:

The following are the intercurrent events (ICEs) for this study:

1. An ostomy or colectomy (partial or total)
2. A protocol-prohibited change in concomitant UC medication(s) (described in Attachment 2)
3. Discontinuation of study intervention due to lack of efficacy or due to an AE of worsening of UC
4. Discontinuation of study intervention due to COVID-19 related reasons (excluding COVID-19 infection)
5. Discontinuation of study intervention for reasons other than lack of efficacy or an AE of worsening of UC or COVID-19-related reasons (this would include COVID-19 infections)

ICEs in categories 1-3 will be handled with the **composite strategy**. ICE category 4 will be handled by the **hypothetical strategy** (as if participants would have not experienced this intercurrent event), and ICE category 5 will be handled by the **treatment policy strategy**. This estimand acknowledges that having an ICE in categories 1-3 is an unfavorable outcome. Participants experiencing ICEs 1-3 will be considered not to have achieved clinical response at Week 12. For participants experiencing ICE 4, data at all visits after an ICE will be set to missing. For participants experiencing ICE 5, their observed clinical response status at Week 12 (if available) will be used. For participants experiencing multiple ICEs, ICEs in categories 1-3 will override ICEs 4 and 5.

Population-level summary:

The difference in proportion of participants achieving clinical response at Week 12 between combination therapy and each monotherapy.

5.3.3. Analysis Methods for the Primary Estimand (Estimand 1)

5.3.3.1. Main Estimator (Analysis) for the Primary Estimand

The primary endpoint, defined in Section 5.3.1, will be analyzed based on the Primary Estimand (Section 5.3.2). After accounting for the ICEs for the primary estimand (Estimand 1), participants who are missing any or all of the Mayo subscores that comprise the primary endpoint at Week 12 will be considered not to be in clinical response at Week 12 (i.e. nonresponder imputation).

In the primary analysis, data from all participants in the FAS (Section 2.2.2) will be analyzed according to the randomized study intervention regardless of the study intervention they actually received. The treatment difference between combination therapy versus each monotherapy will be tested using a 2-sided CMH chi-square test stratified by corticosteroid use at baseline (yes, no) at the 0.2 significance level. The magnitude of the treatment difference will be estimated by the difference in the proportion of participants achieving clinical response at Week 12 between the combination therapy group and each monotherapy group with a 2-sided 80% confidence interval (CI) calculated based on Wald statistic.

The study will be considered positive if the combination therapy group is significantly different from both monotherapy groups at the 2-sided significance level of 0.2 for the primary endpoint.

5.3.3.2. Subgroup Analyses

Subgroup analyses will be performed (if the number of participants within each subgroup level permits) based on demographic and baseline disease characteristics, baseline concomitant UC medication use, and history of UC-related medications specified in Section 2.3. Note that, for subgroup analyses, the analysis sets are the individual subgroups of the FAS. A forest plot will be produced for all subgroups. Odds ratio of the combination therapy group versus each monotherapy group and the associated 80% CI from a logistic regression model will be provided for each of subgroups. The logistic regression model will include treatment group and corticosteroid use at baseline (yes, no) as the factors. For subgroup analyses of corticosteroid use at baseline (yes, no), treatment group will be the only the factor in the model. The primary estimand (Estimand 1) will be used for these subgroup analyses.

5.3.3.3. Sensitivity Analyses

For the primary estimand (Estimand 1), after accounting for the ICEs, the proportion of participants with missing clinical response status at Week 12 is expected to be quite small. Despite the expectation of a very small percentage of missing data at Week 12, the following sensitivity analysis will be performed using a tipping point analysis with exhaustive scenarios.

The following method will be utilized to vary the imputation of clinical response status at Week 12 for missing data after the intercurrent event rules have been applied. For participants with missing data at Week 12, the clinical response status will be imputed in an increasing manner by participant level for each study intervention. Specifically, for each participant, a responder/non-responder status will be imputed starting with the scenario where all participants are non-responders up to the scenario where all participants are responders. This would include all possible

scenarios of responder status for all missing data. For each scenario, an analysis same as described above in Section 5.3.3.1 for the primary analysis for the primary estimand will be performed.

5.3.4. Supplementary Estimands

Two supplementary estimands (Estimands 2 and 3) are considered to support the primary estimand for the primary endpoint (Estimand 1).

5.3.4.1. Estimand 2

In this supplementary estimand for the primary endpoint (Estimand 2), all ICEs (Section 5.3.2) will be addressed by the **composite strategy**.

The attributes of this supplementary estimand are the same as those for the primary estimand (Estimand 1) with the exception of the Variable (Endpoint) and strategy for ICEs, which is described as follows:

Variable (Endpoint): Clinical response at Week 12. Participants who have any ICE in categories 1-5 (listed in Section 5.3.2) prior to the Week 12 visit will be considered not to have achieved clinical response at Week 12, regardless of the observed data.

5.3.4.2. Estimand 3

In this supplementary estimand for the primary endpoint (Estimand 3), all ICEs (Section 5.3.2) will be addressed by the **hypothetical strategy**.

The attributes of this supplementary estimand are the same as those for the primary estimand (Estimand 1) with the exception of the Variable (Endpoint) and strategy for ICEs, which is described as follows:

Variable (Endpoint): Clinical response at Week 12. Participants who have any ICE in categories 1-5 (listed in Section 5.3.2) prior to the Week 12 visit will be considered to have **missing** clinical response status at Week 12, regardless of the observed data.

This supplementary estimand uses a **hypothetical strategy** that considers the clinical response status at Week 12 after the ICE to be the same as for other participants in the same study intervention who had not yet experienced the ICE. The **hypothetical strategy** attempts to estimate the treatment effect if the randomized treatment is taken as specified in the protocol. As such, any clinical response status observed after an ICE is excluded in order to ensure results reflect participant experience on randomized treatment only. Therefore, utilizing the hypothetical estimand allows for an analysis of participants under the scenario when they are compliant with the investigative therapy, and thus, is justified clinically since it attempts to eliminate the potentially confounding effects due to ICEs.

5.3.5. Estimators (Analyses) for the Supplementary Estimands

5.3.5.1. Estimator (Analysis) for Estimand 2

For the supplementary analysis of the primary endpoint based on Estimand 2, after accounting for the ICEs for this supplementary estimand, participants who have a missing clinical response status at Week 12 will be considered not to be in clinical response at Week 12 (i.e. nonresponder imputation).

In this supplementary analysis, data from all participants in the FAS (Section 2.2.2) will be analyzed according to the randomized study intervention regardless of the study intervention they actually received. The analysis method is same as that for the main estimator (analysis) for the primary estimand (Estimand 1). Refer to Section 5.3.3.1 for details.

5.3.5.2. Estimator (Analysis) for Estimand 3

For the supplementary analysis of the primary endpoint based on Estimand 3, after accounting for the ICEs for this supplementary estimand, the missing status of clinical response at Week 12 will be imputed using multiple imputation (MI) as described in Table 3, under the assumption that the data are missing at random (MAR) for all participants in the FAS. The MI will be performed on each Mayo subscore with missing data and then the clinical response status at Week 12 will be derived from such imputed Mayo subscores.

Table 3: Multiple Imputation Methods

Endpoints	MI specification	Analysis method/Summary statistics
Clinical Response at Week 12	Multiple imputation with full conditional specification (FCS) regression of subscores	MIMayoData1 (N=200, Seed=4362478) • Imputation variables: 4 Mayo subscores from Week 0 – 12 (after accounting for ICEs) • Ancillary variables: Treatment group, corticosteroid use at baseline (yes, no)

In this supplementary analysis, data from all participants in the FAS will be analyzed according to the randomized study intervention regardless of the study intervention they actually received. For each of the N imputation datasets, the treatment difference between combination therapy versus each monotherapy will be tested using a 2-sided CMH chi-square test stratified by corticosteroid use at baseline (yes, no). The analysis results from all the imputation datasets will be combined according to Rubin[10], and the p-value for testing the treatment difference will be obtained.

5.4. Major Secondary Endpoint

The major secondary endpoint is clinical remission at Week 12.

5.4.1. Definition

Clinical remission: the Mayo score ≤ 2 with no individual subscore >1.

5.4.2. Main Estimand (Estimand 4)

The attributes and strategies for the ICEs that were used for the primary estimand (Estimand 1, defined in Section 5.3.2) will also be used for the main estimand (Estimand 4) for the major secondary endpoint.

5.4.3. Analysis Methods for the Main Estimand (Estimand 4)

5.4.3.1. Main Estimator (Analysis) for the Main Estimand

The major secondary endpoint will be analyzed based on the Main Estimand (Section 5.4.2). After accounting for the ICEs for the main estimand, participants who have a missing clinical remission status at Week 12 will be considered not to be in clinical remission at Week 12 (i.e. nonresponder imputation).

In this analysis for the major secondary endpoint, data from all participants in the FAS will be analyzed according to the randomized study intervention regardless of the study intervention they actually received. The treatment difference between combination therapy versus each monotherapy will be tested using a CMH chi-square test stratified by corticosteroid use at baseline (yes, no). The magnitude of the treatment difference will be estimated by the difference in the proportion of participants achieving clinical remission at Week 12 between combination therapy and each monotherapy with a 2-sided 80% CI calculated based on Wald statistic.

The testing of the major secondary endpoints will occur at a 2-sided significance level of 0.2 regardless of the significance of the primary endpoint. No adjustment for multiple comparisons will be made and nominal p-values will be presented.

5.4.3.2. Subgroup Analyses

Subgroup analyses will be performed (if the number of participants within each subgroup level permits) based on demographic and baseline disease characteristics, baseline concomitant UC medication use, and history of UC-related medications specified in Section 2.3. Note that, for subgroup analyses, the analysis sets are the individual subgroups of the FAS. A forest plot will be produced for all subgroups. Odds ratio of the combination therapy group versus each monotherapy group and the associated 80% CI from a logistic regression model will be provided for each of subgroups. The logistic regression model will include treatment group and corticosteroid use at baseline (yes, no) as the factors. For subgroup analyses of corticosteroid use at baseline (yes, no), treatment group will be the only the factor in the model. The main estimand (Estimand 4) will be used for these subgroup analyses.

5.4.4. Supplementary Estimands

Two supplementary estimands (Estimands 5 and 6) are considered to support the main estimand for the major secondary endpoint (Estimand 4).

5.4.4.1. Estimand 5

In this supplementary estimand for the major secondary endpoint (Estimand 5), all ICEs (Section 5.3.2) will be addressed by the **composite strategy**.

The attributes and strategies for the ICEs that were used for Estimand 2 for the primary endpoint (Section 5.3.4.1) will also be used for this supplementary estimand for the major secondary endpoint (Estimand 5).

5.4.4.2. Estimand 6

In this supplementary estimand for the major secondary endpoint (Estimand 6), all ICEs (Section 5.3.2) will be addressed by the **hypothetical strategy**.

The attributes and strategies for the ICEs that were used for Estimand 3 for the primary endpoint (Section 5.3.4.2) will also be used for this supplementary estimand for the major secondary endpoint (Estimand 6).

5.4.5. Estimators (Analyses) for the Supplementary Estimands

5.4.5.1. Estimator (Analysis) for Estimand 5

For the supplementary analysis of the major secondary endpoint based on Estimand 5, after accounting for the ICEs for this supplementary estimand, participants who have a missing clinical remission status at Week 12 will be considered not to be in clinical remission at Week 12 (i.e. nonresponder imputation).

In this supplementary analysis for the major secondary endpoint, data from all participants in the FAS (Section 2.2.2) will be analyzed according to the randomized study intervention regardless of the study intervention they actually received. The analysis method is same as that for the main estimator (analysis) for the primary estimand (Estimand 1). Refer to Section 5.3.3.1 for details.

5.4.5.2. Estimator (Analysis) for Estimand 6

For the supplementary analysis of the major secondary endpoint based on Estimand 6, after accounting for the ICEs for this supplementary estimand, the missing status of clinical remission at Week 12 will be imputed using MI. For participants with missing status of clinical remission at Week 12, their clinical remission status at Week 12 will be derived from the imputed Mayo subscores in MIMayoData1 that has been created in Section 5.3.5.2.

The analysis method based on this supplementary estimand (Estimand 6) for the major secondary endpoint is same as the analysis method based on the supplementary estimand (Estimand 3) for the primary endpoint. Refer to Section 5.3.5.2 for details.

5.5. Other Efficacy Endpoints

In addition to the primary and major secondary endpoints, other efficacy endpoints related to disease status, HRQoL outcomes (including fatigue), and inflammatory biomarkers will also be analyzed. This section outlines the definitions (Section 5.5.1) and analyses (Section 5.5.2) of the other efficacy endpoints.

Clinical Endpoints

- Clinical response at Week 38

- Clinical response at Week 38 by clinical response status at Week 12
- Clinical remission at Week 38
- Clinical remission at Week 38 by clinical remission status at Week 12
- Clinical remission at Week 38 by clinical response status at Week 12
- Change from baseline in Mayo score at Weeks 12 and 38.
- Change from baseline in partial Mayo score by visit through Week 38.
- Change from baseline in modified Mayo score at Weeks 12 and 38.
- Change from baseline in Mayo subscores by visit through Week 38.
- Modified Mayo response at Weeks 12 and 38
- Modified Mayo response at Week 38 by modified Mayo response status at Week 12
- Symptomatic remission by visit through Week 38
- Endoscopic healing at Weeks 12 and 38
- Endoscopic healing at Week 38 by endoscopic healing status at Week 12
- Endoscopic normalization at Weeks 12 and 38.
- Endoscopic normalization at Weeks 38 by endoscopic normalization status at Week 12
- Clinical response, clinical remission, and endoscopic healing at Weeks 12 and 38 by negative response signature status at baseline.
- Clinical remission by alternative definitions:
    - Clinical remission (UNIFI definition) at Weeks 12 and 38
    - Clinical remission (UNIFI definition) at Week 38 by clinical remission (UNIFI definition) status at Week 12
    - Clinical remission (Health Authority definition 1) at Weeks 12 and 38
    - Clinical remission (Health Authority definition 1) at Week 38 by clinical remission (Health Authority definition 1) status at Week 12
    - Clinical remission (Health Authority definition 2) by visit at Weeks 12 and 38
    - Clinical remission (Health Authority definition 2) at Week 38 by clinical remission (Health Authority definition 2) status at Week 12
- Histologic healing at Weeks 12 and 38
- Mucosal healing at Weeks 12 and 38
- Histologic remission at Weeks 12 and 38
- Histologic-endoscopic mucosal healing at Weeks 12 and 38
- Deep histologic-endoscopic mucosal healing at Weeks 12 and 38
- Geboes total score, high activity subscore, and low activity subscore at Weeks 0, 12, and 38
- Roberts Histologic Index (RHI)-based histologic remission at Weeks 12 and 38

- Nancy Histologic Index (NHI)-based histologic remission at Weeks 12 and 38
- Ulcerative Colitis Endoscopic Index of Severity (UCEIS) score at Weeks 0, 12, and 38 by the level of Mayo endoscopy score at the corresponding visit
- Change from baseline in UCEIS score at Weeks 12 and 38
- UCEIS score ≤ 4 at Weeks 12 and 38
- Clinical remission and not receiving concomitant corticosteroids at Week 12
- Clinical response and not receiving concomitant corticosteroids at Week 12
- Clinical remission and not receiving concomitant corticosteroids at Week 38
- Clinical response and not receiving concomitant corticosteroids at Week 38
- Clinical remission and not receiving concomitant corticosteroids at Week 38 by status of clinical remission and not receiving concomitant corticosteroids at Week 12
- Clinical response and not receiving concomitant corticosteroids at Week 38 by status of clinical response and not receiving concomitant corticosteroids at Week 12
- ICEs (listed in Section 5.3.2) prior to Week 12 and prior to Week 38
- Mayo score missing at Week 12 by reason causing missing (including COVID-19 related) and at Week 38 by reason causing missing (including COVID-19 related)
- Endoscopy subscore missing at Week 12 by reason causing missing (including COVID-19 related) and at Week 38 by reason causing missing (including COVID-19 related)

Inflammatory Biomarkers (CRP and Fecal Calprotectin)

- Change from baseline in CRP by visit through Week 38
- Change from baseline in CRP by visit through Week 38 among participants with abnormal CRP concentration at baseline
- Change from baseline in fecal calprotectin concentration by visit through Week 38
- Change from baseline in fecal calprotectin concentration by visit through Week 38 among participants with abnormal fecal calprotectin concentration at baseline
- Normalization of CRP concentration by visit through Week 38 among participants with abnormal CRP concentration at baseline
- Normalization of fecal calprotectin concentration by visit through Week 38 among participants with abnormal fecal calprotectin concentration at baseline

HRQoL Endpoints

- Change from baseline in the total score of the Inflammatory Bowel Disease Questionnaire (IBDQ) by visit through Week 38
- A >20-point improvement from baseline in the IBDQ score by visit through Week 38
- Change from baseline in the 7 domain scores (norm-based) and the abdominal pain intensity score of Patient-Reported Outcomes Measurement Information System (PROMIS)-29 by visit through Week 38

- A ≥ 5-point improvement from baseline in PROMIS-29 domain scores (norm-based) and pain intensity score by visit through Week 38
- Change from baseline in the PROMIS Fatigue Short Form 7a total score by visit through Week 38
- Fatigue response (≥ 5-point improvement) based on the PROMIS Fatigue Short Form 7a by visit through Week 38

5.5.1. Definitions

5.5.1.1. Clinical Endpoints

5.5.1.1.1. Modified Mayo Response

Modified Mayo response is defined as a decrease from baseline in modified Mayo score ≥ 30% and ≥ 2 points, with either a decrease in the RBS ≥ 1 or a RBS of 0 or 1.

5.5.1.1.2. Symptomatic remission

Symptomatic remission is defined as a stool frequency subscore of 0 or 1, where the stool frequency subscore has not increased from baseline, and a rectal bleeding subscore of 0.

5.5.1.1.3. Endoscopic healing

Endoscopic healing (i.e., improvement in the endoscopic appearance of the mucosa) is defined as an endoscopy subscore of 0 or 1 with no friability present on the endoscopy.

5.5.1.1.4. Endoscopic normalization (i.e. normalization of endoscopic appearance of mucosa in protocol)

Endoscopic normalization is defined as an endoscopic subscore of 0 with no friability present on the endoscopy.

5.5.1.1.5. Clinical remission by alternative definitions

Clinical Remission (UNIFI definition) is defined as an absolute stool number ≤ 3, a rectal bleeding subscore of 0, and an endoscopy subscore of 0 or 1 with no friability present on the endoscopy.

Clinical Remission (Health Authority definition 1) is defined as a stool frequency subscore of 0 or 1, a rectal bleeding subscore of 0, and an endoscopy subscore of 0 or 1 with no friability present on the endoscopy, where the stool frequency subscore has not increased from baseline.

Clinical Remission (Health Authority definition 2) is defined as a 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.

5.5.1.1.6. Histologic healing

Histologic healing is defined as 0 - < 5% neutrophils in epithelium and no crypt destruction, and no erosion, ulceration, or granulation tissue according to the Geboes grading system[3] (Attachment 3).

5.5.1.1.7. Mucosal healing

Mucosal healing is a combination of histologic healing (Section 5.5.1.1.6) and endoscopic healing (Section 5.5.1.1.3).

5.5.1.1.8. Histologic remission

Histologic remission is 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[3] (Attachment 3).

5.5.1.1.9. Histologic-endoscopic mucosal healing

Histologic-endoscopic mucosal healing is a combination of histologic remission (Section 5.5.1.1.8) and endoscopic healing (Section 5.5.1.1.3).

5.5.1.1.10. Deep histologic-endoscopic mucosal healing

Deep histologic-endoscopic mucosal healing is a combination of histologic remission (Section 5.5.1.1.8) and endoscopic normalization (Section 5.5.1.1.4).

5.5.1.1.11. Geboes Scores

The **Geboes total score**, a continuous histology score, is calculated as the sum of all Geboes Grades (Attachment 3) and may take on values from 0 to 22.

The **Geboes high activity subscore**, a continuous histology score, is calculated as the sum of Geboes Grades 3, 4, and 5 and may take on values from 0 to 10.

The **Geboes low activity subscore**, a continuous histology score, is calculated as the sum of Geboes Grades 0, 1, 2A and 2B and may take on values from 0 to 12.

5.5.1.1.12. Roberts Histologic Index (RHI)-based histologic remission

RHI-based histologic remission is defined as RHI $\leq$ 3 with sub-scores of 0 for lamina propria neutrophils and neutrophils in the epithelium according to the Robarts Histologic Index[7] (Attachment 4).

5.5.1.1.13. Nancy Histologic Index (NHI)-based histologic remission

NHI-based histologic remission is defined as NHI $\leq$ 1 according to the Nancy Histologic Index[6] (Attachment 5).

5.5.1.1.14. Ulcerative Colitis Endoscopic Index of Severity (UCEIS)

The **UCEIS** (Attachment 6) is an index that provides an overall assessment of endoscopic severity of UC, based on mucosal vascular pattern, bleeding, and ulceration[12]. The score ranges from 3 to 11, with a higher score indicating more severe disease by endoscopy. The UCEIS score will be assessed only by the central video readers for all endoscopies.

5.5.1.2. Inflammatory Biomarkers

5.5.1.2.1. C-Reactive Protein

C-reactive protein (CRP) has been demonstrated to be useful as a marker of inflammation in participants with IBD. In subjects with UC, elevated CRP has been associated with more severe clinical activity, an elevated sedimentation rate, and active disease as detected by colonoscopy[11,15].

5.5.1.2.2. Fecal Calprotectin

Fecal calprotectin has been demonstrated to be a sensitive and specific marker in identifying intestinal inflammation and response to treatment in participants with IBD, especially in UC[1].

Assays for fecal calprotectin concentration will be performed by the central laboratory using a validated method. Additional tests may also be performed on the stool samples for additional markers that are related to intestinal inflammation and treatment response such as the microbiome.

5.5.1.3. HRQoL Endpoints

5.5.1.3.1. Inflammatory Bowel Disease Questionnaire (IBDQ)

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 dimensional scores: bowel symptoms (loose stools, abdominal pain), systemic function (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.

5.5.1.3.2. PROMIS-29

The **PROMIS-29** is a validated general health profile instrument that is not disease-specific. It is a collection of short forms containing 4 items for each of 7 domains (depression, anxiety, physical function, pain interference, fatigue, sleep disturbance, and ability to participate in social roles and activities). PROMIS-29 also includes an overall average pain intensity 0-10 numeric rating scale. Norm-based scores have been calculated for each domain on the PROMIS measures, with a score of 50 representing the mean or average of 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.

5.5.1.3.3. PROMIS Fatigue 7-items Short Form

The **PROMIS Fatigue Short Form 7a** contains 7 items evaluating fatigue-related symptoms (ie, tiredness, exhaustion, mental tiredness, and lack of energy) and associated impacts on daily activities (ie, activity limitations related to work, self-care, and exercise). PROMIS Fatigue Short

Form 7a has a recall period of past 7 days. Compared to the fatigue scale of PROMIS-29, PROMIS Fatigue Short Form 7a provides additional information to evaluate severity of fatigue.

Fatigue response is defined as a ≥ 5-point improvement in the PROMIS Fatigue Short Form 7a.

5.5.2. Estimands

Unless otherwise specified, the attributes and strategies for the ICEs that were used for the primary estimand of the primary endpoint (Estimand 1, Section 5.3.2) will also be used for the estimand for each of the other endpoints (Section 5.5).

ICE categories 1-3 will be handled by the **composite strategy**. ICE category 4 will be handled by the **hypothetical strategy** (as if participants would have not experienced this intercurrent event), and ICE category 5 will be handled by the **treatment policy strategy**. To be more specific,

- Participants with an ICE in category 1-3 prior to a visit will be considered, at that visit and all subsequent visits, as **not having achieved the binary endpoints for binary endpoints** and as **having no change from baseline (i.e., the baseline value will be assigned) for continuous endpoints.**
- Participants with an ICE 4 will have their observed data set to missing at all visits after the ICE 4.
- Participants with an ICE 5 will have their observed data used after the ICE 4 used, if available.
- For participants experiencing multiple ICEs, ICEs in categories 1-3 will override ICEs 4 and 5.

5.5.3. Estimators (Analyses) for Estimands

Unless otherwise specified, the other endpoints presented in Section 5.5 will be analyzed based on the FAS (Section 2.2.2) according to randomized study intervention regardless of the study intervention actually received.

Descriptive statistics (i.e., mean, median, SD, IQ range, minimum, and maximum) will be used to summarize continuous endpoints and counts and percentages will be used to summarize binary endpoints by visit.

Treatment comparisons between combination therapy versus each monotherapy will be performed by visit through Week 38. All statistical testing will be performed at the 2-sided significance level of 0.2. No adjustment for multiple comparisons will be made for these other endpoints. Nominal p-values will be presented.

Binary Endpoints

After accounting for the ICEs, participants with missing status for a binary endpoint will be considered not to have achieved the associated binary endpoint (i.e. nonresponder imputation).

Binary endpoints will be summarized with the number and frequency of participants who achieve the endpoint by treatment group. Treatment comparisons (combination therapy versus each monotherapy) will be performed using analyses suitable for categorical data (e.g., a CMH chi-square test stratified by corticosteroid use at baseline, as appropriate) to compare the proportion of participants achieving the endpoints. In case of rare events, the Fisher's exact test will be used for treatment comparisons.

Continuous Endpoints

For continuous endpoints that are assessed only at one post-baseline visit (e.g., change from baseline in Mayo score and change from baseline in modified Mayo score at the Week 12 DBL), the missing data after accounting for ICEs will be imputed by the MI methods as described in Table 3. The MI will be performed on each Mayo subscore with missing data and then the Mayo scores will be derived from such imputed subscores. These endpoints will be compared between combination therapy versus each monotherapy using an analysis of covariance (ANCOVA) with treatment group, corticosteroid use at baseline (yes, no), and baseline score as the explanatory factors.

For continuous endpoints that are assessed at more than one post-baseline visit, to account for the missing data after the ICEs have been accounted for, a MMRM will be used, under the assumption of MAR, to test the treatment difference between combination therapy versus each monotherapy. In MMRM, missing data will not be imputed, but rather missing data will be accounted for through correlation of repeated measures in the MMRM model. If the MMRM normality assumption is in question, an appropriate transformation may be implemented before fitting the MMRM model.

The explanatory variables of the MMRM model will include treatment group, corticosteroid use at baseline (yes, no), visit, respective baseline score, and an interaction term of visit with treatment group. An unstructured covariance matrix for repeated measures within a subject will be used. The F-test will use Kenward-Roger's approximation for degree 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. The treatment difference between combination therapy and each monotherapy will be estimated by the difference in the least squares means (LSmeans). The 80% CI for the difference in LSmeans and p-values will be calculated based on the MMRM.

5.6. Exploratory Endpoints

Exploratory endpoints include:

- Change from baseline in average daily number of BSFS types 6 and 7 stools by visit through Week 12
- Change from baseline in average daily number of BSFS types 5, 6, and 7 stools by visit through Week 12
- A ≥ 2 reduction in average daily number of BSFS types 6 and 7 stools by visit through Week 12

- A ≥ 2 reduction in average daily number of BSFS types 5, 6, and 7 stools by visit through Week 12
- Proportion of participants in each category of Patient's Global Impression of Change (PGIC) of Severity of UC by visit through Week 38
- A ≥ 1-point improvement in PGIC by visit through Week 38
- A ≥ 2-point improvement in PGIC by visit through Week 38

5.6.1. Definitions

5.6.1.1. Bristol Stool Form Scale

The **BSFS** is a medical aid to classify the form (or consistency) of human feces into 7 categories[5]. It has been used as a research tool to evaluate the effectiveness of treatments for various diseases of the bowel (e.g., irritable bowel syndrome). Participants will complete the BSFS as a daily diary entry from Week 0 through Week 12.

Average number of BSFS types 6 and 7 stools per day is defined as: the sum of number of BSFS types 6 and 7 stools in previous 7 days in a dairy card ÷ total days assessment performed. Similarly, average number of BSFS types 5, 6 and 7 stools per day is defined as: the sum of number of BSFS types 5, 6 and 7 stools in previous 7 days in a dairy card ÷ total days assessment performed. Average number of BSFS stools per day at a scheduled visit will not be calculated if total days of assessment is less than 5 within the previous 7 days prior to a scheduled visit.

5.6.1.2. Patient's Global Impression of Change (PGIC) of Severity of Ulcerative Colitis

Participants' perceived change (improvement or deterioration) in the severity of their UC will be assessed using the **PGIC**. Participants will rate how their UC has changed since the beginning of the study using a 7-point scale ranging from "a lot better now" to "a lot worse now" with a neutral center point ("neither better nor worse").

5.6.2. Estimands

Unless otherwise specified, the attributes and strategies for the ICEs that were used for the primary estimand of the primary endpoint (Estimand 1, Section 5.3.2) will also be used for the estimand for each exploratory efficacy endpoint (Section 5.6).

ICE categories 1-3 will be handled by the **composite strategy**. ICE category 4 will be handled by the **hypothetical strategy** (as if participants would have not experienced this intercurrent event), and ICE category 5 will be handled by the **treatment policy strategy**. To be specific,

- Participants with an ICE in category 1-3 prior to a visit will be considered, at that visit and all subsequent visits, as **not having achieved the binary endpoints for binary endpoints** and as **having no change from baseline (i.e., the baseline value will be assigned) for continuous** and **ordinal endpoints.**

- PGIC has no baseline data. When PGIC is analyzed as an ordinal endpoint, participants with an ICE in category 1-3 prior to a visit will be considered to have a UC severity change of "neither better, nor worse (no change)" from that visit onwards.

- Participants with an ICE 4 will have their observed data set to missing at all visits after the ICE 4.

- Participants with an ICE 5 will have their observed data used after the ICE 4 used, if available.

- For participants experiencing multiple ICEs, ICEs in categories 1-3 will override ICEs 4 and 5.

5.6.3. Estimators (Analyses) for Estimands

Unless otherwise specified, exploratory endpoints presented in Section 0 will be analyzed based on the FAS (Section 2.2.2) according to randomized study intervention regardless of the study intervention actually received.

Descriptive statistics (i.e., mean, median, SD, IQ range, minimum, and maximum) will be used to summarize continuous endpoints and counts and percentages will be used to summarize binary endpoints by visit.

Treatment comparisons between combination therapy versus each monotherapy will be performed by visit through Week 38. No adjustment for multiple comparisons will be made for these other endpoints. Nominal p-values will be presented.

Binary Endpoints

After accounting for the ICEs, participants with missing status for a binary endpoint will be considered not to have achieved the associated binary endpoint (i.e. nonresponder imputation). Treatment comparisons (combination therapy versus each monotherapy) will be performed using analyses suitable for categorical data (e.g., a CMH chi-square test stratified by corticosteroid use at baseline, as appropriate) to compare the proportion of participants achieving the endpoints. In case of rare events, the Fisher's exact test will be used for treatment comparisons.

Continuous Endpoints

After accounting for the ICEs, to account for the missing data for continuous endpoints, a MMRM as described in Section 5.5.3 will be used, under the assumption of MAR, to test the treatment comparison between combination therapy versus each monotherapy. In MMRM, missing data will not be imputed, but rather missing data will be accounted for through correlation of repeated measures in the model. If the normality assumption is in question, an appropriate transformation may be implemented before fitting the MMRM model.

Ordinal Endpoints

After accounting for the ICEs, treatment comparisons (combination therapy versus each monotherapy) will be performed using a CMH chi-square test (Row Mean Scores) stratified by corticosteroid use at baseline (yes, no).

6. SAFETY

Safety data, including but not limited to, AEs, serious adverse events (SAEs), infections, serious infections, and changes in laboratory assessments will be summarized through Week 12 and through Week 38. Treatment-emergent AEs will be summarized by treatment group and Medical Dictionary for Regulatory Activities (MedDRA) system organ class and preferred terms.

Unless otherwise specified, safety analyses will be provided for participants in the Safety Analysis Set, which includes all randomized participants who received at least 1 dose of study intervention, according to the study intervention that they actually received, regardless of the study intervention they were randomized to. No formal statistical comparisons are planned.

6.1. Adverse Events

The verbatim terms used in the eCRF by investigators to identify AEs will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). Treatment-emergent AEs (TEAEs) are AEs with onset during the intervention phase or that are a consequence of a pre-existing condition that has worsened since baseline. All reported AEs which are treatment-emergent will be included in the analysis. For each AE, the count and percentage of participants who experience at least 1 occurrence of the given event will be summarized by intervention group.

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

- Frequency and type of AEs
- Frequency and type of SAEs
- Frequency and type of reasonably related AEs as assessed by the investigator
- Frequency and type of AEs leading to discontinuation of study intervention
- Frequency and type of infections (including COVID-19 infections)
- Frequency and type of serious infections
- Frequency and type of infections requiring oral or parenteral antimicrobial treatment
- Frequency and type of AEs temporally associated with infusion
- Frequency and type of injection-site reactions

Since safety should be assessed relative to exposure and follow-up, all AE summary tables will summarize the average weeks of follow-up and average exposure (number of administrations) for each treatment group.

In addition to the summary tables, a by-subject listing will be provided for deaths that occurred during the study and, respectively, for the following TEAEs:

1. SAEs
2. AEs that led to permanent discontinuation of study intervention
3. Serious infections including TB
4. COVID-19 infections
5. Anaphylactic reactions or serum sickness reactions
6. Malignancies

6.2. Clinical Laboratory Tests

Routine laboratory data for hematology and clinical chemistry will be collected at study visits from Week 0 through Week 38 according to SoA in the Protocol Section 1.3. The following laboratory assessments will be collected:

1. Hematology: hemoglobin, hematocrit, platelet count, total and differential white blood cell count (WBC) count.

2. Chemistry: total and direct bilirubin, alanine aminotransferase (ALT), aspartate aminotransferase (AST), alkaline phosphatase, albumin, total protein, calcium, phosphate, sodium, potassium, chloride, blood urea nitrogen (BUN), and creatinine.

The following summaries of clinical laboratory tests will be provided for participants in the Safety Analysis Set (Section 2.2.3):

1. Summary of laboratory parameters and change from baseline in laboratory parameters over time
2. Plots of laboratory parameters and changes from baseline in selected clinical laboratory parameters over time
3. Summary of maximum NCI-CTCAE toxicity grade for postbaseline laboratory values
4. Shift tables for maximum NCI-CTCAE toxicity grade for selected laboratory parameters (hematology: hemoglobin, platelets, total WBC, absolute lymphocytes, and absolute neutrophils; chemistry: ALT, AST, and alkaline phosphatase) from baseline to corresponding postbaseline laboratory values
5. Summary of maximum postbaseline measurement for ALT, AST, alkaline phosphatase and total bilirubin relative to ULN
6. Summary of maximum postbaseline elevated liver function tests (AST or ALT > 5xULN)
7. Patterns of change in ALT and AST will be assessed by summarizing the number of participants with only one ALT/AST > 1 ULN or ≥ 3 ULN or ≥ 5 ULN; with 2 or more consecutive ALT/AST measurements >1 ULN or ≥ 3 ULN or ≥ 5 ULN; with 2 or more non-consecutive ALT/AST measurements
8. Summary of laboratory data missed or collected outside window due to COVID-19 Pandemic

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.

Clinical laboratory test values are to be graded based on National Cancer Institute Common Terminology Criteria for Adverse Events (NCI-CTCAE) version 5.0 (Attachment 7). The laboratory tests not included in Attachment 7 will not be presented in the corresponding tables or listings.

Listings of participants with any abnormal post-baseline laboratory values of NCI-CTCAE grade ≥ 2 and maximum postbaseline elevated liver function tests (AST or ALT > 5xULN) will also be provided.

6.3. Other Safety Parameters

6.3.1. Suicidal Ideation and Behavior

The Columbia-Suicide Severity Rating Scale (C-SSRS) will be used as a screening tool to prospectively evaluate suicidal ideation and behavior in this study, as part of a comprehensive evaluation of safety. The C-SSRS is an investigator-administered questionnaire that defines 5 subtypes of suicidal ideation and 4 possible suicidal behaviors, as well as self-injurious behavior with no suicidal intent and completed suicide, and is a fully-structured subject self-report questionnaire, including standardized questions, follow-up prompts, error handling routines, and scoring conventions[8,9].

The baseline is defined as the most severe/maximum score at screening and Week 0. Suicidal ideation and behavior will be summarized by the most severe/maximum post baseline outcome. A shift table from baseline to post-baseline will also be provided. Participants with positive (i.e., score > 0) ideation and behavior will be presented in a data listing.

7. PHARMACOKINETICS/PHARMACODYNAMICS

7.1. Pharmacokinetics

PK analyses will be based on the PK Analysis Set (Section 2.2.4). Participants will be analyzed according to the study intervention that they actually received, regardless of the treatments they are randomized to. No imputation for missing data will be performed.

7.1.1. Serum Guselkumab and Golimumab Concentrations

Blood samples for determining the serum guselkumab and golimumab concentrations will be drawn from all participants according to the Schedule of Activities in the Protocol. Unless otherwise mentioned, serum guselkumab and golimumab concentration summaries will be provided by study intervention at each visit through Week 12 and through Week 38. Descriptive statistics of the serum guselkumab and golimumab concentrations will be calculated by treatment group at each sampling time point when appropriate, including n, arithmetic mean, SD, coefficient of variation (%CV), median, interquartile range, range (minimum and maximum). PK data may be displayed graphically. The following analyses will be performed as appropriate:

- Summary of serum guselkumab and golimumab concentrations at each visit by study intervention

- Proportion of participants without detectable serum guselkumab and golimumab concentration (below the lower limit of quantification [< LLOQ]) at each visit by study intervention
- Summary of serum guselkumab and golimumab concentrations at each visit by study intervention and baseline body weight (quartiles)
- Plot of median serum guselkumab and golimumab concentrations over time by study intervention
- Plot of median serum guselkumab and golimumab concentrations over time by study intervention and baseline body weight (≤ median, > median)

7.1.1.1. Data Handling Rules

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

- A concentration not quantifiable (< LLOQ) will be treated as 0 in the summary statistics and shown as '< LLOQ' in the data listings.

- The data from a participant who meets 1 of the following dosing deviation criteria will be excluded from the by-visit data analyses from that point onwards:

    - Discontinued guselkumab or golimumab administrations.
    - Skipped a guselkumab or golimumab administration.
    - Received an incomplete/ incorrect dose.
    - Received an incorrect study agent.
    - Received an additional guselkumab or golimumab dose.

In addition, if a participant has an administration outside of dosing windows (Table 4), the sample data collected at that visit and after that visit prior to the next administration will be excluded from the by-visit data analyses.

Table 4: Dosing Window

Visit	Window
Week 0 through Week 38	± 4 days from scheduled visit day
Final Safety and Efficacy Follow-up visits	± 7 days from scheduled visit day

7.1.2. PK vs Efficacy

The relationship between serum guselkumab and/or golimumab concentrations and efficacy endpoints may be explored, e.g.:

1. The relationship between serum guselkumab and golimumab concentrations (quartiles) and clinical response, clinical remission, change in Mayo score, symptomatic remission, and endoscopic healing status at Week 12 and Week 38 may be explored by treatment group.

2. The relationship between serum guselkumab and golimumab concentrations (quartiles) and change from baseline in CRP concentration (mg/L) and Fecal Calprotectin concentration (mg/kg) at Week 12 and Week 38 may be explored by treatment group.

7.1.3. Population PK Analysis

When appropriate, population PK analysis will be performed using serum guselkumab concentration-time data in PK analysis set with the nonlinear mixed-effects modeling (NONMEM) approach. Population PK analysis for golimumab may also be explored when appropriate. Details will be provided in a separate document.

7.2. Immunogenicity

Immunogenicity analyses will be based on the Immunogenicity Analysis Set (Section 2.2.5). Participants will be analyzed according to the study intervention that they actually received. No imputation for missing data will be performed.

7.2.1. Antibodies to Guselkumab and to Golimumab

Blood samples will be collected to examine the formation of antibodies to guselkumab and/or to golimumab at the specified visits as shown in the SoA (Protocol Section 1.3). Serum samples will also be collected at the final visit from participants who terminate study participation early.

The antibodies to guselkumab and/or to golimumab status (positive at any time, negative) and peak titers will be summarized by treatment group through Week 12 and through Week 38.

A listing of participants who are positive for antibodies to guselkumab and/or to golimumab will be provided. The sample antibodies status, the titer, and the neutralizing antibodies status will be listed by visit. This listing will also provide information regarding dose administered, injection-site reactions and/or reactions temporally associated with infusion, guselkumab and/or golimumab serum concentration, partial Mayo score, and Mayo score for all visits. In addition, a list of antibodies to guselkumab and/or to golimumab status in participants who discontinued study agent early will be provided.

7.2.1.1. Neutralized Antibodies to Guselkumab and to Golimumab

The incidence of neutralizing antibodies (NAbs) to guselkumab and/or to golimumab will be summarized through Week 12 and through Week 38 for participants who are positive for antibodies to guselkumab and/or to golimumab, respectively, and have samples evaluable for NAbs to guselkumab and/or to golimumab.

7.2.2. Antibody vs PK/Efficacy/Safety

To explore the relationship between antibodies to guselkumab status and to golimumab status and serum guselkumab and serum golimumab concentrations, and efficacy and safety, the following analyses may be performed as appropriate:

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

- Plots of median trough serum golimumab concentrations over time by antibodies to golimumab status
- Summary of clinical response, clinical remission, change in Mayo score, and endoscopic healing, all at Week 12 and Week 38, by antibodies to guselkumab and to golimumab status if sufficient participants have antibodies
- Summary of injection-site reactions by antibodies to guselkumab and to golimumab status
- Summary of AEs temporally associated with infusion by antibodies to guselkumab

7.3. Pharmacokinetic/Pharmacodynamic Relationships

If data permit, the relationships between serum guselkumab and serum golimumab concentration and efficacy may be analyzed graphically. If any visual trend is observed, a suitable population PK/PD model may be developed to describe the exposure-response relationship. Details will be given in a population PK/PD analysis plan and results of the population PK/PD analysis will be presented in a separate technical document.

8. BIOMARKERS

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

Changes in serum protein analytes, fecal biomarkers, and biopsy and whole blood RNA obtained over time will be summarized by study intervention. Associations between baseline levels and changes from baseline in selected markers and response to treatment will be explored. Biomarker analyses will be summarized in a separate technical report.

9. MEDICAL RESOURCE UTILIZATION AND HEALTH ECONOMICS

Medical resource utilization evaluations, including but not limited to UC-related emergency department visits, hospitalizations, and surgeries are collected in this study. The following is a list of the endpoints to be analyzed:

- Proportion of participants having any UC-related ER/hospitalizations/surgeries through Week 12 and though Week 38 (final safety follow-up)
- Proportion of participants with a UC-related hospitalization through Week 12 and though Week 38 (final safety follow-up)
- Proportion of participants with a UC-related surgery through Week 12 and though Week 38

The health economics data will be analyzed based on the FAS (Section 2.2.2) according to randomized study intervention regardless of the study intervention actually received.

The ICEs in categories 1 - 5 specified in Section 5.3.2 will not be used. No imputation will be performed for missing data and the missing values will remain as missing.

Treatment comparisons will be performed. Nominal p-values will be reported. These endpoints will not be adjusted for multiplicity.

Binary endpoints will be summarized with the number and frequency of participants by treatment group. Treatment comparisons (combination therapy versus each monotherapy) will be performed using a CMH chi-square test stratified by corticosteroid use at baseline (yes, no). In case of rare events, the Fisher's exact test will be used for treatment comparisons.

REFERENCES

1. Abraham BP and Kane S. Fecal Markers: Calprotectin and Lactoferrin. Gastroenterol Clin N Am. 2012;41:483-495.
2. Baron JH, Connell AM, Lennard-Jones JE. Variation between observers in describing mucosal appearances in proctocolitis. Brit Med J. 1964;1:89-92.
3. 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
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. Lewis SJ, Heaton KW. Stool form scale as a useful guide to intestinal transit time. Scand J Gastroenterol. 1997;32(9):920–924.
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. Rubin DB, Multiple Imputations for Nonresponse in Surveys. New York: John Wiley & Sons. 1987
11. Solem CA, Loftus EV, Tremaine WJ, et al. Correlation of C-reactive protein with clinical, endoscopic, histologic, and radiographic activity in inflammatory bowel disease. Inflamm Bowel Dis. 2005;11(8):707-712.
12. Travis SP, Schnell D, Krzeski P, et al. Developing an instrument to assess the endoscopic severity of ulcerative colitis: the Ulcerative Colitis Endoscopic Index of Severity (UCEIS). Gut. 2012;61:535-542.
13. Telesco S, Greenbaum L, Zhang H, et al. Prediction of Non-Response to Anti-TNF Therapy in Ulcerative Colitis: Implications for Patient Stratification for IBD Trials and Novel Treatment Paradigms. Gastroenterology. 2017;152(5):S984-S985
14. Truelove SC, Witts LJ. Cortisone in ulcerative colitis: final report on a therapeutic trial. Brit Med J. 1955;2:1041-1048.
15. Vermeire S, Van Assche G, Rutgeerts P. C-reactive protein as a marker of inflammatory bowel disease. Inflamm Bowel Dis. 2004;10(5):661-665.

ATTACHMENTS

Attachment 1: MAYO SCORE

Mayo scoring system for assessment of ulcerative colitis activity

Stool frequency[a]
0 = Normal number of stools for this patient
1 = 1-2 stools more than normal
2 = 3-4 stools more than normal
3 = 5 or more stools more than normal

Rectal bleeding[b]
0 = No blood seen
1 = Streaks of blood with stool less than half the time
2 = Obvious blood with stool most of the time
3 = Blood alone passed

Findings of endoscopy
0 = Normal or inactive disease
1 = Mild disease (erythema, decreased vascular pattern, mild friability)
2 = Moderate disease (marked erythema, absent vascular pattern, friability, erosions)
3 = Severe disease (spontaneous bleeding, ulceration)

Physician's global assessment[c]
0 = Normal
1 = Mild disease
2 = Moderate disease
3 = Severe disease

[a] At the screening visit, each person indicates the number of stools he/she passed in a 24-hour period when in remission or before his/her UC diagnosis, thereby serving as his/her own control to establish the degree of abnormality of stool frequency.
[b] The daily bleeding score represents the most severe bleeding of the day.
[c] The physician's global assessment acknowledges the 3 other criteria, the patient's recall of abdominal discomfort and general sense of well-being, and other observations, such as physical findings and the patient's performance status.

Attachment 2: PROHIBITED CHANGES IN UC MEDICATIONS (ICE 2)

The following are the-protocol-prohibited changes in UC medication(s):

(i) Restricted or prohibited medications

- Initiation of restricted or prohibited medications or therapies as defined in the protocol (see Protocol Section 6.5.2), except for antibiotics used to treat UC, TPN, and apheresis

(ii) Oral corticosteroids

Over the Combination Comparison Phase (Through Week 12)

- Initiation of oral corticosteroids (including budesonide and beclomethasone dipropionate) due to worsening of disease after the baseline visit for participants who were not receiving oral corticosteroids at baseline
- Increase in the dose of oral corticosteroids (including budesonide and beclomethasone dipropionate) above the baseline dose, as specified below, due to worsening of disease
    i. Oral corticosteroids > 5 mg/day (prednisone equivalent)
    ii. Oral budesonide > 3 mg/day
    iii. Oral beclomethasone dipropionate > 5 mg/day
- Any switch among oral budesonide, oral beclomethasone dipropionate or other oral corticosteroids (excluding prednisone equivalent changes) due to worsening of disease

Over the Monotherapy Phase (After Week 12 Through Week 38)

- Initiation of oral corticosteroids (including budesonide and beclomethasone dipropionate) due to worsening of disease that lasts for more than 7 days after the Week 26 visit (i.e., approximately 90 days prior to Week 38) for participants who were not receiving oral corticosteroids at baseline
- Increase in the dose of oral corticosteroids (including budesonide and beclomethasone dipropionate) above the baseline dose, as specified below, due to worsening of disease for more than 7 days after the Week 26 visit (i.e., approximately 90 days prior to Week 38).
    i. Oral corticosteroids > 5 mg/day (prednisone equivalent)
    ii. Oral budesonide > 3 mg/day
    iii. Oral beclomethasone dipropionate > 5 mg/day
- Initiation of oral corticosteroids (including budesonide and beclomethasone dipropionate) due to reasons other than worsening of disease that lasts for more than 28 days after the Week 26 visit (i.e., approximately 90 days prior to Week 38) for participants who were not receiving oral corticosteroids at baseline
- Increase in the dose of oral corticosteroids (including budesonide and beclomethasone dipropionate) above the baseline dose, as specified below, due to reasons other than worsening of disease for more than 28 days after the Week 26 visit (i.e., approximately 90 days prior to Week 38)
    i. Oral corticosteroids > 5 mg/day (prednisone equivalent)
    ii. Oral budesonide > 3 mg/day

iii. Oral beclomethasone dipropionate > 5 mg/day

- Any switch among oral budesonide, oral beclomethasone dipropionate or other oral corticosteroids (excluding prednisone equivalent changes) due to worsening of disease

(iii) 5-ASA compounds

- Initiation of oral or rectal 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

Attachment 3: GRADING CRITERIA FOR THE HISTOLOGICAL EVALUATION OF DISEASE ACTIVITY IN ULCERATIVE COLITIS

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

Attachment 4: ROBARTS HISTOLOGIC INDEX

Component

Chronic inflammatory infiltrate
0=No increase
1=Mild but unequivocal increase
2=Moderate increase
3=Marked increase

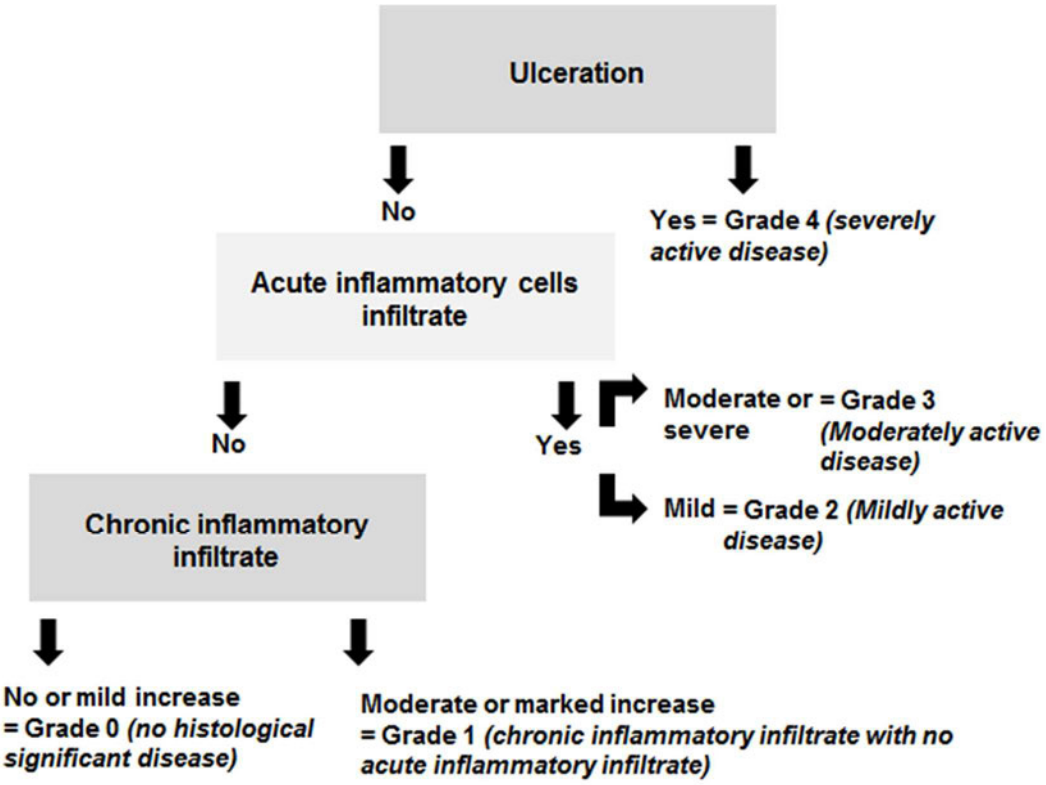
Lamina propria neutrophils
0=None
1=Mild but unequivocal increase
2=Moderate increase
3=Marked increase

Neutrophils in epithelium
0=None
1=<5% crypts involved
2=<50% crypts involved
3=>50% crypts involved

Erosion or ulceration
0=No erosion, ulceration or granulation tissue
1=Recovering epithelium+adjacent inflammation
1=Probable erosion-focally stripped
2=Unequivocal erosion
3=Ulcer or granulation tissue

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

Attachment 5: NANCY HISTOLOGIC INDEX



Attachment 6: UCEIS

The UCEIS is an index that provides an overall assessment of endoscopic severity of UC based upon mucosal vascular pattern, bleeding, and ulceration. The score ranges from 3 to 11. The UCEIS score will be assessed only by the central readers for all endoscopies received.
UCEIS descriptors and definitions:

Descriptor (score most severe lesions)	Likert scale anchor points	Definition
Vascular pattern	Normal (1)	Normal vascular pattern with arborisation of capillaries clearly defined, or with blurring or patchy loss of capillary margins
	Patchy obliteration (2)	Patchy obliteration of vascular pattern
	Obliterated (3)	Complete obliteration of vascular pattern
Bleeding	None (1)	No visible blood
	Mucosal (2)	Some spots or streaks of coagulated blood on the surface of the mucosa ahead of the scope, which can be washed away
	Luminal mild (3)	Some free liquid blood in the Lumen
	Luminal moderate or severe (4)	Frank blood in the lumen ahead of endoscope or visible oozing from mucosa after washing intraluminal blood, or visible oozing from a haemorrhagic mucosa
Erosions and ulcers	None (1)	Normal mucosa, no visible erosions or ulcers
	Erosions (2)	Tiny (# 5mm) defects in the mucosa, of a white or yellow color with a flat edge
	Superficial ulcer (3)	Larger (> 5 mm) defects in the mucosa, which are discrete fibrincovered ulcers in comparison with erosions, but remain superficial
	Deep ulcer (4)	Deeper excavated defects in the mucosa, with a slightly raised edge

Attachment 7: NCI-CTCAE GRADING CRITERIA FOR HEMATOLOGY AND CHEMISTRY LABORATORY TESTS [CTCAE VERSION 5.0]

Hematology Tests		**Criteria**			
Test	**Direction**	**1**	**2**	**3**	**4**
Hemoglobin (g/dL)	Increase	>0 - 2 x ULN	>2 - 4 x ULN	>4 x ULN	
Hemoglobin (g/dL)	Decrease	<LLN - 10.0	<10.0 - 8.0	<8.0	
Lymphocytes (/mm^3)	Increase		>4000 - 20,000	>20,000	
Lymphocytes (/mm^3)	Decrease	<LLN - 800	<800 - 500	<500 - 200	<200
Neutrophils (/mm^3)	Decrease	<LLN - 1500	<1500 - 1000	<1000 - 500	<500
Platelets (/mm^3)	Decrease	<LLN - 75,000	<75,000 - 50,000	<50,000 - 25,000	<25,000
Total WBC count (/mm^3)	Increase			>100,000	
Total WBC count (/mm^3)	Decrease	<LLN - 3000	<3000 - 2000	<2000 - 1000	<1000

Chemistry Tests		**Criteria**			
Test	**Direction**	**1**	**2**	**3**	**4**
ALT	Increase	>ULN - 3.0 x ULN if baseline was normal; 1.5 - 3.0 x baseline if baseline was abnormal	>3.0 - 5.0 x ULN if baseline was normal; >3.0 - 5.0 x baseline if baseline was abnormal	>5.0 - 20.0 x ULN if baseline was normal; >5.0 - 20.0 x baseline if baseline was abnormal	>20.0 x ULN if baseline was normal; >20.0 x baseline if baseline was abnormal
AST	Increase	>ULN - 3.0 x ULN if baseline was normal; 1.5 - 3.0 x baseline if baseline was abnormal	>3.0 - 5.0 x ULN if baseline was normal; >3.0 - 5.0 x baseline if baseline was abnormal	>5.0 - 20.0 x ULN if baseline was normal; >5.0 - 20.0 x baseline if baseline was abnormal	>20.0 x ULN if baseline was normal; >20.0 x baseline if baseline was abnormal
Albumin (g/L)	Decrease	<LLN - 30	<30 - 20	<20	
Alkaline Phosphatase	Increase	>ULN - 2.5 x ULN if baseline was normal; 2.0 - 2.5 x baseline if baseline was abnormal	>2.5 - 5.0 x ULN if baseline was normal; >2.5 - 5.0 x baseline if baseline was abnormal	>5.0 - 20.0 x ULN if baseline was normal; >5.0 - 20.0 x baseline if baseline was abnormal	>20.0 x ULN if baseline was normal; >20.0 x baseline if baseline was abnormal
Bilirubin (total)	Increase	>ULN - 1.5 x ULN if baseline was normal; > 1.0 - 1.5 x baseline if baseline was abnormal	>1.5 - 3.0 x ULN if baseline was normal; >1.5 - 3.0 x baseline if baseline was abnormal	>3.0 - 10.0 x ULN if baseline was normal; >3.0 - 10.0 x baseline if baseline was abnormal	>10.0 x ULN if baseline was normal; >10.0 x baseline if baseline was abnormal
Corrected Calcium (mmol/L)	Increase	>ULN - ≤2.9	>2.9 - 3.1	>3.1 - 3.4	>3.4

Chemistry Tests		Criteria			
Test	**Direction**	**1**	**2**	**3**	**4**
Corrected Calcium (mmol/L)	Decrease	<LLN – 2.0	<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	125 - 129	120 - 124	<120