Janssen Research & Development *

Statistical Analysis Plan
For the Maintenance Study

A Phase 3, Randomized, Double-blind, Placebo-controlled, Parallel-group, Multicenter Protocol to Evaluate the Safety and Efficacy of Ustekinumab Induction and Maintenance Therapy in Subjects with Moderately to Severely Active Ulcerative Colitis

UNIFI

Protocol CNTO1275UCO3001; Phase 3
AMENDMENT 1

CNTO1275 (CNTO1275)

Status:	Approved
Date:	25 July 2018
Prepared by:	Janssen Research & Development, LLC
Document No.:	EDMS-ERI-128681086

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

AMENDMENT HISTORY

This analysis plan was modified to implement the following changes:

1) The following new analyses have been added:

   a) Major secondary endpoints will be evaluated for the following subgroups:

      CRP and Fecal Biomarkers (calprotecin and lactoferrin), both at Week 0 of the induction study and at Week 0 of the maintenance study

      - CRP (≤3 mg/L, >3 mg/L)
      - CRP (≤ median, > median)
      - CRP (≤10 mg/L, >10 mg/L)
      - Fecal calprotectin (≤250 mg/kg, >250 mg/kg)
      - Fecal lactoferrin (≤7.24 µg/g, >7.24 µg/g)

      Induction baseline Weight (≤ 1st quartile, >1st quartile and ≤ 2nd quartile, >2nd quartile and ≤ 3rd quartile, >3rd quartile)

   b) Apart from those subgroup analyses by biologic failure status that were prespecified for the primary and the first 2 major secondary endpoints, subgroup analyses based on biologic failure status will also be performed for the 3rd and 4th major secondary endpoints as well as for the mucosal healing endpoint.

   c) Primary and major secondary endpoints as well as mucosal healing endpoint will be evaluated based on the following biologic failure profiles: subjects who were biologic-naïve, subjects who were biologic failures to at least 1 anti-TNF (regardless of vedolizumab), and subjects who were biologic failures to both anti-TNF and vedolizumab. These subpopulations were identified as they represent the general treatment pathway for patients with UC and indicate progressively more refractory disease.

   d) Sensitivity analyses will be performed for mucosal healing at Week 44 and histologic healing at Week 44 by considering subjects with an unevaluable biopsy at Week 44 as nonresponders.

   e) The following IBDQ endpoints will be analyzed:

      - A >20-point improvement from induction baseline in the IBDQ score at Week 44.
      - A ≥16-point improvement from induction baseline in the IBDQ score at Week 44.
      - IBDQ remission at Week 44
      - Maintenance of IBDQ remission among subjects with IBDQ remission at the maintenance baseline.

f) The following endpoints will be analyzed based on the endoscopy subscore assigned by the local endoscopist:

- Normal or inactive mucosal disease at Week 44
- Mucosal healing at Week 44

g) The following endpoints will be summarized based on the Non-randomized Analysis Set.

- Corticosteroid-free clinical remission at Week 44
- Mucosal healing at Week 44

In addition, maintenance of clinical response through Week 44, instead of clinical response at Week 44, will be summarized based on the Non-randomized Analysis Set

h) The following endpoints will be analyzed:

- Subjects with a UC-related hospitalization and/or surgery through Week 44
- The time to the first UC-related hospitalization or surgery through Week 44

2) The definition of rescue medication was clarified.

3) Minor errors were noted and corrected.

TABLE OF CONTENTS

LIST OF IN-TEXT TABLES

LIST OF IN-TEXT FIGURES

ABBREVIATIONS

6-MP	6-mercaptopurine
AE	adverse event
ALT	alanine aminotransferase
ANCOVA	analysis of covariance
AST	aspartate aminotransferase
ATC	Anatomical Therapeutic Chemical Classification
AZA	Azathioprine
CI	confidence interval
CMH	Cochran-Mantel-Haenszel
CRF	case report form
CRP	C-reactive protein
CSR	Clinical Study Report
DBL	database lock
DMC	Data Monitoring Committee
ECG	Electrocardiogram
eCRF	electronic case report form
EQ-5D	EuroQoL-5D Health Questionnaire
EQ-VAS	EQ Visual Analogue scale
FDA	Food and Drug Administration
IBDQ	Inflammatory Bowel Disease Questionnaire
ICH	International Conference on Harmonization
IEC	Independent Ethics Committee
IRB	Institutional Review Board
ITT	Intent-to-Treat
IV	Intravenous
IWRS	interactive web response system
LOCF	last observation carried forward
LTE	long term extension
MAR	missing at random
MCS	mental component summary
MedDRA	Medical Dictionary for Regulatory Activities
MTX	methotrexate
NCI-CTC	National Cancer Institute's Common Terminology Criteria
PCS	physical component summary
PD	Pharmacodynamic
PI	principal investigator
PK	pharmacokinetic(s)
SAE	serious adverse event
SAP	Statistical Analysis Plan
SC	Subcutaneous
SD	standard deviation
SF-36	36-item Short Form Health Survey
SUA	serious unexpected associated adverse reaction
TNF	more tumor necrosis factor
UC	ulcerative colitis
UCEIS	Ulcerative Colitis Endoscopic Index of Severity
WHO	World Health Organization
WPAI-GH	Work Productivity and Activity Impairment Questionnaire-General Health

1. INTRODUCTION

This Statistical Analysis Plan contains definitions of analysis sets, derived variables, and statistical methods for all planned analyses of the maintenance study in protocol CNTO1275UCO3001. The analyses for the induction study in this protocol will be provided in a separate induction SAP.

1.1. Trial Objectives

The primary objectives of the maintenance study in protocol CNTO1275UCO3001 are:

- To evaluate clinical remission for subcutaneous (SC) maintenance regimens of ustekinumab in subjects with moderately to severely active ulcerative colitis (UC) induced into clinical response with ustekinumab.
- To evaluate the safety of SC maintenance regimens of ustekinumab in subjects with moderately to severely active UC induced into clinical response with ustekinumab.

The secondary objectives are:

- To evaluate the efficacy of ustekinumab in maintaining clinical response in subjects induced into clinical response with ustekinumab.
- To evaluate endoscopic healing (i.e., improvement in the endoscopic appearance of the mucosa) in subjects induced into clinical response with ustekinumab.
- To evaluate the efficacy of ustekinumab in achieving corticosteroid-free clinical remission in subjects induced into clinical response with ustekinumab.
- To evaluate the efficacy of ustekinumab in maintaining clinical remission in subjects induced into clinical remission with ustekinumab.
- To evaluate the efficacy of ustekinumab treatment on mucosal healing (i.e., endoscopic healing and histologic healing).
- To evaluate the impact of SC ustekinumab on disease-specific health-related quality of life.
- To evaluate the efficacy of maintenance therapy with SC ustekinumab by biologic failure status.
- To evaluate the pharmacokinetics (PK), immunogenicity, and pharmacodynamics (PD) of ustekinumab induction therapy in subjects with moderately to severely active UC, including changes in C-reactive protein (CRP), fecal calprotectin, fecal lactoferrin, and other PD biomarkers.

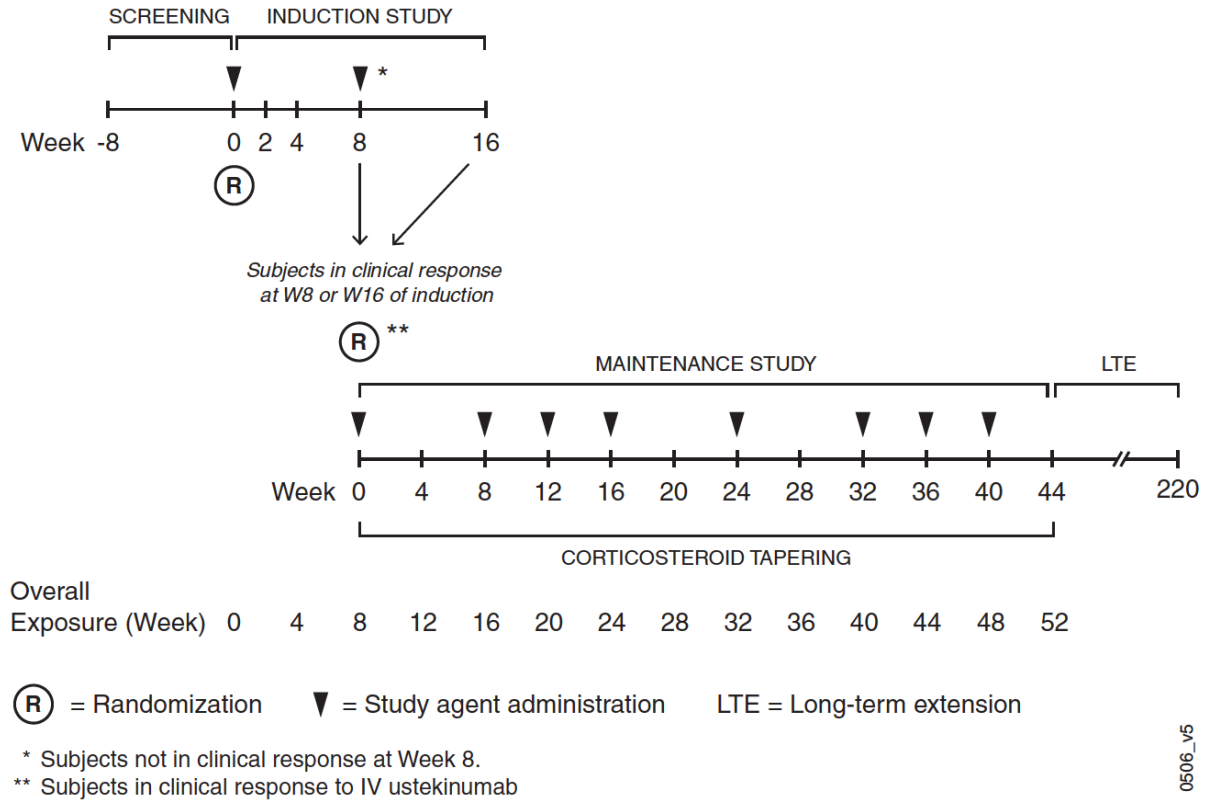
1.2. Trial Design

The Phase 3 development program for ustekinumab in the treatment of UC will be conducted under a single protocol but will be designed and analyzed as 2 separate studies, an induction study and a maintenance study. Both studies will be Phase 3, randomized, double-blind, placebo-controlled, parallel-group, multicenter studies of ustekinumab in subjects with moderately to severely active UC. The induction study will target subjects with moderately to severely active UC who have demonstrated an inadequate response or failure to tolerate conventional or biologic

therapy. The maintenance study will be a randomized withdrawal study targeting subjects with moderately to severely active UC who have demonstrated a clinical response to induction treatment with IV ustekinumab. Overall, the program will evaluate ustekinumab treatment in subjects with moderately to severely active UC through at least 1 year of induction and maintenance therapy; after completion of the maintenance study through Week 44, a long-term extension study will follow eligible subjects for an additional 3 years.

Overall schema for CNTO1275UCO3001 protocol is shown in Figure 1.

Figure 1: Overall schema for CNTO1275UCO3001 protocol



The target population is men or women 18 years of age or older with moderately to severely active UC, as defined by a Mayo score of 6 to 12, inclusive, at Week 0 of the induction study, including an endoscopy subscore ≥2 as assessed by the central reader. Subjects must not be at imminent risk of colectomy.

A broad subject population will be evaluated and will include both subjects who have failed biologic therapy and those who have not:

- Subjects may be biologic failures, i.e., have received treatment with 1 or more tumor necrosis factor (TNF) antagonists or vedolizumab at a dose approved for the treatment of UC, and either did not respond initially, responded initially but then lost response, or were intolerant to the medication. A minimum of 40% and a maximum of 50% of the total subject population in the induction study will be biologic failures.
- Subjects may be biologic-naive or may have been exposed to biologic therapy but not demonstrated an inadequate response or intolerance to treatment with a biologic agent (i.e., a TNF antagonist or vedolizumab). These subjects must have demonstrated an inadequate response to, or have failed to tolerate, at least 1 of the following conventional UC therapies: oral or intravenous (IV) corticosteroids or the immunomodulators azathioprine (AZA) or 6-mercaptopurine (6-MP). Subjects who have demonstrated corticosteroid dependence (i.e., an inability to successfully taper corticosteroids without a return of the symptoms of UC) are also eligible for entry into the study.

Induction Study

In the induction study, a target of 951 subjects will be randomized in a 1:1:1 ratio to receive 1 of the following 3 treatment regimens at Week 0 (I-0): Placebo IV, Ustekinumab 130 mg IV, and Weight-range-based ustekinumab doses approximating ustekinumab 6 mg/kg IV (i.e., ustekinumab ~6 mg/kg IV). At Week 8 (I-8), subjects who are in clinical response are eligible to enter the maintenance study; subjects who are not in clinical response will receive ustekinumab as follows:

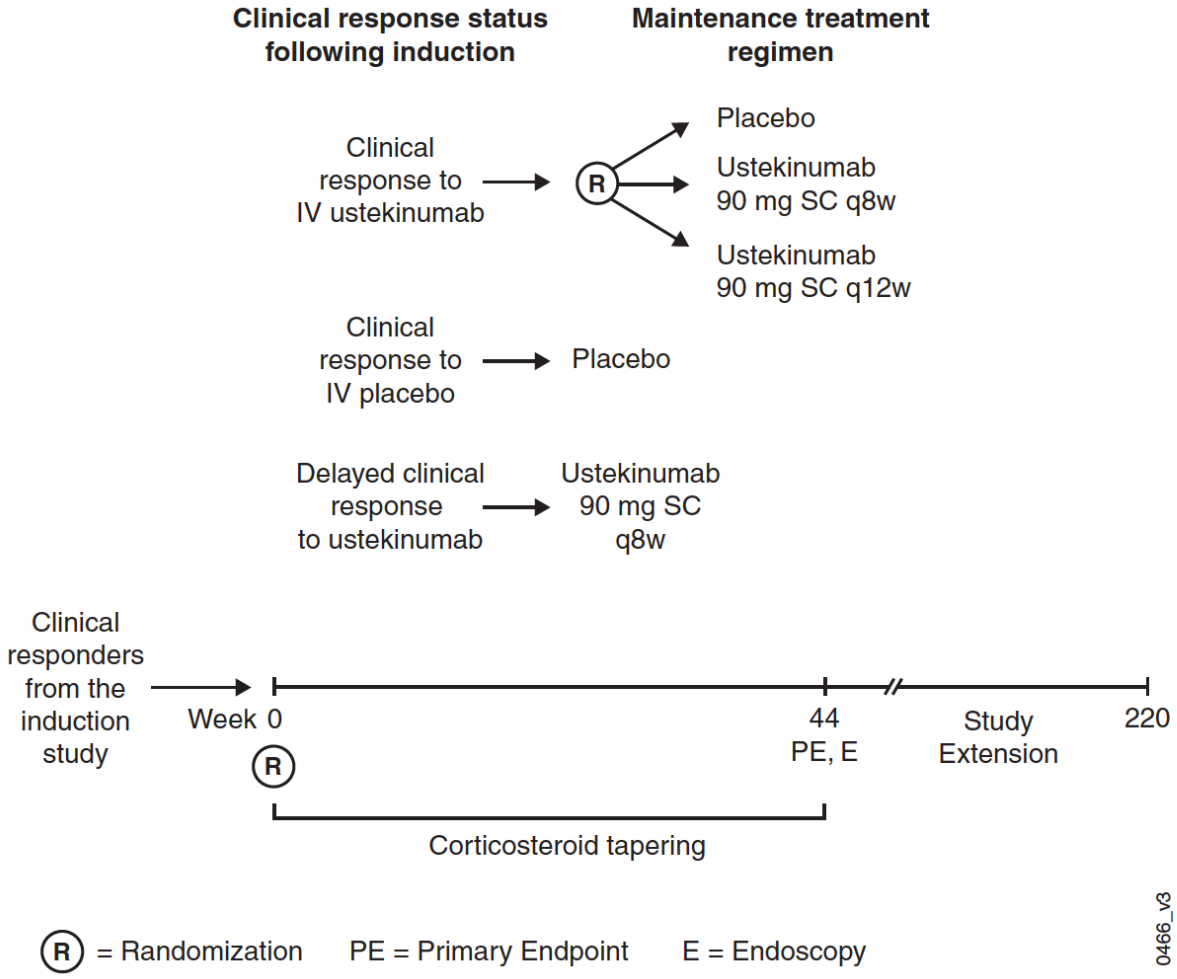
- Subjects who are randomized to placebo at Week 0 will receive 1 dose of ustekinumab ~6 mg/kg IV plus placebo SC (to maintain the blind) at I-8.
- Subjects who are randomized to ustekinumab at Week 0 will receive 1 dose of ustekinumab 90 mg SC plus placebo IV (to maintain the blind) at I-8.

At Week 16 of the induction study, the subjects who are not in clinical response at I- 8 will be re-evaluated for clinical response. Subjects who achieve clinical response at I-16 are eligible to enter the maintenance study. All subjects in the induction study who do not enter the maintenance study will have a safety follow-up visit approximately 20 weeks after their last administration of study agent

Maintenance Study

In the randomized-withdrawal maintenance study, all subjects enrolled will be responders to study agent administered in the induction study. The schema for the maintenance study is shown in Figure 2.

Figure 2: Schema for the CNTO1275UCO3001 maintenance study



Subjects who are in clinical response to IV ustekinumab during induction will comprise the primary population in the maintenance study; this population will include the following:

- Subjects who are randomized to receive ustekinumab at Week 0 of the induction study and are in clinical response at I-8.
- Subjects who are randomized to receive placebo at Week 0 of the induction study and are not in clinical response at I-8, but are in clinical response at I-16 after receiving an induction dose of IV ustekinumab (~6 mg/kg) at I-8.

Note that subjects who are delayed clinical responders to ustekinumab induction (i.e., the subjects who are randomized to ustekinumab at Week 0 of the induction study, not in clinical response at I- 8 but are in clinical response at I- 16 after receiving 1 dose of ustekinumab 90 mg SC at I-8) will not be included in the primary population.

A target of 327 subjects who are in clinical response to IV ustekinumab induction will be randomized in a 1:1:1 ratio to 1 of 3 treatment groups at the maintenance Week 0/baseline (M-0) visit of the maintenance study:

- Placebo SC.
- Ustekinumab 90 mg SC every 12 weeks (q12w).
- Ustekinumab 90 mg SC q8w.

Eligible subjects will be allocated to a treatment group using a permuted block randomization with clinical remission status at maintenance baseline (yes/no), oral corticosteroid use at maintenance baseline (yes/no), and induction treatment (placebo IV [I-0]→ ustekinumab ~6 mg/kg IV [I-8], ustekinumab 130 mg IV [I-0], or ustekinumab ~6 mg/kg IV [I-0]) as stratification variables.

Additional subjects entering the maintenance study will include the following; these subjects will not be part of the primary population:

- Subjects who are in clinical response to placebo IV induction will receive placebo SC.
- Subjects who are delayed responders to ustekinumab induction will receive ustekinumab 90 mg SC q8w.

All subjects will receive their assigned dose of SC study agent at the M-0 visit. Thereafter, to maintain the blind, all subjects will receive study agent at all scheduled study agent administration visits. Subjects will be assessed for clinical flare at every visit and, if loss of response is confirmed (based on the Mayo score that includes the endoscopy subscore assigned by the local endoscopist), may be eligible for rescue medication. See the protocol for more details.

Concomitant medical therapy for UC must have been stable from the I-0 visit through the M-0 visit unless, in the judgment of the investigator, the therapy had to be discontinued or reduced in dose because of toxicity or medical necessity. Subjects who initiated or increased the dose of a UC-specific medication (or any restricted/prohibited medication) during the induction study are prohibited from entering the maintenance study.

With the exception of corticosteroids, which should be tapered, UC-specific medical therapies (i.e., oral 5-ASA compounds, or the immunomodulators 6-MP, AZA, or methotrexate [MTX]) must be maintained at stable doses through M-44 unless investigator judgment requires that the therapy be discontinued or the dose reduced because of toxicity or medical necessity, or unless there is a documented loss of response that makes the subject eligible for rescue medication. Corticosteroids will be tapered beginning at the M-0 visit for all subjects who enter the maintenance study.

The primary endpoint of the maintenance study is clinical remission at Week 44. The major secondary endpoints are (presented in the order in which they will be tested): maintenance of clinical response through Week 44; endoscopic healing at Week 44; clinical remission and not

receiving concomitant corticosteroids at Week 44; and maintenance of clinical remission through Week 44 among the subjects who had achieved clinical remission at maintenance baseline.

The primary database lock (DBL) will occur when all subjects in the maintenance study have either completed the M-44 visit or have terminated study participation before M-44. No interim analysis is planned during the maintenance study.

Long Term Extension

Subjects who complete the safety and efficacy evaluations at Week 44 and who may benefit from continued treatment, in the opinion of the investigator, will have the opportunity to participate in the long term extension (LTE). The LTE begins after the assessments listed for the M-44 visit (i.e., the last assessments in the maintenance study) have been completed and will continue through Week 220 or until the sponsor decides not to pursue an indication in UC, whichever occurs first.

Subjects will continue to receive the same treatment regimen during the LTE that they were receiving at the end of the maintenance study (either placebo or ustekinumab 90 mg SC q8w or q12w), with the first dose in the LTE being administered at Week 48. The study blind will be maintained during the LTE until the last subject in the maintenance study has completed the M-44 visit evaluations and the Week 44 analyses have been completed. After the study is unblinded to the investigative sites, subjects receiving placebo will be terminated from study participation, and subjects receiving ustekinumab will continue to receive ustekinumab, but will have their study visits scheduled to coincide with their dose regimen (either q8w or q12w, as appropriate for their dose regimen).

During the LTE, all subjects will be assessed for worsening of disease activity based on the clinical judgment of the investigator. Subjects in the primary population whose UC disease activity worsens will be eligible for a dose adjustment as follows:

- Placebo SC → ustekinumab 90 mg SC q8w
- Ustekinumab 90 mg SC q12w → ustekinumab 90 mg SC q8w
- Ustekinumab 90 mg SC q8w → continue on ustekinumab 90 mg SC q8w

The first visit at which a subject can be considered for a dose adjustment is at Week 56. Subjects will be allowed 1 dose adjustment during the LTE. The interactive web response system (IWRS) will ensure that SC ustekinumab is not administered more frequently than q8w.

Subjects who are not in the primary population (i.e., induction placebo responders, delayed ustekinumab responders) will not be eligible for a dose adjustment during the LTE.

Any subject who has not shown improvement in his or her UC disease activity by 16 weeks after worsening of their UC disease activity (as assessed by the investigator) will be discontinued from further study agent administration.

During the LTE, all concomitant medications, including UC-specific medications (with the exception of the prohibited medications listed in Section 8.1.1.1.2 of the protocol), may be administered at the discretion of the investigator.

Efficacy evaluations during the LTE will generally be based on the partial Mayo score, markers of inflammation (i.e. CRP, fecal lactoferrin and calprotectin), and corticosteroid use. Selected patient-reported outcomes (PRO) and health economics data will also be collected. All study evaluations to be performed during the LTE are listed in the Time and Events Schedule of the protocol.

The final DBL will occur when all subjects have either completed the M-220 visit or have terminated study participation before M-220. One or more additional DBLs may occur between Week 44 and Week 220 for publication purposes or for regulatory requirements.

Unless stated otherwise, the procedures in the following sections are all described in the context of the maintenance study, and the analysis time points are in reference to Week 0 of the maintenance study. For example, Week 44 of the endpoints refers to 44 weeks after Week 0 of the maintenance study, *not* 44 weeks after Week 0 of the induction study. Week X of the induction study will be denoted as I-X.

1.3. Statistical Hypotheses for Trial Objectives

The primary endpoint is clinical remission at Week 44 among subjects with moderately to severely active UC who were responders to IV ustekinumab induction. The null hypothesis to be tested is that there is no difference between ustekinumab maintenance therapy (90 mg SC q8w or 90 mg SC q12w) and placebo in achieving clinical remission at Week 44.

1.4. Randomization and Blinding

Randomization and Treatment Allocation

Treatment allocation of this study will be performed using a centralized IWRS. At Week 0 upon a subject's entry into the maintenance study, subjects in the primary population (i.e., the subjects who were in clinical response to IV ustekinumab during the induction) will be randomized to 1 of 3 treatment groups (placebo SC, ustekinumab 90 mg SC q8w, or ustekinumab 90 mg SC q12w) using a permuted block randomization with clinical remission status at maintenance baseline (yes/no), oral corticosteroid use at maintenance baseline (yes/no), and induction treatment (placebo IV [I-0]→ ustekinumab ~6 mg/kg IV [I-8], ustekinumab 130 mg IV [I-0], or ustekinumab ~6 mg/kg IV [I-0]) as stratification variables. Other subject populations will not be randomized, but will be assigned to treatment at Week 0 upon a subject's entry into the maintenance study as described in Section 1.2. A computer-generated randomization schedule will be prepared under the supervision of the sponsor.

At each call to the IWRS for a treatment allocation, the IWRS will assign a treatment code that will dictate the treatment assignment and matching study agent kit for each subject. The requestor must use his or her own personal identification number when contacting the IWRS.

The values of the stratification factors in the IWRS that are used for the Week 0 randomization for a subject will not be changed once the subject is randomized. Dose adjustment in the LTE is based on the status of worsening of UC disease activity assessed by the investigator. The status of worsening of UC disease activity cannot be changed once the subject has requested a dose adjustment.

Blinding

The sponsor will be blinded to maintenance treatment assignment until after the Week 44 database lock has occurred. Treatment assignment blinding will be maintained for investigative sites, site monitors, and subjects participating in this protocol until the Week 44 analyses for the maintenance study have been completed.

To maintain the study blind, the study agent container will have a label containing the study name, medication number, and reference number. A tear-off label is designed to be separated from the study agent container, and attached to the subject's source documents. The label will not identify the study agent in the container. The medication number will be entered in the case report form (CRF) when the drug is dispensed. The study agents will be identical in appearance and packaging.

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 subject. The date and reason for the unblinding will be documented in the appropriate section of the CRF and in the source document. The documentation received from the IWRS indicating the code break will be retained with the subject's source documents in a secure manner. Subjects who had their treatment assignment unblinded by the investigator will be discontinued from study agent.

Data that may potentially unblind the treatment assignment (i.e., study agent serum concentrations, antibodies to study agent, and treatment allocation) will be handled through the Secure Data Office to ensure that the integrity of the blind is maintained and the potential for bias is minimized.

The postbaseline results of CRP, fecal lactoferrin, and fecal calprotectin tests performed by the central laboratory will be blinded to the investigative sites. If an investigative site requests these data, it will be provided to them after the Week 44 analyses for the maintenance study have been completed.

Additionally, a given subject’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 serious unexpected associated adverse reactions (SUAs).

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

An independent DMC will regularly review unblinded safety data. The same DMC will also review the interim analysis results from the induction study and form a recommendation to the sponsor per the pre-specified decision rules detailed in the Interim Analysis Plan. The committee and its statistical support group will have access to unblinded treatment information. They will not divulge any information to the Sponsor that may potentially unblind an individual subject's treatment group.

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. Baseline Measurements

Unless otherwise specified, the baseline measurement for each parameter is defined as the closest measurement taken prior to or on the same day as the Week 0 study agent administration in the maintenance study.

2.3. Analysis Sets

2.3.1. Efficacy Analysis Sets

Unless otherwise specified, all efficacy analyses will be based on the intent-to-treat (ITT) principle. Therefore, subjects will be analyzed according to the treatment group to which they were assigned regardless of the treatment they actually received.

The **Primary Efficacy Analysis Set** consists of all subjects randomized at Week 0 of this maintenance study, that is, subjects in clinical response to IV ustekinumab induction as determined by the IWRS (i.e., subjects who are in clinical response to IV ustekinumab induction at I-8, or subjects who are not in clinical response to IV placebo induction at I-8 but are in clinical response at I-16 after receiving an induction dose of IV ustekinumab at I-8). Some prespecified efficacy analyses will also be conducted in the **Non-randomized Analysis Set**, which includes:

- Subjects in clinical response to placebo IV induction dosing at I-8.
- Subjects who are delayed responders to ustekinumab induction.

2.3.2. Safety Analysis Set

All subjects who receive at least 1 administration of study agent in this maintenance study will be included in the **Safety Analysis Set**, which includes subjects who are randomized and subjects who are not randomized. Some prespecified safety analyses will also be conducted in the **Randomized Safety Analysis Set**, which includes subjects in the **Safety Analysis Set** who are randomized at Week 0 of this maintenance study. Subjects will be analyzed according to the treatment group to which they were assigned regardless of the treatment they actually received.

2.3.3. Pharmacokinetics Analysis Set

PK summaries will be provided for the **PK Analysis Set** (i.e., subjects who receive at least 1 administration of study agent [either ustekinumab or placebo] in this maintenance study **and** receive at least one administration of ustekinumab in the induction study). Some prespecified PK analyses will also be conducted in the **Randomized PK Analysis Set**, which includes subjects in the **PK Analysis Set** who are randomized at Week 0 of this maintenance study. Subjects will be analyzed according to the treatment group to which they were assigned regardless of the treatment they actually received.

2.4. Definition of Subgroups

The consistency of treatment effect for 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, all at Week 0 of the induction study, as well as maintenance stratification factors and UC clinical disease characteristics at Week 0 of the maintenance study, when the number of subjects within each level of the subgroup permits.

Demographics at Week 0 of the induction study

- Gender (male, female)
- Race (Caucasian, non-Caucasian)
- Age (≤ median age, > median age)
- Weight (≤ 1st quartile, >1st quartile and ≤ 2nd quartile, >2nd quartile and ≤ 3rd quartile, >3rd quartile)
- Smoking status (nonsmoker, prior smoker, current smoker)
- Region
  - Asia: Japan, South Korea
  - Eastern Europe: Belarus, Bulgaria, Czech Republic, Hungary, Poland, Romania, Russia, Serbia, Slovakia, Ukraine
  - Rest of World: Australia, Austria, Belgium, Canada, Denmark, France, Germany, Israel, Italy, Netherlands, New Zealand, United Kingdom, United States

Baseline UC clinical disease characteristics at Week 0 of the induction study

- UC disease duration (≤ 5 years, > 5 years to ≤ 15 years, > 15 years)
- Extent of disease (limited, extensive)
- Severity of UC disease (moderate: 6≤ Mayo score ≤ 10, severe: Mayo score >10)
- Extraintestinal manifestations (absent, present)
- CRP (≤3 mg/L, >3 mg/L)
- CRP (≤ median, > median)
- CRP (≤10 mg/L, >10 mg/L)

- Fecal calprotectin (≤250 mg/kg, >250 mg/kg)
- Fecal lactoferrin (≤7.24 µg/g, >7.24 µg/g)

UC-related concomitant medications at Week 0 of the induction study

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

UC-related medication history at Week 0 of the induction study

- 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)
- Biologic failure status (yes, no)
- Subjects with biologic failure
  - Primary nonresponse, secondary nonresponse, or intolerance to
    - At least one anti-TNF (yes, no)
    - Anti-TNF only (yes, no)
    - Vedolizumab (yes, no)
    - Vedolizumab and at least one anti-TNF (yes, no)
  - For subjects with biologic failure to anti-TNF only
    - Primary nonresponse (yes, no)
    - Secondary nonresponse (yes, no)
    - Intolerance (yes, no)
- Subjects without biologic failure (naïve, bio-experienced [but not documented failure])

Stratification variables for the maintenance study

- Clinical remission status at maintenance baseline as determined by the IWRS (yes, no)
- Induction treatment (placebo IV [I-0] → ustekinumab ~6 mg/kg IV [I-8], ustekinumab 130 mg IV [I-0], ustekinumab ~6 mg/kg IV [I-0])
- Oral corticosteroid use at maintenance baseline as recorded in the IWRS (yes, no)

UC clinical disease characteristics at Week 0 of the maintenance study

- Endoscopic healing status (yes, no)
- CRP (≤3 mg/L, >3 mg/L)
- CRP (≤ median, > median)
- CRP (≤10 mg/L, >10 mg/L)
- Fecal calprotectin (≤250 mg/kg, >250 mg/kg)
- Fecal lactoferrin (≤7.24 µg/g, >7.24 µg/g)

In addition, the consistency of treatment effect for the major secondary endpoints will be evaluated for the following subgroups, when the number of subjects within each level of the subgroup permits.

CRP and Fecal Biomarkers (calprotecin and lactoferrin), both at Week 0 of the induction study and at Week 0 of the maintenance study

- CRP (≤3 mg/L, >3 mg/L)
- CRP (≤ median, > median)
- CRP (≤10 mg/L, >10 mg/L)
- Fecal calprotectin (≤250 mg/kg, >250 mg/kg)
- Fecal lactoferrin (≤7.24 µg/g, >7.24 µg/g)

Induction baseline Weight (≤ 1st quartile, >1st quartile and ≤ 2nd quartile, >2nd quartile and ≤ 3rd quartile, >3rd quartile)

3. INTERIM ANALYSIS AND DATA MONITORING COMMITTEE REVIEW

No interim analysis is planned for the maintenance study.

A DMC will be established to monitor data on an ongoing basis to ensure the continuing safety of the subjects enrolled in this protocol. As mentioned in Section 1.4, the committee will meet periodically to review unblinded safety data. The same DMC will also review the interim analysis results from the induction study and form a recommendation to the sponsor per the pre-specified decision rules detailed in the Interim Analysis Plan. The DMC roles and responsibilities, and the general procedures of the DMC review are defined and documented in the DMC Charter.

4. SUBJECT INFORMATION

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

4.1. Demographics and Baseline Characteristics

To assess the comparability of subjects, the following data will be summarized by treatment group for all subjects enrolled in the maintenance study:

- Demographic data at I-0 (i.e., age, gender [male, female], race [White, Black or African American, Asian, American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, Other], weight, height, and region [Asia, Eastern Europe, Rest of World])
- Disease characteristics at I-0 (i.e., UC disease duration, extent of disease [limited, extensive], Mayo score, severity of UC disease [moderate: 6≤ Mayo score ≤10, severe: Mayo score >10], extraintestinal manifestations [absent, present], CRP, fecal calprotectin, fecal lactoferrin, abnormal CRP (>3 mg/L), abnormal fecal calprotectin (>250 mg/kg), abnormal fecal lactoferrin (>7.24 µg/g)
- Smoking status at I-0 (nonsmoker, prior smoker, current smoker)
- Past medical history and current diagnoses at I-0
- Disease characteristics at M-0 (i.e., Mayo score, CRP, abnormal CRP [>3 mg/L], fecal calprotectin, abnormal fecal calprotectin (>250 mg/kg), fecal lactoferrin, abnormal fecal lactoferrin (>7.24 µg/g), clinical remission status [yes, no], endoscopic healing [yes, no])

4.2. Prior and Concomitant Medications

Summaries of UC medication history (i.e., subjects who took medications for UC and their length of exposure prior to I-0), UC-related concomitant medications at I-0, UC-related non-biologic medication history at I-0 (i.e., history of response to or tolerance of 6-MP/AZA and history of response to, dependence or tolerance of corticosteroids) and UC-related biologic medication history at I-0 will be provided for all subjects enrolled in the maintenance study.

4.3. Disposition Information

Subjects who discontinued study agent prior to Week 44 along with the reasons for discontinuation, and subjects who terminated study participation prior to Week 44 along with the reasons for termination, will be summarized by treatment group for all subjects enrolled in the maintenance study. In addition, listings of the subjects who discontinued study agent and the subjects who terminated study participation prior to Week 44 will be provided.

4.4. Treatment Compliance

The number of subjects receiving each scheduled administration through Week 44 will be summarized for all subjects enrolled in the maintenance study. In addition, a listing of subjects who were assigned treatment but were never treated and a listing of subjects who were unblinded prior to Week 44 during the maintenance study will be provided.

4.5. Extent of Exposure

The cumulative dose of ustekinumab received will be summarized by treatment group through Week 44 based on the **Safety Analysis Set**. In addition, the number of administrations of study agent and the average duration of follow-up (weeks) will be summarized through Week 44 by treatment group for all subjects in the **Safety Analysis Set** as part of the safety tables.

4.6. Protocol Deviations

Subjects with a major protocol deviation through Week 44 will be summarized for all subjects enrolled in the maintenance study. In addition, subjects 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). A listing of subjects who have a major protocol deviation and a more detailed listing of subjects who do not meet study entry criteria will be provided.

5. EFFICACY

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

Analyses suitable for categorical data (e.g., chi-square tests, Cochran-Mantel-Haenszel chi-square tests, or logistic regression, as appropriate) will be used to compare the proportions of subjects achieving selected endpoints (e.g., clinical remission). In cases of rare events, Fisher's exact test will be used for treatment comparisons. Continuous response parameters will be compared using an analysis of variance (ANOVA) or covariance (ANCOVA), unless otherwise specified. If the normality assumption is in question, an ANOVA or ANCOVA on the van der Waerden normal scores will be used.

Nominal p-values will be presented for all analyses unless specified otherwise.

5.1. Analysis Specifications

5.1.1. Data Handling Rules

5.1.1.1. Treatment Failure

Subjects who have any of the following events are considered to be a treatment failure from the time of event onward:

- An ostomy or colectomy (partial or total)
- Discontinue study agent due to lack of therapeutic effect or due to an AE of worsening of UC
- Use a rescue medication after clinical flare as defined in protocol Section 3.1.2.1
- A prohibited change in UC medication:
  - Initiation of restricted or prohibited medications or therapies as defined in the protocol (see protocol Section 8.1.1.1), except for antibiotics used to treat UC, TPN, and apheresis.
  - Increase in the dose of oral corticosteroids (excluding budesonide and beclomethasone dipropionate) > 5 mg/day (prednisone equivalent) above the baseline dose (Week 0 of the maintenance study) for more than 7 days after the Week 32 visit (i.e., approximately 90 days prior to Week 44) due to worsening of disease. This includes initiation of oral

corticosteroids due to worsening of disease that lasts for more than 7 days after the Week 32 visit for subjects who were not receiving oral corticosteroids at Week 0 of the maintenance study.

- Increase in the dose of oral budesonide > 3 mg/day above the baseline dose (Week 0 of the maintenance study) for more than 7 days after the Week 32 visit (i.e., approximately 90 days prior to Week 44) due to worsening of disease. This includes initiation of oral budesonide due to worsening of disease that lasts for more than 7 days after the Week 32 visit for subjects who were not receiving oral budesonide at Week 0 of the maintenance study.
- Increase in the dose of oral beclomethasone dipropionate > 5 mg/day above the baseline dose (Week 0 of the maintenance study) for more than 7 days after the Week 32 visit (i.e., approximately 90 days prior to Week 44) due to worsening of disease. This includes initiation of oral beclomethasone dipropionate due to worsening of disease that lasts for more than 7 days after the Week 32 visit for subjects who were not receiving oral beclomethasone dipropionate at Week 0 of the maintenance study.
- Increase in the dose of oral corticosteroids (excluding budesonide and beclomethasone dipropionate) > 5 mg/day (prednisone equivalent) above the baseline dose (Week 0 of the maintenance study) for more than 28 days after the Week 32 visit (i.e., approximately 90 days prior to Week 44) due to reasons other than worsening of disease. This includes initiation of oral corticosteroids due to reasons other than worsening of disease that lasts for more than 28 days after the Week 32 visit for subjects who were not receiving oral corticosteroids at Week 0 of the maintenance study.
- Increase in the dose of oral budesonide > 3 mg/day (prednisone equivalent) above the baseline dose (Week 0 of the maintenance study) for more than 28 days after the Week 32 visit (i.e., approximately 90 days prior to Week 44) due to reasons other than worsening of disease. This includes initiation of oral budesonide due to reasons other than worsening of disease that lasts for more than 28 days after the Week 32 visit for subjects who were not receiving oral budesonide at Week 0 of the maintenance study.
- Increase in the dose of oral beclomethasone > 5 mg/day (prednisone equivalent) above the baseline dose (Week 0 of the maintenance study) for more than 28 days after the Week 32 visit (i.e., approximately 90 days prior to Week 44) due to reasons other than worsening of disease. This includes initiation of oral beclomethasone dipropionate due to reasons other than worsening of disease that lasts for more than 28 days after the Week 32 visit for subjects who were not receiving oral beclomethasone dipropionate at Week 0 of the maintenance study.
- Any switch among oral budesonide, oral beclomethasone dipropionate or other oral corticosteroids (excluding prednisone equivalent changes) due to worsening of disease.
- Initiation of oral 5-ASA compounds due to worsening of disease.
- Increase above baseline in the dosage of oral 5-ASA compounds due to worsening of disease.
- Change from one oral 5-ASA compound to another 5-ASA compound due to worsening of disease.

- Initiation of 6-MP/AZA/MTX due to worsening of disease.
- Increase above baseline in the dosage of 6-MP/AZA/MTX due to worsening of disease.
- Any switch between 6-MP/AZA and MTX due to worsening of disease.

Following a clinical flare, any initiation or increase in the dose of 6-MP/AZA/MTX, oral budesonide > 3 mg/day, oral beclomethasone dipropionate > 5 mg/day, or oral corticosteroids (excluding budesonide and beclomethasone dipropionate) > 5 mg/day (prednisone equivalent) above the dose received at maintenance baseline was considered a rescue medication.

For dichotomous endpoints, subjects who have a treatment failure are considered as not achieving the respective endpoints. For continuous endpoints, subjects who have a treatment failure have their induction baseline values carried forward from the time of the treatment failure onwards.

5.1.1.2. Missing Data

For subjects with missing data, unless otherwise specified, the last observation will be carried forward for continuous endpoints, with the exception of the Mayo and partial Mayo scores (defined in Section 5.2.1.1), where the last available Mayo subscores will be carried forward. For dichotomous endpoints, subjects with missing data will be considered as not achieving the respective endpoints. More detailed descriptions of missing data for the Mayo score, IBDQ and SF-36 will be provided later in the document. **Note that treatment failure rules override missing data rules.** This means that if a subject has an event of treatment failure, induction baseline values will be assigned from the point of treatment failure onward for continuous endpoints, and subjects will be considered as not achieving the respective endpoints for dichotomous endpoints, regardless of whether the data were observed or missing.

5.2. Primary Efficacy Endpoint

The primary endpoint is clinical remission at Week 44, which is derived from the Mayo score.

5.2.1. Definitions

5.2.1.1. Mayo Score

The Mayo score (Schroeder et al, 1987)[8] was developed from the criteria of Truelove and Witts (1955)[10] for mild, moderate, and severe UC; and from the criteria of Baron et al (1964)[1] for grading the mucosal 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 of stool frequency, rectal bleeding, physician's global assessment, and the findings of endoscopy. Thus, the Mayo score may take on values in the range of 0 to 12. **The partial Mayo score,** which is the Mayo score without taking into account the findings of endoscopy, is calculated as the sum of the stool frequency, rectal bleeding, and physician's global assessment 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.

The Mayo score can be calculated when at least 1 of the 4 subscores is available. If 1 or more of the 4 Mayo subscores is missing at a specific visit, but not all 4 subscores are missing, the last available value for each missing subscore will be carried forward to compute a full Mayo score (where the full Mayo score is scheduled to be collected). When the Mayo score cannot be calculated (i.e., all 4 subscores are missing), then the Mayo score is considered to be missing.

Similarly, if 1 or more of the following 3 Mayo subscores (stool frequency subscore, rectal bleeding subscore, and PGA subscore) is missing at a specific visit, but not all 3 subscores are missing, the last available value for each missing subscore will be carried forward to compute a partial Mayo score. If all 3 subscores are missing at a specific visit, the partial Mayo score will be considered missing at that visit. The modified Mayo score (the sum of stool frequency subscore, rectal bleeding subscore, and endoscopy subscore) will be calculated in a similar fashion.

Stool Frequency and Rectal Bleeding Subscores

The eCRFs capture seven days of rectal bleeding data and the number of stools per day prior to each visit at which the partial Mayo score or Mayo score is collected. Sites are instructed to check the boxes next to the 3 days for which are used to calculate stool frequency and rectal bleeding subscores. **Absolute stool number** is the average of the daily stool number over the three days, and rectal bleeding subscores is calculated using the average rectal bleeding number for the three days based on the criteria in Attachment 1. 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.

Sites are directed to use the most recent 3 consecutive days within the week 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 (i.e. diphenoxylate hydrochloride with atropine sulfate or loperamide)
- The 48 hours immediately following a colonoscopy

If three consecutive days are not available, the sites are instructed to choose two consecutive days and the closest nonconsecutive day. If two consecutive days are not available, then three nonconsecutive days closest to the visit should be chosen. If 3 days (within the week 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.

Endoscopy Subscore

The endoscopic findings will be based on the criteria of the Mayo endoscopy subscore described in Attachment 1. The endoscopic subscore will be assessed by the investigator (i.e., local endoscopist) during the endoscopy procedure and by a central reader reviewing a video of the endoscopy. The central readers will be blinded to the local endoscopist subscores, treatment assignment, and study visit (except for the initial or the final endoscopy videos when the central readers know that these videos are or are not from the screening visit).

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 subscores, the video endoscopy will be submitted to a second central reader designated for adjudication. The adjudicator will be blinded to the scores of the local and first central reader. From the three Mayo endoscopic subscores, the score with which two readers agree will be reported as the final Mayo endoscopic subscore. In the unlikely event that no two readers agree on the Mayo endoscopic subscore, the median score of the three completed reads (i.e., local read, central read #1 and central read #2 designated for adjudication) will be chosen as the final reported endoscopic subscore.

The final reported endoscopic subscore at screening will be used to determine eligibility (i.e., Mayo endoscopy subscore ≥2) and to calculate the baseline Mayo score. Further details regarding video acquisition, standardization, readings, and data transfer are provided in the Imaging Charter.

Unless otherwise specified, the analysis of endpoints related to the endoscopy subscore, including the Mayo score, will be based on the final reported endoscopic subscore. If the final reported endoscopic subscore 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 also not available, then the endoscopy subscore for the analysis will be left missing.

The central reader will perform a friability assessment for each endoscopy (sigmoidoscopy or colonoscopy) received by providing a response to the following question: “Is hemorrhage detected with incidental trauma caused by the endoscopic procedure?” If friability is present, the central reader will also provide a confidence level (High or Low) when evaluating friability. Further details are provided in the Imaging Charter.

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.

5.2.1.2. Clinical Remission

The definition of the primary endpoint of clinical remission will be different for countries outside the United States and for the United States. Each definition of clinical remission will be applied to all subjects in the efficacy analysis populations.

The global definition of clinical remission is (for countries outside the United States**):** A Mayo score ≤2 points, with no individual subscore >1. In addition to the clinical remission status based on the Mayo score, treatment failure rules will be applied to determine the final clinical remission status for a subject. Subjects who have a treatment failure (see Section 5.1.1.1) prior to Week 44 will be considered not to be in clinical remission at Week 44, regardless of the actual computation of clinical remission based on the Mayo score. Subjects who have all 4 Mayo subscores missing at Week 44 will be considered not to be in clinical remission.

The US-specific definition of clinical remission is: An absolute stool number ≤3, a Mayo rectal bleeding subscore of 0, and a Mayo endoscopy subscore of 0 or 1. The Absolute stool number as defined in Section 5.2.1.1 is the average of the daily stool number over the three days. Subjects who have a treatment failure (see Section 5.1.1.1) prior to Week 44 will be considered not to be in clinical remission regardless of the actual computation of clinical remission based on the Mayo score. Subjects who are missing all three Mayo components pertaining to the primary endpoint (i.e., the absolute stool number, rectal bleeding subscore, and Mayo endoscopy subscore) at Week 44 will be considered not to be in clinical remission.

5.2.2. Analysis Methods

The primary analysis will be based on the **Primary Efficacy Analysis Set**.

The proportions of subjects in clinical remission will be compared between each ustekinumab treatment group and the placebo group using a Cochran-Mantel-Haenszel (CMH) chi-square test stratified by clinical remission status at baseline (yes/no as determined by the IWRS) and induction treatment (placebo IV [I-0] → ustekinumab ~6 mg/kg IV [I-8], ustekinumab 130 mg IV [I-0], or ustekinumab ~6 mg/kg IV [I-0]). Summaries of the proportion of subjects in clinical remission by treatment group, the adjusted treatment difference (with Cochran-Mantel-Haenszel weight) between each ustekinumab treatment group and the placebo group, as well as the associated 95% confidence interval, will be provided.

The multiple testing procedure to control the Type 1 error will be different for the United States and the countries outside the United States, as described below.

Type I error control in countries outside the United States (global testing procedure): A fixed-sequence testing procedure will be used to control the overall Type I error rate at the 0.05 level for the primary endpoint. Specifically, the high maintenance dose group (i.e., ustekinumab 90 mg SC q8w) will be considered significant if its p-value is < 0.05. The low maintenance dose group (ustekinumab 90 mg SC q12w) will be significant if the p-value for both high and low maintenance dose groups are < 0.05. The study will be considered positive if the test involving the high maintenance dose group (ustekinumab 90 mg SC q8w) shows a statistically significant difference versus placebo for the primary endpoint of clinical remission at Week 44, regardless of the result of the test for the low maintenance dose (ustekinumab 90 mg SC q12w).

Type I error control in the United States (US-specific testing procedure): A fixed-sequence testing procedure will be employed for the United States to strongly control the overall Type 1 error rate at the 0.05 level across the primary and all 4 major secondary endpoints and across the 2 ustekinumab doses, starting with the high maintenance dose group (ustekinumab 90 mg SC q8w) of the primary endpoint. The exact testing procedure is detailed in Section 5.3.2. The study will be considered positive if the test involving the high maintenance dose group (ustekinumab 90 mg SC q8w) is positive for clinical remission at Week 44.

5.2.3. Sample Size Justification

The major efficacy analyses in the maintenance study will be based on the primary population, i.e., subjects who were in clinical response to IV ustekinumab induction. Unless otherwise stated, the sample size/power calculations in this section refer to this population.

For both global and US-specific testing procedures, a fixed-sequence testing procedure, starting with the high dose group (q8w), will be used to control the overall Type I error rate at the 0.05 level (2-sided). Therefore, sample size/power calculations were based on the chi-square test to detect a significant difference between subjects receiving SC ustekinumab 90 mg q8w and those receiving placebo.

The definition of the primary endpoint of clinical remission will be different for countries outside the United States and for the United States, as described in Section 5.2.1.2.

The treatment effect for the global definition of the primary endpoint of clinical remission at Week 44 in the maintenance study, was based on maintenance data from similarly designed studies of the anti-TNFα golimumab and of vedolizumab.[7,3] In the golimumab UC maintenance study, the proportions of subjects in clinical remission (based on the global definition) at Week 54 (among subjects in clinical response to golimumab induction) were 34% and 22% in the 100 mg group and the placebo group, respectively. In the vedolizumab UC study, the proportions of subjects in clinical remission (based on the global definition) at Week 52 (among subjects in clinical response at Week 6) were 42% and 16% in the vedolizumab q8w group and the placebo group, respectively. For the CNTO1275UCO3001 maintenance study, it was assumed that clinical remission rates (based on the global definition) at Week 44 were 40% and 20%, respectively, for the ustekinumab 90 mg SC q8w and placebo groups. The clinical remission rates at Week 44 based on the US definition are also assumed to be 40% and 20%,

respectively, for the ustekinumab 90 mg SC q8w and placebo groups based on data from the golimumab UC maintenance study (C0524T18).

Assuming a 20% clinical remission rate (for both global and US definitions) in the placebo group and 40% in the SC ustekinumab 90 mg q8w group, 109 subjects in each treatment group (327 subjects in total) will provide statistical power of 90% at a significance level of 0.05 (2-sided). Table 1 shows the power for detecting a treatment difference between the SC ustekinumab 90 mg q8w and the placebo group based on different proportions of subjects in clinical remission (for both the global and US definitions) at Week 44 with a fixed sample size of 327.

Table 1: Power for detecting a treatment effect based on different proportions of subjects in clinical remission (both global and US definitions) at Week 44 with a fixed sample size of 327 subjects (109 in each treatment group)

Proportion of Subjects in Clinical Remission at Week 44 (%)		
Placebo	**Ustekinumab**	**Power[a] (%)**
20	45	98
	42	94
	40	90
	37	80
	35	70

a: Based on testing the ustekinumab 90 mg SC q8w group versus placebo at α=0.05 (2-sided).

Table 2 shows the power for detecting a treatment difference between the ustekinumab 90 mg SC q8w and the placebo group for each of the major secondary endpoints with 327 subjects in the primary population. The assumptions about the proportion of subjects achieving each major secondary endpoint have been based on data from the golimumab[7] and vedolizumab[3] maintenance studies in subjects with moderately to severely active UC.

Table 2: Power for detecting a treatment effect for each of the major secondary endpoints with 327 subjects in the primary population (109 in each treatment group)

	Proportion of subjects achieving the endpoint		
Major secondary endpoints	**Placebo**	**Ustekinumab**	**Power[a] (%)**
Maintenance of clinical response through Week 44	25	50	97
Endoscopic healing at Week 44	25	45	88
Clinical remission and not receiving concomitant corticosteroids at Week 44	15	30	76
Maintenance of clinical remission through Week 44 *among the subjects who had achieved clinical remission at maintenance baseline (both global and US definitions)* [b]	25	50	65

a: Based on testing the SC ustekinumab 90 mg q8w group versus placebo at α=0.05 (2-sided).
b: It is estimated that about 37% of subjects in the primary population (40 subjects per treatment group) will be in clinical remission at Week 0 of maintenance.

The number of subjects in the primary analysis population of the maintenance study will depend on the number of subjects from the following 2 groups of the induction study: 1) subjects in clinical response to IV ustekinumab induction at Week 8 of the induction study (Group A), and 2) subjects who were not in clinical response to IV placebo induction at Week 8 of the induction study but were in clinical response at induction Week 16 after receiving an induction dose of IV ustekinumab at Week 8 (Group B). If the average clinical response rate to IV ustekinumab induction is 45%, 317 subjects in each induction treatment group (for a total of 951subjects) will result in about 328 subjects in the primary population of the maintenance study, assuming 15% attrition from the induction study to the maintenance study. However, the average clinical response rate to IV ustekinumab induction could range from 40% to 50% and attrition from the induction study to the maintenance study could range from 10% to 15%. With 317 subjects in each induction treatment group, the number of subjects in the primary population of the maintenance study could range from 290 to 385 (Table 3).

Table 3: Projected number of subjects in the primary population of the maintenance study with a sample size of 951 subjects (317 per treatment group) in the induction study

Attrition From Induction Study to Maintenance Study	Clinical Response Rate to IV Ustekinumab Induction	Subjects in Group A Entering Maintenance	Subjects in Group B Entering Maintenance[a]	Number of Subjects in Primary Population of Maintenance
15%	40%	216	75	291
	45%	243	85	328
	50%	269	94	363
10%	40%	228	80	308
	45%	257	90	347
	50%	285	100	385

Group A=Subjects in clinical response to IV ustekinumab induction at induction Week 8; Group B=Subjects not in clinical response to IV placebo induction at induction Week 8 but in clinical response at induction Week 16 after receiving an induction dose of IV ustekinumab at Week 8.
a: The proportion of subjects not in clinical response to intravenous placebo induction at Week 8 of the induction study is assumed to be 70%.

5.2.4. Subgroup Analyses

The consistency of treatment effect for the primary endpoint will be evaluated for the subgroups defined in Section 2.4. For each of these subgroups, the odds ratio of each ustekinumab dose group vs placebo and the associated 95% confidence interval will be provided. The odds ratios and confidence intervals will be provided based on the logistic regression model that includes factors for treatment group, clinical remission status at baseline and induction treatment (for the subgroup analyses based on clinical remission status at baseline or induction treatment factor, the corresponding factor will not be included in the model).

Subgroup analyses will also be performed for Japanese subjects. These analyses will be included in a separate report to support the submission in Japan.

5.2.5. Sensitivity Analysis

To examine the robustness of the primary analysis, sensitivity analyses will be conducted for the primary endpoint of clinical remission (for both global and US-specific definitions) at Week 44 based on observed case, worst case, and multiple imputation method as described below.

1. Sensitivity Analysis 1 (Observed case): exclude subjects with missing Mayo data at Week 44 and have not had an event of treatment failure prior to Week 44.

    - For the global definition, subjects with missing data are those who are missing all four Mayo subscores.
    - For the US-specific definition, subjects with missing data are those who are missing all three Mayo components pertaining to the primary endpoint (i.e., the average stool number, rectal bleeding subscore, and Mayo endoscopy subscore).

2. Sensitivity Analysis 2 (nonresponder if any missing subscore): subjects with missing data for any of the Mayo components pertaining to the primary endpoint (i.e., the four Mayo subscores for the global definition and the average stool number, Mayo rectal bleeding subscore, and Mayo endoscopy subscore for the US-specific definition) at Week 44 are considered not to be in clinical remission.

3. Sensitivity Analysis 3 (Multiple imputation):

    1) Any missing Mayo components pertaining to the primary endpoint at Week 44 will be imputed 5 times to generate 5 complete data sets using the Markov chain Monte Carlo method, assuming missing at random (MAR) and a multivariate normal distribution. The following variables will be included in the imputation model: Mayo subscores at Week 0 of an induction study and at Week 0 and Week 44 of the maintenance study, induction dose factor and the maintenance treatment group.

    2) Before the multiple imputation method is applied, the Mayo data after treatment failure will be set to missing. For each multiple-imputed data set, each subject's clinical remission status at Week 44 will be calculated.

    3) Each of the 5 resulting data sets will be analyzed using a logistic regression model with treatment group, baseline Mayo score (for global definition) or baseline modified Mayo score (for US-specific definition), and induction treatment as covariates. The modified Mayo score was used as the baseline disease activity measure for the US-specific definition of clinical remission as it contains the same Mayo components as those for the US-specific definition of clinical remission

    4) The results from the 5 data sets will be combined to produce inferential results.

4. Sensitivity Analysis 4 (Worst case): subjects receiving placebo who have a missing Mayo score at Week 44 and have not had an event of treatment failure prior to Week 44 (as defined in Sensitivity Analysis 1) will be considered to be in clinical remission, and subjects receiving ustekinumab who have a missing Mayo score at Week 44 will be considered not to be in clinical remission.

In addition, the following sensitivity analyses will be performed:

5. Sensitivity Analysis 5 ("modified ITT" analysis): excludes subjects who are randomized but not treated.

6. Sensitivity analysis 6 (Per-Protocol analysis): Excludes subjects with at least 1 of the following deviations that might affect efficacy: 1) subjects who are found to have Crohn's disease instead of UC, 2) subjects who are randomized, but do not receive study agent, 3) subjects who are a treatment failure in the induction study, 4) subjects who are excluded in the Per-Protocol analysis in the induction study, 5) subjects who are randomized to the placebo group at Week 0 but receive a ustekinumab injection at any administration visit through Week 44 or subjects who are randomized to one of the ustekinumab groups at Week 0 but receive only placebo through Week 44 , 6) subjects who are not in clinical response on entry into the maintenance study based on the final reported endoscopy score and the eCRF Mayo score data, or have other deviations that will be identified and documented prior to unblinding.

7. Sensitivity analysis 7 (logistic regression): The primary endpoint will be analyzed using a logistic regression model.

    1) For the global definition, the primary endpoint will be analyzed based on a logistic regression model with maintenance baseline Mayo score, induction treatment, and maintenance treatment group as covariates.

    2) For the US-specific definition, the primary endpoint will be analyzed based on a logistic regression model with maintenance baseline modified Mayo score, induction treatment, and maintenance treatment group as covariates.

8. Sensitivity Analysis 8 (clinical remission based on local endoscopy subscores): the primary analysis will be repeated with clinical remission derived based on local endoscopy subscores, instead of the final reported endoscopic subscores.

The following sensitivity analyses will be performed for the US-specific definition of clinical remission:

9. Sensitivity Analysis 9 (nonresponder if friability present on endoscopy): subjects with friability at Week 44 are considered not to have achieved clinical remission. Note that subjects with missing friability or a friability that cannot be determined will be considered not to have friability.

10. Sensitivity Analysis 10 (Tipping Point Analysis): the adjusted treatment difference (with Cochran-Mantel-Haenszel weight) in proportion of subjects in clinical remission at Week 44 between each ustekinumab group and placebo group and the corresponding p-values will be provided for possible scenarios of remission status for subjects with missing data (0%-100% of subjects in clinical remission, in increments of at least 10%, for the placebo and the ustekinumab group).

The same treatment failure rules as were used in the primary analysis will be applied. For Sensitivity Analyses 5 through 9, the same missing data rules as are used in the primary analysis will also be applied. Note that treatment failure rules override the missing data rules. Subjects who have a treatment failure prior to Week 44 are considered not to be in clinical remission at Week 44, regardless of the observed or missing data.

5.3. Major Secondary Endpoints

The following are the **major secondary endpoints**, presented in the order in which they will be tested:

1. Maintenance of clinical response through Week 44;
2. Endoscopic healing at Week 44;
3. Clinical remission and not receiving concomitant corticosteroids (corticosteroid-free clinical remission) at Week 44;
4. Maintenance of clinical remission through Week 44 among the subjects who had achieved clinical remission at maintenance baseline

5.3.1. Definitions

- **Clinical response** is defined as a decrease from induction baseline in the Mayo score by ≥30% and ≥3 points, with either a decrease from induction baseline in the rectal bleeding subscore ≥1 or a rectal bleeding subscore = 0 or 1.
- **Endoscopic healing** (i.e., improvement in the endoscopic appearance of the mucosa) is defined as endoscopy subscore of the Mayo score = 0 or 1.
- **Clinical remission** is defined in Section 5.2.1.2. Two separate definitions for clinical remission will be applied to accommodate the global and US-preferred definitions of clinical remission. Note that, for the United States, for the fourth major secondary endpoint (maintenance of clinical remission through Week 44 among the subjects who had achieved clinical remission at maintenance baseline), the US-specific definition of clinical remission should also be used to determine whether or not a subject had achieved clinical remission at maintenance baseline.

Treatment failure and missing data rules will be applied to each of these major secondary endpoints. Subjects who have a treatment failure prior to the M-44 visit will be considered not to have achieved the respective endpoints. At Week 44, subjects who have a missing Mayo endoscopy subscore will be considered not to have endoscopic healing; subjects who have all 4 Mayo subscores missing will be considered not to be in clinical response or clinical remission (for the global definition of remission). For the US-specific definition of clinical remission,

subjects who are missing the average stool number, rectal bleeding subscore, and Mayo endoscopy subscore at Week 44 will be considered not to be in clinical remission. For subjects without corticosteroid information at Week 44, the last available corticosteroid dose will be carried forward to Week 44.

5.3.2. Analysis Methods

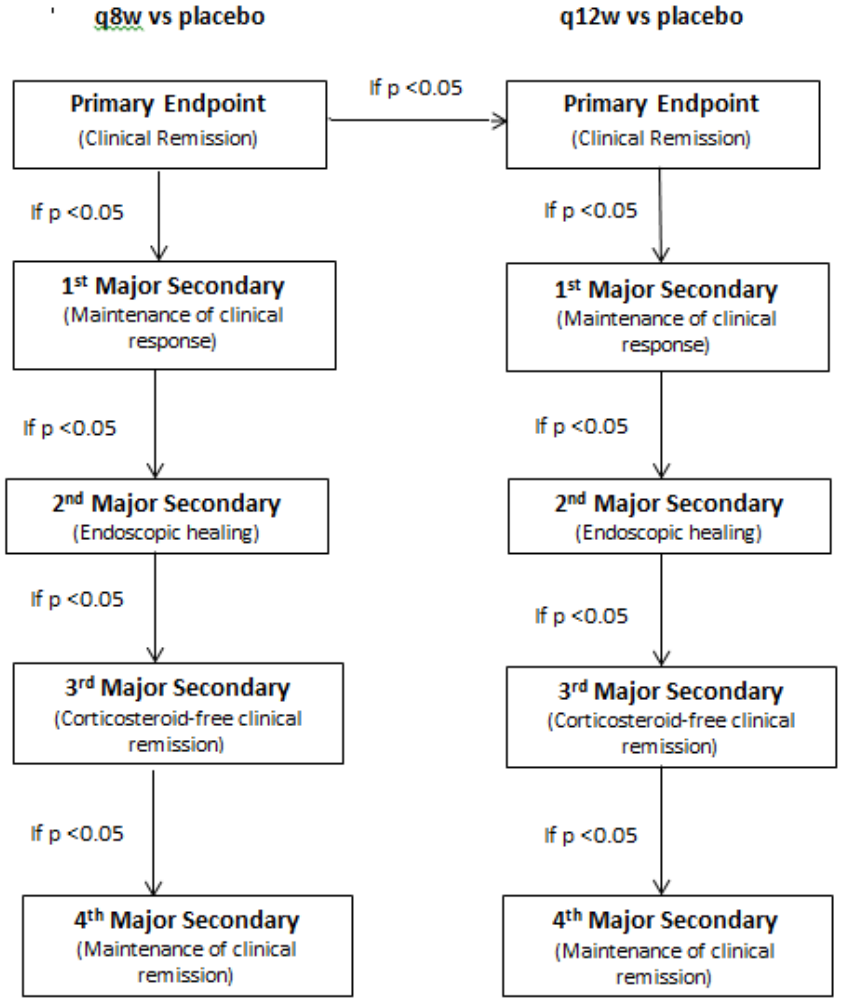
The major secondary endpoint analyses will be based on the **Primary Efficacy Analysis Set**.

Except for the fourth major secondary endpoint of maintenance of clinical remission, analyses of major secondary endpoints will be conducted using a CMH chi-square test stratified by clinical remission status at maintenance baseline (yes/no as determined by the IWRS) and induction treatment. For the fourth major secondary endpoint, a CMH chi-square test stratified by induction treatment will be used. Summaries of the proportion of subjects by treatment group, the adjusted treatment difference (with Cochran-Mantel-Haenszel weight) between each ustekinumab treatment group and the placebo group, as well as the associated 95% confidence interval, will be provided for all the major secondary endpoints.

The United States and the global regions will employ a different multiple testing strategy as described below to control the overall Type I error.

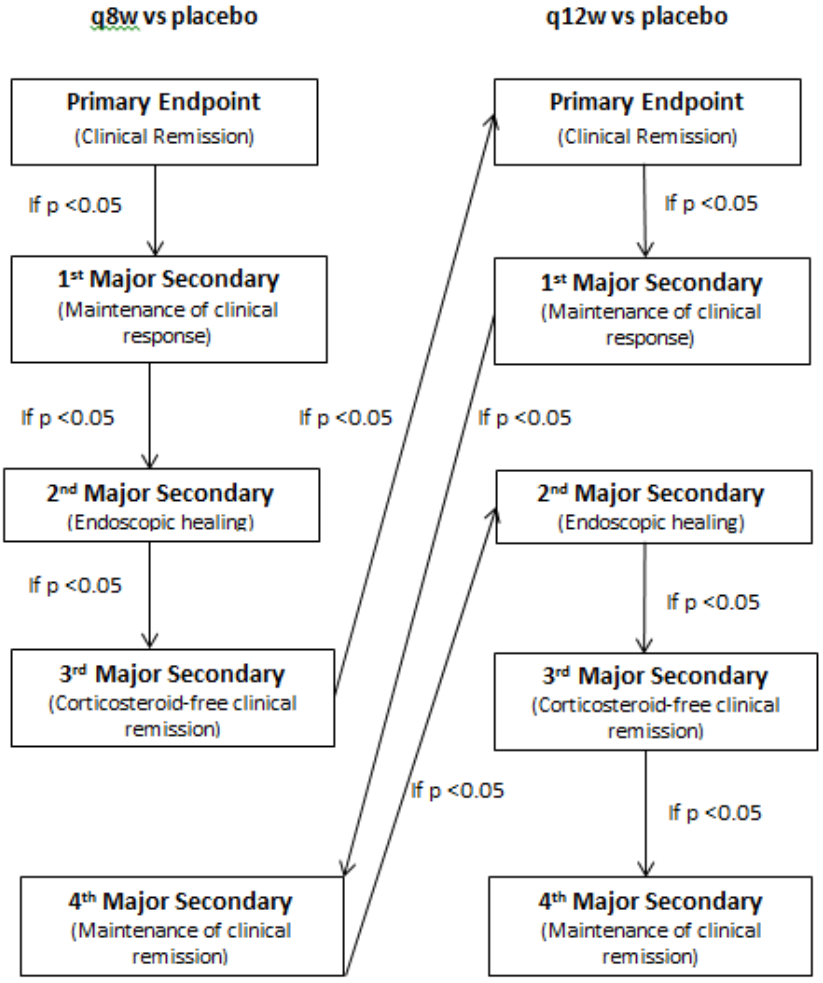
Type I error control in countries outside the United States (global testing procedure): A hierarchical testing procedure as shown in Figure 3 will be employed to control the overall Type 1 error rate over the 4 major secondary efficacy analyses at the (2-sided) 0.05 significance level **within** a dose group. A major secondary endpoint for a dose group will be considered significant only if both the previous endpoints in the hierarchy and current endpoint test positive at the 2-sided 0.05 level of significance. If an endpoint is not significant, all subsequent tests in the hierarchy will be considered not to be significant. Nominal p-values will be reported for all analyses.

Figure 3: Global testing procedure for primary and major secondary endpoints



Type I error control in the United States (US-specific testing procedure): A hierarchical testing procedure as shown Figure 4 (and listed in Attachment 4) will be employed for the United States to strongly control the overall Type 1 error rate at the 0.05 level across the primary and all 4 major secondary endpoints and across the 2 ustekinumab doses. An endpoint will be considered significant only if both the previous endpoints in the hierarchy and current endpoint test positive at the 2-sided 0.05 level of significance. If an endpoint is not significant, all subsequent tests in the hierarchy will be considered not to be significant. Nominal p-values will be reported for all analyses.

Figure 4: US-specific testing procedure for the primary and major secondary endpoints



5.3.3. Subgroup Analyses

The primary and major secondary endpoints will be analyzed using a CMH chi-square test described above based on the following subgroups, when the number of subjects within each level of the subgroup permits. Summaries of the proportion of subjects by treatment group, the adjusted treatment difference (with Cochran-Mantel-Haenszel weight) between each ustekinumab treatment group and the placebo group, as well as the associated 95% confidence interval, will be provided

CRP and Fecal Biomarkers (calprotecin and lactoferrin), both at Week 0 of the induction study and at Week 0 of the maintenance study

- CRP (≤3 mg/L, >3 mg/L)
- CRP (≤ median, > median)
- CRP (≤10 mg/L, >10 mg/L)

- Fecal calprotectin (≤250 mg/kg, >250 mg/kg)
- Fecal lactoferrin (≤7.24 µg/g, >7.24 µg/g)

Induction baseline Weight (≤ 1st quartile, >1st quartile and ≤ 2nd quartile, >2nd quartile and ≤ 3rd quartile, >3rd quartile)

5.4. Other Efficacy Endpoints

Regardless of the significance of the endpoints in the hierarchical testing procedures described above, the testing of the other endpoints in the following sections will proceed. These endpoints will not be adjusted for multiplicity, and statements of significance for these endpoints will be based on nominal p-values.

5.4.1. Endpoints

5.4.1.1. Clinical Endpoints

The following endpoints will be summarized and compared between each of the ustekinumab treatment groups and the placebo treatment group:

- The change from maintenance baseline in the Mayo score at Week 44.
- The change from maintenance baseline in the partial Mayo score over time through Week 44.
- The change from maintenance baseline in the modified Mayo score (i.e., Mayo score without the PGA) at Week 44.
- Partial Mayo remission over time through Week 44.
- Durable partial Mayo remission through Week 44.
- Clinical remission (global definition) at Week 44 with a rectal bleeding subscore of 0.
- Remission at Week 44 based on a stool frequency subscore of 0 or 1, a rectal bleeding subscore of 0, and an endoscopy subscore of 0 or 1.
- Remission at Week 44 based on a stool frequency subscore of 0, a rectal bleeding subscore of 0, and an endoscopy subscore of 0 or 1.
- Remission at Week 44 based on a stool frequency subscore of 0 or 1, a rectal bleeding subscore of 0, and an endoscopy subscore of 0 or 1 (Nonresponder if friability present on endoscopy).
- Remission at Week 44 based on a stool frequency subscore of 0, a rectal bleeding subscore of 0, and an endoscopy subscore of 0 or 1 (Nonresponder if friability present on endoscopy).
- Symptomatic remission through Week 44.
- Maintenance of symptomatic remission through Week 44 among the subjects who had achieved symptomatic remission at maintenance baseline.
- Endoscopic healing at Week 44 among the subjects who had achieved endoscopic healing at maintenance baseline.

- Histologic healing at Week 44.
- Endoscopic healing at Week 44 (Nonresponder if friability present on endoscopy).
- Mucosal healing at Week 44.
- Mucosal healing at Week 44 (Nonresponder if friability present on endoscopy).
- Normal or inactive mucosal disease at Week 44.
- Subjects in clinical response and not receiving corticosteroids at Week 44.
- The change from maintenance baseline in the average daily prednisone-equivalent corticosteroid dose (excluding budesonide and beclomethasone dipropionate) over time through Week 44 among the subjects who were receiving concomitant corticosteroids at maintenance baseline.
- Subjects who are not receiving concomitant corticosteroids at Week 44 among the subjects who were receiving concomitant corticosteroids at maintenance baseline.
- Subjects who were not receiving corticosteroids for at least 90 days prior to Week 44 among subjects who were receiving corticosteroids at maintenance baseline.
- Subjects who were not receiving corticosteroids for at least 30 days prior to Week 44 among subjects who were receiving corticosteroids at maintenance baseline.
- The time to loss of clinical response through Week 44.
- Clinical response through Week 44 based on the Mayo score that includes the endoscopy subscore assigned by the local endoscopist.
- Endoscopic healing at Week 44 based on the endoscopy subscore assigned by the local endoscopist.
- Normal or inactive mucosal disease at Week 44 based on the endoscopy subscore assigned by the local endoscopist.
- Mucosal healing at Week 44 based on the endoscopy subscore assigned by the local endoscopist.
- Endoscopic healing at Week 44 by biologic failure status.
- Maintenance of clinical response through Week 44 by biologic failure status.
- Mucosal healing at Week 44 by biologic failure status

In addition, primary and major secondary endpoints as well as mucosal healing endpoint will be evaluated based on the following biologic failure profiles: subjects who were biologic-naïve, subjects who were biologic failures to at least 1 anti-TNF (regardless of vedolizumab), and subjects who were biologic failures to both anti-TNF and vedolizumab, when the number of subjects within each subpopulation permits. Furthermore, Sensitivity analyses will be performed for mucosal healing at Week 44 and histologic healing at Week 44 by considering subjects with an unevaluable biopsy at Week 44 as nonresponders.

The following endpoints will be summarized and compared for both global and US-specific definitions of clinical remission:

- Subjects in clinical remission at Week 44 and not receiving corticosteroids for at least 90 days prior to Week 44.
- Subjects in clinical remission at Week 44 and not receiving corticosteroids for at least 30 days prior to Week 44.
- Subjects in clinical remission at Week 44 and not receiving corticosteroids at Week 44 among the subjects receiving concomitant corticosteroids at maintenance baseline.
- Subjects in clinical remission at Week 44 and not receiving corticosteroids for at least 90 days prior to Week 44 among the subjects receiving concomitant corticosteroids at maintenance baseline.
- Subjects in clinical remission at Week 44 and not receiving corticosteroids for at least 30 days prior to Week 44 among the subjects receiving concomitant corticosteroids at maintenance baseline.
- The time to loss of clinical remission through Week 44 among subjects who had achieved clinical remission at maintenance baseline.
- Clinical remission at Week 44 by biologic failure status.
- Corticosteroid-free clinical remission at Week 44 by biologic failure status
- Maintenance of clinical remission through Week 44 among the subjects who had achieved clinical remission at maintenance baseline by biologic failure status.

In addition, the following endpoints will be summarized:

- The change from induction baseline in the Mayo score over time through Week 44.
- The change from induction baseline in the partial Mayo score over time through Week 44.
- Mayo subscores through Week 44.
- Clinical remission (both global and US-specific definitions) at Week 44 by investigative site.

5.4.1.2. CRP and Fecal Biomarkers

The following endpoints will be summarized and compared between each of the ustekinumab treatment groups and the placebo treatment group.

- The change from maintenance baseline in CRP concentration over time through Week 44.
- The change from maintenance baseline in fecal lactoferrin concentration over time through Week 44.
- The change from maintenance baseline in fecal calprotectin concentration over time through Week 44.
- Normalization of CRP over time through Week 44 among subjects with abnormal CRP at induction baseline.

- Normalization of fecal lactoferrin over time through Week 44 among subjects with abnormal fecal lactoferrin at induction baseline.
- Normalization of fecal calprotectin over time through Week 44 among subjects with abnormal fecal calprotectin concentration at induction baseline.

In addition, CRP concentration, fecal lactoferrin concentration and fecal calprotectin concentration through Week 44 will be provided by CRP (≤ 3 mg/L, > 3 mg/L), fecal lactoferrin (≤7.24 ug/g, > 7.24 µg/g) and fecal calprotectin (≤ 250 mg/kg, > 250 mg/kg) at Week 0 of the induction study, respectively.

5.4.1.3. Patient Reported Outcomes

The following endpoints will be summarized and compared between each of the ustekinumab treatment groups and the placebo treatment group.

- The change from maintenance baseline in the total IBDQ score and each of the 4 IBDQ dimensions through Week 44.
- A >20-point improvement from induction baseline in the IBDQ score at Week 44.
- A ≥16-point improvement from induction baseline in the IBDQ score at Week 44.
- IBDQ remission at Week 44
- Maintenance of 20-point improvement in IBDQ through Week 44 among subjects with a >20-point improvement in IBDQ at the maintenance baseline from the induction baseline.
- Maintenance of 16-point improvement in IBDQ through Week 44 among subjects with a ≥16-point improvement in IBDQ at the maintenance baseline from the induction baseline.
- Maintenance of IBDQ remission among subjects with IBDQ remission at the maintenance baseline.
- The change from maintenance baseline in the SF-36 physical and mental component summary scores and the individual scale scores through Week 44.
- Maintenance of 5-point improvement in the SF-36 physical component score (PCS) through Week 44 among subjects with a ≥5-point improvement in the SF-36 PCS at the maintenance baseline from the induction baseline
- Maintenance of 5-point improvement in the SF-36 mental component score (MCS) through Week 44 among subjects with a ≥5-point improvement in the SF-36 MCS at the maintenance baseline from the induction baseline
- A ≥5-point improvement from induction baseline in the SF-36 PCS through Week 44.
- A ≥5-point improvement from induction baseline in the SF-36 MCS through Week 44.
- The changes from maintenance baseline in the EQ-5D dimensions, EQ-5D index, and health state VAS scores through Week 44.

5.4.2. Definitions

The Mayo/Partial Mayo scores, and global and US-specific definitions of clinical remission are defined in Section 5.2.1, Clinical response and endoscopic healing are defined in Section 5.3.1. The definitions for other efficacy endpoints not defined previously are as follows:

Clinical Endpoints:

- **Partial Mayo remission:** A partial Mayo score ≤ 2.
- **Durable partial Mayo remission:** Achieving partial Mayo remission at ≥80% of all visits (at least 9 out of 11 visits) prior to Week 44 and in partial Mayo remission at last visit (Week 44).
- **Symptomatic remission:** A stool frequency subscore of 0 or 1 and a rectal bleeding subscore of 0.
- **Maintenance of symptomatic remission:** Achieving symptomatic remission at ≥80% of all visits from Week 4 to Week 40 (at least 8 out of 10 visits) and in symptomatic remission at last visit (Week 44) among subjects who had achieved symptomatic remission at maintenance baseline.
- **Normal or inactive mucosal disease:** Endoscopy subscore = 0.
- **Histological healing:** It is based on the Geboes score (Attachment 2; Geboes et al, 2000)[4] and is defined as 0-<5% neutrophils in epithelium and no crypt destruction, erosions, ulcerations or granulations.
- **Mucosal healing:** Endoscopic healing (i.e., endoscopy subscore of the Mayo score = 0 or 1) and histological healing.
- **The time to loss of clinical remission/response through Week 44:** It is defined as the number of days elapsed from the date of the Week 0 study agent administration in this maintenance study to the date of first loss of clinical remission/response at or prior to Week 44. Subjects who have missing data or who do not have an event at or prior to Week 44 will be censored at Week 44, date of study completion, or early termination, whichever happens earlier. Subjects who have a treatment failure (see Section 5.1.1.1) prior to Week 44 will be considered to have an event (loss of clinical remission/response) at the time of treatment failure.
- **Average daily prednisone-equivalent corticosteroid dose:** For corticosteroid medication (excluding budesonide and beclomethasone dipropionate) use at a specific time point, the average daily prednisone-equivalent corticosteroid dose is calculated using the corticosteroid dosage levels from the 6 days prior to the day before the visit.

CRP and Fecal Biomarkers:

- **Normalization of CRP concentration**: CRP concentration ≤3 mg/L.
- **Normalization of fecal lactoferrin concentration**: Fecal lactoferrin concentration ≤7.24 µg/g.
- **Normalization of fecal calprotectin concentration**: Fecal calprotectin concentration ≤250 mg/kg.

Patient Reported Outcomes:

- **Inflammatory Bowel Disease Questionnaire (IBDQ):** The IBDQ is a 32-item questionnaire for subjects with IBD that will be used to evaluate the disease-specific health-related quality of life across 4 dimensional scores: bowel (loose stools, abdominal pain), systemic (fatigue, altered sleep pattern), social (work attendance, need to cancel social events), and emotional (anger, depression, irritability). The range of the IBDQ score is 32 to 224. Higher scores indicate better quality of life. Each of the individual IBDQ dimensions will be calculated when no more than 1 item is missing in the dimension. The missing item will be estimated using the average value across the nonmissing items. If any one of the dimensions within the IBDQ cannot be calculated, then the total IBDQ score cannot be calculated.
- **SF-36:** The SF-36, a generic measure of health status, evaluates eight individual subscales (physical functioning[PF], role physical[RP], bodily pain[BP], general health[GH], vitality[VT], social functioning[SF], roleemotional[RE], and mental health[MH]) and the physical component summary (PCS, calculated from the subscales PF, RP, BP and GH) and the mental component summary (MCS, calculated from the subscales VT, SF, RE and MH) scores.

  Each of the SF-36 subscales and two summary scores are derived and are scaled to scores with a mean of 50 and standard deviations of 10, based upon general US population norms of year 2009. Maximum Data Recovery option will be used to estimate scores with missing items. Based on the user's guide from the developer, the software applies a value to a scale item rendered missing if at least one of the items in that scale has valid data. A scale receives a "missing" score (".") only if all the items in that scale are missing. PCS and MCS are calculated when at least seven of the eight profile scales have valid data, either actual or estimated. However, to calculate PCS, the PF scale must be one of the seven scales having valid data. Also, to calculate MCS, the MH scale must be one of the seven scales having valid data.

- **IBDQ remission: IBDQ**≥170 (Irvine et al, 1994; Higgins et al, 2005)[5,6]
- **Maintenance of 20-point improvement in IBDQ through Week 44**: Among subjects with a >20-point improvement in IBDQ at the maintenance baseline from the induction baseline, a >20-point improvement in IBDQ at both Week 20 and Week 44 from Week 0 of the induction study.
- **Maintenance of 16-point improvement in IBDQ through Week 44**: Among subjects with a ≥16-point improvement in IBDQ at the maintenance baseline from the induction baseline, a ≥16-point improvement in IBDQ at both Week 20 and Week 44 from Week 0 of the induction study.
- **Maintenance of IBDQ remission:** in IBDQ remission at both Week 20 and Week 44 among subjects with IBDQ remission at the maintenance baseline.
- **Maintenance of 5-point improvement in the SF-36 PCS through Week 44**: A ≥5-point improvement in the SF-36 PCS at both Week 20 and Week 44 from Week 0 of an induction study.

- **Maintenance of 5-point improvement in the SF-36 MCS through Week 44**: A ≥5-point improvement in the SF-36 MCS at both Week 20 and Week 44 from Week 0 of an induction study.
- **EuroQoL-5D Health Questionnaire:** The EQ-5D is designed for self-completion by subjects and consists of 2 pages - the EQ-5D descriptive system and the EQ visual analog scale (EQ VAS). The EQ-5D descriptive system comprises the following 5 dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each dimension has 5 levels: no problems, slight problems, moderate problems, severe problems, and extreme problems. The EQ VAS records the respondent's self-rated health on a vertical, visual analog scale where the endpoints are labeled 'Best imaginable health state' (score of 100) and 'Worst imaginable health state' (score of 0). The EQ-5D descriptive system can be converted into a single summary EQ-5D Index. EQ-5D index scores in this analysis will be derived with the algorithm and software provided by the developer based on the UK model.[2]

5.4.3. Analysis Methods for Other Efficacy Endpoints

The analyses in this section will be based on the **Primary Efficacy Analysis Set**. A 2-sided significance level of 0.05 will be used for all tests. The treatment failure rules and missing data rules described in Section 5.1.1 will be applied to each of the above endpoints unless otherwise specified. For endpoints involving friability, subjects with missing friability or a friability that cannot be determined will be considered not to have friability.

Dichotomous endpoints will be summarized and compared between each of the ustekinumab treatment groups and the placebo treatment group using a CMH chi-square test, stratified by clinical remission status at maintenance baseline and induction treatment. In cases of rare events, Fisher's exact test will be used for treatment comparisons.

The change from maintenance baseline in the Mayo, partial Mayo score, modified Mayo score, and the average daily prednisone-equivalent corticosteroid dose will be analyzed using analysis of covariance (ANCOVA) with the respective baseline value, clinical remission status at maintenance baseline, induction treatment, and maintenance treatment group as covariates.

The normality assumption of the other continuous endpoints (i.e., the change from maintenance baseline in CRP, Fecal Lactoferrin, Fecal Calprotectin, total IBDQ, IBDQ dimensions, SF-36 PCS and MCS, EQ-5D index, and health state VAS score etc.) is in question based on results from previous studies, therefore for them, the comparison between each ustekinumab treatment group and the placebo group will be made using an ANCOVA on the van der Waerden normal scores with the respective baseline value, clinical remission status at maintenance baseline, induction treatment, and maintenance treatment group as covariates.

For the change from baseline in EQ-5D dimensions scores, comparisons between each of the ustekinumab treatment groups and placebo will be made using a CMH chi-square (Row Mean Scores) test stratified by clinical remission status at maintenance baseline and induction treatment.

Time to event endpoints will be compared between each of the ustekinumab treatment groups and the placebo treatment group using the stratified log-rank test with clinical remission status at maintenance baseline and induction treatment as the stratification factors. The Kaplan-Meier curve by treatment group will be provided.

5.5. Efficacy Endpoints in Other Populations

The following endpoints will be summarized based on the **Non-randomized Analysis Set**. Treatment failure and missing data rules will be applied.

- Clinical remission (for both global and US-specific definitions) at Week 44
- Maintenance of clinical response through Week 44
- Endoscopic healing at Week 44
- Partial Mayo remission over time through Week 44
- Corticosteroid-free clinical remission at Week 44
- Mucosal healing at Week 44

5.6. Exploratory Endpoints

The following exploratory endpoints will be summarized and compared between each of the ustekinumab treatment groups and the placebo treatment group. The analyses will be based on the **Primary Efficacy Analysis Set**.

5.6.1. Clinical Endpoints

- Modified Mayo score response (i.e., response based on the Mayo score without the PGA subscore) at Week 44 by two definitions as defined in Section 5.2.1.1.
- The change from baseline in Ulcerative Colitis Endoscopic Index of Severity (UCEIS) score at Week 44.
- UCEIS score ≤4 at Week 44.
- The UCEIS score at Week 0 and at Week 44 of the maintenance study by the level of the Mayo endoscopy subscore at the corresponding visit.

5.6.2. Definitions

Modified Mayo score response definition 1: decrease in the modified Mayo score of ≥2 points and ≥ 35% and either a decrease in the rectal bleeding subscore of ≥ 1 or a rectal bleeding subscore of 0 or 1.

Modified Mayo score response definition 2: decrease in the modified Mayo score of ≥2 points and ≥ 30% and either a decrease in rectal bleeding of ≥ 1 or a rectal bleeding score of 0 or 1.

Ulcerative Colitis Endoscopic Index of Severity (UCEIS):[9] The UCEIS (Attachment 3) is an index that provides an overall assessment of endoscopic severity of UC and is the summation of the scores from the following three components:

- Mucosal vascular pattern (1=Normal; 2=Patchy obliteration; 3=Obliterated)
- Bleeding (1=None; 2=Mucosal; 3=Luminal mild; 4=Luminal moderate or severe)
- Ulceration (1=None; 2=Erosions; 3=Superficial ulcer; 4=Deep ulcer)

The UCEIS 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.

The treatment failure rules and missing data rules described in Section 5.1.1 will be applied to these endpoints, except for the UCEIS score by the level of Mayo endoscopy subscore. For the UCEIS score by the level of Mayo endoscopy subscore, no treatment failure and missing data rules will be applied.

5.6.3. Analysis Methods

The endpoints in Section 5.6.1 will not be controlled for multiplicity. Statements of significance will be based on nominal p-values.

For dichotomous endpoints, the comparison between each ustekinumab treatment group and the placebo group will be conducted using a CMH chi-square test, stratified by clinical remission status at maintenance baseline and induction treatment.

The change from baseline in the UCEIS score at Week 44 will be analyzed using an ANCOVA with the respective baseline value, clinical remission status at maintenance baseline, induction treatment, and maintenance treatment group as covariates.

The UCEIS score at Week 0 and at Week 44 of the maintenance study will be summarized by the level of Mayo endoscopy subscore at the corresponding visit to assess the association between the UCEIS score and the Mayo endoscopy subscore. The endoscopy subscores assigned by the central reader will be used.

6. SAFETY

Safety will be assessed by summarizing the frequency and type of AEs and by summarizing laboratory parameters (hematology and chemistry) and vital signs parameters.

Unless otherwise mentioned, safety summaries will be provided for subjects in the **Safety Analysis Set**, overall and by randomization status. In general, subjects will be analyzed according to their assigned treatment.

6.1. Adverse Events

Treatment-emergent AEs through Week 44 will be summarized by treatment group and Medical Dictionary for Regulatory Affairs (MedDRA) system organ class and preferred term. A treatment-emergent AE is any AE that occurs at or after the initial administration of study agent of this maintenance study.

The following treatment-emergent summary tables will be provided:

1. Any AEs.
2. SAEs.
3. Reasonably related AEs.
4. AEs leading to discontinuation of study agent.
5. Infections, including infections requiring with an oral or parenteral antibiotic treatment.
6. Serious infections.
7. Injection-site reactions.

These summary tables will provide counts and percentages of subjects with 1 or more of the specified AEs by treatment group. In addition to the summary tables, listings of subjects with SAEs, and AEs leading to discontinuation of study agent will be provided. Any deaths, possible anaphylactic or serum-sickness like reactions, malignancies, or major cardiovascular events will either be presented in a listing or described in the clinical study report.

The number of adverse events per hundred subject-years of follow-up through Week 44 will also be provided.

AEs, SAEs, and serious infections will also be summarized by treatment group and induction baseline body weight (≤ median body weight or > median body weight).

A reasonably related AE is defined as any event with a relationship to study agent of 'Very Likely', 'Probable', or 'Possible' on the AE eCRF page or if the relationship to study agent is missing.

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

A study agent injection-site reaction is any adverse reaction at an SC study agent injection site. The injection sites will be evaluated for reactions and any injection site reaction will be recorded as an AE (and an injection-site reaction) by the investigator on the eCRF.

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

If any subject receives ustekinumab outside of this study protocol (i.e commercially available Stelara), AE data after the receipt of ustekinumab will be excluded from the AE analyses to preserve the integrity and accuracy of the randomized treatment groups.

6.2. Clinical Laboratory Tests

Routine laboratory data for hematology and clinical chemistry will be collected at study visits from Week 0 through Week 44 according to the Time and Events schedule in the Protocol. The following laboratory assessments will be collected:

1. Hematology: hemoglobin, hematocrit, platelet count, total and differential 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 subjects in the **Safety Analysis Set**:

1. Line plots of the observed values and changes from maintenance baseline over time through Week 44.
2. Summary of maximum National Cancer Institute Common Terminology Criteria for Adverse Events (NCI-CTCAE) toxicity grade for postbaseline laboratory values through Week 44.
3. Shift tables 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 through Week 44.
4. Summary of maximum postbaseline measurement through Week 44 for ALT, AST, alkaline phosphatase and total bilirubin relative to ULN.

In addition, summaries for #'s 2 and 4 (above) will be provided by randomization status. Listings of subjects with any abnormal postbaseline laboratory values of CTCAE grade ≥2 will also be provided.

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

Clinical laboratory test values are to be graded based on CTCAE version 4.03 (Table 4). The laboratory tests not included in Table 4 will not be presented in the corresponding tables or listings.

Table 4: Grading Criteria for Clinical Laboratory Tests [CTCAE Version 4.03]

Hematology Tests		**Criteria**			
Test	**Direction**	**1**	**2**	**3**	**4**
Hemoglobin (g/L)	Decrease	≥100 - <LLN	≥80 - <100.0	≥65 - <80	<65
Leukocytes (WBC) (10^9/L)	Decrease	≥3.0 - <LLN	≥2.0 - <3.0	≥1.0 - <2.0	<1.0
Lymphocytes (10^9/L)	Decrease	≥0.8 - <LLN	≥0.5 - <0.8	≥0.2 - <0.5	<0.2
Neutrophils (10^9/L)	Decrease	≥1.5 - <LLN	≥1.0 - <1.5	≥0.5 - <1.0	<0.5
Platelets (10^9/L)	Decrease	≥75.0 - <LLN	≥50.0 - <75.0	≥25.0 - <50.0	<25.0
Chemistry Tests		**Criteria**			
Test	**Direction**	**1**	**2**	**3**	**4**
ALT	Increase	>ULN - ≤3.0 xULN	>3.0 xULN - ≤5.0 xULN	>5.0 xULN - ≤20.0 xULN	>20.0 xULN
Albumin (g/L)	Decrease	≥30 - <LLN	≥20 - <30	<20	
Alkaline Phosphatase	Increase	>ULN - ≤2.5 xULN	>2.5 xULN - ≤5.0 xULN	>5.0 xULN - ≤20.0 xULN	>20.0 xULN
AST	Increase	>ULN - ≤3.0 xULN	>3.0 xULN - ≤5.0 xULN	>5.0 xULN - ≤20.0 xULN	>20.0 xULN
Bicarbonate (mmol/L)	Decrease	≥16 - <LLN	≥11 – <16	≥8 – <11	<8
Bilirubin	Increase	>ULN - ≤1.5 xULN	>1.5 xULN - ≤3.0 xULN	>3.0 xULN - ≤10.0 xULN	>10.0 xULN
Calcium (mmol/L)	Increase	>ULN - ≤2.9	>2.9 - ≤3.1	>3.1 - ≤3.4	>3.4
Calcium (mmol/L)	Decrease	[Albumin ≥40 g/L or missing and calcium ≥2.0 - <LLN]; or [Albumin < 40 g/L and (calcium – 0.8 x (albumin – 40)) ≥2.0 - <LLN]	[Albumin ≥40 g/L or missing and calcium ≥1.75 - <2.0]; or [Albumin < 40 g/L and (calcium – 0.8 x (albumin – 40)) ≥1.75 - <2.0]	[Albumin ≥40 g/L or missing and calcium ≥1.5 - <1.75]; or [Albumin < 40 g/L and (calcium – 0.8 x (albumin – 40)) ≥1.5 - <1.75]	[Albumin ≥40 g/L or missing and calcium <1.5]; or [Albumin < 40 g/L and (calcium – 0.8 x (albumin – 40)) <1.5]
Creatinine	Increase	>ULN - ≤1.5 xULN	>1.5 xULN - ≤3.0 xULN	>3.0 xULN - ≤6.0 xULN	>6.0 xULN
Phosphate (mmol/L)	Decrease	≥0.8 - <LLN	≥0.6 - <0.8	≥0.3 - <0.6	<0.3
Potassium (mmol/L)	Increase	>ULN - ≤5.5	>5.5 - ≤6.0	>6.0 - ≤7.0	>7.0
Potassium (mmol/L)	Decrease	≥3.0 - <LLN		≥2.5 - <3.0	<2.5
Sodium (mmol/L)	Increase	>ULN - ≤150	>150 - ≤155	>155 - ≤160	>160
Sodium (mmol/L)	Decrease	≥130 - <LLN		≥120 - <130	<120

6.3. Vital Signs

Vital signs will be measured at each study visit. The measurements of the vital signs and the changes from maintenance baseline will be summarized by visit and treatment group.

7. PHARMACOKINETICS/PHARMACODYNAMICS

In general, subjects will be analyzed according to their assigned treatment for PK analyses.

7.1. Pharmacokinetics

Blood samples for determining the serum ustekinumab concentrations will be drawn from all subjects according to the Time and Events Schedule in the Protocol. Unless otherwise mentioned, serum ustekinumab concentration summaries will be provided based on the **PK Analysis Set** and **Randomized PK Analysis Set**. These summaries will be based on observed data (i.e., no data imputation). Serum ustekinumab concentrations through Week 44 will be summarized. In addition, serum ustekinumab concentrations through Week 44 will be summarized by induction baseline body weight quartiles for subjects in the **Randomized PK Analysis Set**. The proportion of subjects with concentrations below the lowest quantifiable concentration at each visit through Week 44 will also be summarized. Subjects with concentrations below the lowest quantifiable concentration will be treated as 0 in the summary statistics and labeled as 0 in any listings that contain serum concentration data.

All serum ustekinumab concentration summaries will exclude, from the time of occurrence, data collected for subjects who 1) discontinue study agent, 2) skip an injection, 3) receive an incomplete injection, 4) receive an incorrect injection, 5) receive an additional injection, and/or 6) receive commercial ustekinumab. In addition, PK samples taken outside the scheduled visit window (±10 days of each scheduled visit) will be excluded from the summaries.

The relationship between ustekinumab serum concentrations (at Week 24 and at Week 44) and change in Mayo score, clinical remission (for both global and US-specific definitions), clinical response, endoscopic healing and CRP concentration at Week 44 will be explored based on the **Randomized PK Analysis Set**. Subject data will be summarized according to their assigned treatment.

A population PK analysis will be performed to characterize the PK of ustekinumab. The influence of important covariates on the population PK parameter estimates may be evaluated. The plan for the population PK analyses will be in a separate document and results will be presented in a separate technical report.

7.2. Immune Response

Blood samples will be collected to determine immune response to ustekinumab at Week 0, and at Weeks 4, 12, 24, 36 and 44 using a drug tolerant assay. The antibody to ustekinumab status (positive at any time, negative) and titer through Week 44 will be summarized by treatment group for subjects who received a dose of ustekinumab either in the induction study or in this study and have appropriate samples for detection of antibodies to ustekinumab (i.e., subjects with at least 1 sample obtained after their first dose of ustekinumab). In addition, the Neutralizing Antibody (NAB) Status will be summarized by treatment group through Week 44 weeks.

An evaluation of antibody to ustekinumab status versus change in Mayo score, clinical remission (for both global and US-specific definitions), clinical response, and endoscopic healing at Week 44 will be performed based on the **Randomized PK Analysis Set** to determine the influence of antibodies to ustekinumab on the efficacy of ustekinumab, if the number of subjects with positive antibody status permits.

The relationship between antibody to ustekinumab status and safety will be assessed based on the **Randomized PK Analysis Set** with a summary of injection-site reactions through Week 44 by antibody to ustekinumab status throughWeek 44, if the number of subjects with positive antibody status permits.

Median serum concentrations over time through Week 44 by antibody to ustekinumab status through Week 44 will also be summarized by treatment group based on the **Randomized PK Analysis Set**, if the number of subjects with positive antibody status permits.

A listing of subjects who are positive for antibodies to ustekinumab will be provided. This listing will provide information regarding dose administered, injection-site reactions/AEs within 1 hour of infusion, concomitant medications, partial Mayo score/Mayo Score, ustekinumab serum concentration, and antibody status.

7.3. Pharmacodynamics

Blood samples for serum-based biomarker analyses, and fecal and mucosal biopsy samples will be collected at visits indicated in Time and Events schedule of the protocol. Changes in the concentration of individual serum markers from baseline to the selected post treatment time points will be summarized. RNA analyses and microbiome analysis will be performed. Biomarker analyses are considered exploratory and will be summarized in separate technical reports.

Whole blood samples will be collected for genetic and epigenetic analyses as specified in the Time and Events Schedules of the protocol. Genetic and epigenetic (DNA) analyses will be conducted only in subjects who sign the consent form to participate in the genetic and epigenetic assessments. These analyses are considered exploratory and will be summarized in a separate technical report.

C-reactive protein, fecal lactoferrin, fecal calprotectin and histology assessments are included in the efficacy analysis section.

8. HEALTH ECONOMICS

The health economics analyses described in this section will be based on the **Primary Efficacy Analysis Set**. The data for each subject will be analyzed according to the assigned treatment regardless of the actual treatment received.

Medical resource utilization, including UC-related hospitalizations and UC-related surgeries, will be collected. Other heath economics data including Work Productivity and Activity Impairment Questionnaire-General Health (WPAI-GH) and the productivity visual analog scale (VAS) will be collected.

The productivity VAS will measure the impact of disease on subjects' daily productivity using a VAS (0=no impact at all to 10=impact productivity very much).

The Work Productivity and Activity Impairment Questionnaire-General Health (WPAI-GH) is a validated instrument created as a patient-reported quantitative assessment of the amount of absenteeism, presenteeism, and daily activity impairment attributable to general health. The WPAI-GH consists of 6 questions to determine employment status, hours missed from work due to UC, hours missed from work for other reasons, hours actually worked, the degree to which general health affected work productivity while at work, and the degree to which general health affected activities outside of work. Four impairment percentages (Percent work time missed due to health, Percent impairment while working due to health, Percent overall work impairment due to health and Percent activity impairment due to health) can be derived based on WPAI, each ranging from 0 to 100, with higher numbers indicating greater impairment and less productivity, i.e., worse outcomes.

No imputation will be performed for missing health economics values. The treatment failure rules only apply to WPAI-GH and productivity VAS endpoints. These endpoints are not adjusted for multiplicity; nominal p-values will be reported.

1. The proportion of subjects with a UC-related hospitalization through Week 44 will be summarized and compared between treatment groups using a CMH chi-square test, stratified by clinical remission status at maintenance baseline and induction treatment.

2. The time to the first UC-related hospitalization through Week 44 will be compared using the stratified log-rank test with clinical remission status at maintenance baseline and induction treatment as the stratification factors. The Kaplan-Meier curve by treatment group will be provided. The time to the first UC-related hospitalization is defined as the number of days elapsed from the date of the Week 0 study agent administration in this maintenance study to the date of the first hospitalization prior to or at Week 44. Subjects who are not hospitalized prior to Week 44 will be censored at Week 44, date of study completion, or early termination, whichever happens earlier.

3. The proportion of subjects with a UC-related surgery through Week 44 will be summarized and compared between treatment groups using a CMH chi-square test, stratified by clinical remission status at baseline and induction treatment.

4. The time to first colectomy (partial or full) through Week 44 will be compared between treatment groups using the stratified log-rank test with clinical remission status at baseline and induction treatment as the stratification factors. This analysis takes into account all reported colectomies that occurred prior to 44 weeks following the first maintenance study

agent administration. The time to the first colectomy is defined as the number of days elapsed from the date of Week 0 study agent administration in this maintenance study to the date of colectomy prior to or at Week 44. Subjects who do not have an event prior to Week 44 following the first maintenance study agent administration will be censored at Week 44, date of study completion, or early termination, whichever happens earlier.

5. The proportion of subjects with a UC-related hospitalization or surgery through Week 44 will be summarized and compared between treatment groups using a CMH chi-square test, stratified by clinical remission status at maintenance baseline and induction treatment.

6. The time to the first UC-related hospitalization or surgery through Week 44 will be compared using the stratified log-rank test with clinical remission status at maintenance baseline and induction treatment as the stratification factors. The Kaplan-Meier curve by treatment group will be provided. The time to the first UC-related hospitalization or surgery is defined as the number of days elapsed from the date of the Week 0 study agent administration in this maintenance study to the date of the first hospitalization or surgery prior to or at Week 44. Subjects who are not hospitalized nor had a surgery prior to Week 44 will be censored at Week 44, date of study completion, or early termination, whichever happens earlier.

7. The change from maintenance baseline in productivity VAS at Week 44 will be summarized and compared between treatment groups using ANCOVA on the van der Waerden normal scores with the respective baseline value, clinical remission status at maintenance baseline, induction treatment, and maintenance treatment group as covariates.

8. The change from maintenance baseline in each of four impairment percentages derived from WPAI-GH at Week 44 will be summarized and compared between treatment groups using ANCOVA on the van der Waerden normal scores with the respective baseline value, clinical remission status at maintenance baseline, induction treatment, and maintenance treatment group as covariates.

REFERENCES

1. Baron JH, Connell AM, Lennard-Jones JE. Variation between observers in describing mucosal appearances in proctocolitis. Brit Med J. 1964; 1:89-92.
2. Devlin N, Shah K, Feng Y, Mulhern B, Hout B V. Valuing Health-Related Quality of Life: An EQ-5D-5L Value Set for England. Office of Health Economics, Research Paper 16/01.
3. Feagan BG, Rutgeerts P, Sands BE, et al. Vedolizumab as induction and maintenance therapy for ulcerative colitis. N Engl J Med. 2013; 369:699-710.
4. Geboes K , Riddell R , Ost A et al. A reproducible grading scale for histological assessment of inflammation in ulcerative colitis. Gut 2000; 47: 404 – 409.
5. Higgins PDR, Schwartz M, Mapili J, Krokos I, Leung J, Zimmermann EM. Patient defined dichotomous endpoints for remission and clinical improvement in ulcerative colitis. Gut. 2005;54(6):782-788.
6. 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.
7. Sandborn WJ, Feagan BG, Marano C, et al. Subcutaneous golimumab maintains clinical response in patients with moderate-to-severe ulcerative colitis. Gastroenterology. 2014; 146:96-109.
8. Schroeder KW, Tremaine WJ, Ilstrup DM. Coated oral 5-aminosalicylic acid therapy for mildly to moderately active ulcerative colitis. A randomized study. N Engl J Med. 1987; 317(26):1625-1629.
9. 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.
10. Truelove SC, Witts LJ. Cortisone in ulcerative colitis: final report on a therapeutic trial. Brit Med J. 1955;2:1041-1048.

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: 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 3: 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 4: LIST OF PRIMARY AND MAJOR SECONDARY ENDPOINTS IN THE ORDER OF THE US-SPECIFIC TESTING

The following list details the order of the testing that will be performed at the 0.05 level of significance:

- Clinical remission at Week 44 in the ustekinumab high dose group (90 mg SC q8w)
- Maintenance of Clinical response through Week 44 in the ustekinumab high dose group (90 mg SC q8w)
- Endoscopic healing at Week 44 in the ustekinumab high dose group (90 mg SC q8w)
- Clinical remission and not receiving concomitant corticosteroids (corticosteroid-free clinical remission) at Week 44 in the ustekinumab high dose group (90 mg SC q8w)
- Clinical remission at Week 44 in the ustekinumab low dose group (90 mg SC q12w)
- Maintenance of Clinical response through Week 44 in the ustekinumab low dose group (90 mg SC q12w)
- Maintenance of clinical remission through Week 44 among the subjects who had achieved clinical remission at maintenance baseline in the ustekinumab high dose group (90 mg SC q8w)
- Endoscopic healing at Week 44 in the ustekinumab low dose group (90 mg SC q12w)
- Clinical remission and not receiving concomitant corticosteroids (corticosteroid-free clinical remission) at Week 44 in the ustekinumab low dose group (90 mg SC q12w)
- Maintenance of clinical remission through Week 44 among the subjects who had achieved clinical remission at maintenance baseline in the ustekinumab low dose group (90 mg SC q12w)