STATISTICAL ANALYSIS PLAN

NCT Number: NCT03259334

Study Title: A Phase 3 Randomized, Double-blind, Placebo-controlled, Parallel-group Efficacy and Safety Study of SHP647 as Induction Therapy in Subjects with Moderate to Severe Ulcerative Colitis (FIGARO UC 301)

Study Number: SHP647-301

SAP Version and Date:

Version 3.0: 04 Nov 2020

STATISTICAL ANALYSIS PLAN

SHP647
PHASE 3

A Phase 3 Randomized, Double-blind, Placebo-controlled, Parallel group Efficacy and Safety Study of SHP647 as Induction Therapy in Subjects with Moderate to Severe Ulcerative Colitis (FIGARO UC 301)

PROTOCOL IDENTIFIER: SHP647-301

Study Sponsor(s):	Shire Human Genetic Therapies, Inc. ("Shire"), 300 Shire Way, Lexington, MA 02421 USA
Author:	[redacted], [redacted] PPD
Protocol:	V1 (06 Jul 2017); Amendment 1, 05 Sep 2018; Amendment 2, 11 Nov 2019
SAP Version #:	3.0
SAP Date:	04 Nov 2020
Status:	Final

REVISION HISTORY

Version	Issue Date	Summary of Changes
1.0	30 Aug 2019	Final
2.0	13 Aug 2020	Incorporated Protocol Amendment 2 Updates. - Added Hypersensitivity AEs. - Changed study treatment name. - Changed sample size calculation. - Changed the term 'dimension' to 'domain' for IBDQ Domain Scores. - [redacted] Updated Prior medication definition to remove restriction for stop date. Due to the early discontinuation of the study before full enrollment, previously planned supplementary and subgroup analyses were removed. - Removed Per-protocol Set and Completers Set. - Removed supplementary analyses for efficacy endpoints (per-protocol, completer, missing data assumptions, etc). - Removed subgroup for efficacy endpoints beyond the actual randomization factors. Updated PCI percentage summary definition. Updated skip logic for Patient-reported UC signs and symptom data in Section 12.3.5. Added Section 12.5.2 Missing Date of Ulcerative Colitis Diagnoses. Added Section 7.5.3 Pregnancy Test and Follicle-stimulating Hormone Test. Added Section 7.5.4 Contraception Check. Added Section 7.5.5 Stool Microbiology. Added Section 7.5.6 Physical Examination. Added Section 10.1 Coronavirus. Updated Appendix 3 Rescue Therapy for UC. Moved Smoking history from Section 5.2 to Section 5.3. Minor editorial and formatting changes.
3.0	04 Nov 2020	Updated text in Section 6.1.1 and Section 6.2.1. Minor editorial and formatting changes.

TABLE OF CONTENTS

[redacted]

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDIX

ABBREVIATIONS

ADA	anti-drug antibodies
ANCOVA	analysis of covariance
AE	adverse event
BMI	body mass index
BLQ	below the limit of quantification
CI	confidence interval
CDF	cumulative distribution function
CMH	Cochran-Mantel Haenszel
[redacted]	[redacted]
DMC	data monitoring committee
eCRF	electronic case report form
ECG	electrocardiogram
ED	emergency department
[redacted]	[redacted]
ET	early termination
EOF	End of Follow-up
EOT	End of Treatment
FAS	full analysis set
FWER	family-wise type I error rate
HEOR	Health Economics and Outcomes Research
HRQL	health-related quality of life
IBDQ	Inflammatory Bowel Disease Questionnaire
ICF	informed consent form
LLOQ	lower limit of quantification
IP	investigational product
LTS	long-term safety extension
NAB	neutralizing antibodies
[redacted]	[redacted]
MedDRA	Medical Dictionary for Regulatory Activities
PCI	potentially clinically important
[redacted]	[redacted]
PDF	probability density function
[redacted]	[redacted] s
PML	progressive multifocal leukoencephalopathy
PRO	patient-reported outcome
QoL	quality of life

REML	restricted maximum likelihood
SAE	serious adverse event
SAP	statistical analysis plan
SC	subcutaneous/subcutaneously
SF-36	Short Form-36 Health Survey
SOC	system organ class
TEAE	treatment-emergent adverse event
TNF	tumor necrosis factor
[redacted]	[redacted]
UC	ulcerative colitis
VAS	visual analogue scale
WHO	World Health Organization
[redacted]	[redacted]

1. INTRODUCTION

This statistical analysis plan (SAP) provides a technical and detailed elaboration of the statistical analyses of efficacy, safety, and Health Economics and Outcomes Research (HEOR) data, as well as descriptive summaries of [redacted], as described in Protocol Amendment 2 dated 11 Nov 2019 (original protocol dated 06 Jul 2017 and Protocol Amendment 1 dated 05 Sep 2018). Specifications for tables, figures, and listings are contained in a separate document. The analysis plans for [redacted] and HEOR patient-reported outcome (PRO) validation are prepared separately.

On May 29, 2020, Takeda announced the decision to discontinue the ontamalimab clinical trial program in ulcerative colitis and Crohn's disease. The planned analyses have been updated to reflect the planned early discontinuation of this study.

2. OBJECTIVES, ESTIMAND(S), AND ENDPOINTS

2.1 Objectives

2.1.1 Primary Objective

To evaluate the efficacy of ontamalimab in inducing remission, based on composite score of patient-reported symptoms and centrally read endoscopy, in subjects with moderate to severe ulcerative colitis (UC).

2.1.2 Key Secondary Objectives

The key secondary objectives are as follows:

- To evaluate the efficacy of ontamalimab in achieving endoscopic remission, based on centrally read endoscopy
- To evaluate the efficacy of ontamalimab in achieving clinical remission, based on composite score of patient-reported symptoms
- To evaluate the efficacy of ontamalimab in inducing clinical response, based on composite score of patient-reported symptoms and centrally read endoscopy
- To evaluate the efficacy of ontamalimab in achieving mucosal healing, based on endoscopic and histological assessment using the Geboes Score grading system.

2.1.3 Other Secondary Objectives

The other secondary objectives are as follows:

- To evaluate the safety and tolerability of ontamalimab
- To evaluate the effect of ontamalimab induction treatment on other clinical and endoscopic outcomes (including Mayo-based remission and clinical response, partial Mayo score over time, clinical remission over time, endoscopic remission, and deep remission)
- To evaluate the effect of ontamalimab on abdominal pain, urgency, diarrhea, and absolute stool frequency and bleeding scores
- To evaluate the effect of ontamalimab on health-related quality of life (as measured by the Inflammatory Bowel Disease Questionnaire [IBDQ] and the Short Form-36 Health Survey [SF-36])
- To evaluate the effect of ontamalimab on incidence of hospitalizations and total inpatient days.

2.1.4 Exploratory Objective(s)

The exploratory objectives of the study are as follows:

[redacted]

2.2 Estimand(s)

The primary, key secondary, and secondary estimands are described in Table 1.

Table 1 List of Select Estimands

Estimand	Definition	Attributes			
		A: Population	**B: Variable**	**C: Strategy for addressing intercurrent event**	**D: Population-level summary**
Primary	The primary estimand is the effect of ontamalimab compared to placebo at Week 12 in remission	16-80 year old adolescents and adults with UC defined through inclusion and exclusion criteria as stated in the protocol	Remission at the Week 12 visit without rescue therapy defined in Appendix 3 and discontinuation	Composite: intercurrent events captured in variable definition.	Difference in proportions of subjects in remission at the Week 12 visit between each active treatment group (ontamalimab 25 mg, ontamalimab 75 mg) and placebo
Key Secondary	1st key secondary estimand is the effect of ontamalimab compared to placebo at Week 12 in impact on endoscopic remission	16-80 year old adolescents and adults with UC defined through inclusion and exclusion criteria as stated in the protocol	Endoscopic remission at the Week 12 visit, defined by centrally read endoscopic subscore 0 or 1 (modified, excludes friability), at the Week 12 visit without rescue therapy defined in Appendix 3 and discontinuation	Composite: intercurrent events captured in variable definition.	Difference in proportions of subjects in endoscopic remission at the Week 12 visit between each active treatment group (ontamalimab 25 mg, ontamalimab 75 mg) and placebo
Key Secondary	2nd key secondary estimand is the effect of ontamalimab compared to placebo at Week 12 in impact on clinical remission	16-80 year old adolescents and adults with UC defined through inclusion and exclusion criteria as stated in the protocol	Clinical remission at the Week 12 visit, defined by stool frequency subscore of 0 or 1 with at least a 1-point change from baseline in stool frequency subscore, and rectal bleeding subscore of 0, at the Week 12 visit without rescue therapy defined in Appendix 3 and discontinuation	Composite: intercurrent events captured in variable definition.	Difference in proportions of subjects in clinical remission at the Week 12 visit between each active treatment group (ontamalimab 25 mg, ontamalimab 75 mg) and placebo
Key	3rd key secondary	16-80 year old	Clinical response at the Week	Composite:	Difference in proportions

Table 1 List of Select Estimands

Estimand	Definition	Attributes			
		A: Population	**B: Variable**	**C: Strategy for addressing intercurrent event**	**D: Population-level summary**
Secondary	estimand is the effect of ontamalimab compared to placebo at Week 12 in impact on clinical response	adolescents and adults with UC defined through inclusion and exclusion criteria as stated in the protocol	12 visit, defined as the decrease from baseline in the composite score of patient-reported symptoms using daily e-diary and centrally read endoscopy of at least 2 points and at least 30%, with an accompanying decrease in the subscore for rectal bleeding ≥1 point or a subscore for rectal bleeding ≤1 without rescue therapy defined in Appendix 3 and discontinuation	intercurrent events captured in variable definition.	of subjects with clinical response at the Week 12 visit between each active treatment group (ontamalimab 25 mg, ontamalimab 75 mg) and placebo
Key Secondary	4th key secondary estimand is the effect of ontamalimab compared to placebo at Week 12 in impact on mucosal healing	16-80 year old adolescents and adults with UC defined through inclusion and exclusion criteria as stated in the protocol	Mucosal healing at the Week 12 visit, defined as mucosal healing based on endoscopic and histological assessment at the Week 12 visit. Mucosal healing is defined by centrally read endoscopic subscore 0 or 1 (modified, excludes friability) and centrally read Geboes score of ≤2 without rescue therapy defined in Appendix 3 and discontinuation	Composite: intercurrent events captured in variable definition.	Difference in proportions of subjects with mucosal healing at the Week 12 visit between each active treatment group (ontamalimab 25 mg, ontamalimab 75 mg) and placebo

2.3 Endpoints

2.3.1 Primary Endpoint

The primary efficacy endpoint is remission at the Week 12 visit. Remission is defined as a composite score of patient-reported symptoms using daily e-diary and centrally read endoscopy as follows:

- stool frequency subscore of 0 or 1 with at least a 1-point change from baseline

AND

- rectal bleeding subscore of 0

AND

- endoscopic subscore of 0 or 1 (modified, excludes friability).

2.3.2 Key Secondary Endpoints

- Endoscopic remission, as defined by centrally read endoscopic subscore 0 or 1 (modified, excludes friability), at the Week 12 visit.
- Clinical remission, as defined by stool frequency subscore of 0 or 1 with at least a 1-point change from baseline in stool frequency subscore, and rectal bleeding subscore of 0, at the Week 12 visit.
- Clinical response based on composite score at the Week 12 visit. Clinical response (composite) is defined as a decrease from baseline in the composite score of patient-reported symptoms using daily e-diary and centrally read endoscopy of at least 2 points and at least 30%, with an accompanying decrease in the subscore for rectal bleeding ≥1 point or a subscore for rectal bleeding ≤1.
- Mucosal healing based on endoscopic and histological assessment at the Week 12 visit. Mucosal healing is defined by centrally read endoscopic subscore 0 or 1 (modified, excludes friability) and centrally read Geboes score of ≤2.

2.3.3 Other Secondary Endpoints

- Remission, defined as a total Mayo score ≤2 with no individual subscore (stool frequency, rectal bleeding, endoscopy [modified, excludes friability], and physician's global assessment) exceeding 1, at the Week 12 visit.
- Clinical response based on total Mayo score at the Week 12 visit. Clinical response (Mayo) is defined as a decrease from baseline in the total Mayo score of at least 3 points and at least 30%, with an accompanying decrease in the subscore for rectal bleeding ≥1 point or a subscore for rectal bleeding ≤1.
- Partial Mayo score ≤2 with no individual subscore >1 at the Week 4, 8, and 12 visits. The partial Mayo score does not include the endoscopy subscore.

- Clinical remission as defined by stool frequency subscore of 0 or 1 with at least a 1-point change from baseline in stool frequency subscore, and rectal bleeding subscore of 0, at Weeks 4 and 8.
- Endoscopic remission at the Week 12 visit with endoscopic subscore of 0.
- Clinical remission at the Week 4, 8, and 12 visits with both rectal bleeding and stool frequency subscores of 0.
- Deep remission at the Week 12 visit. Deep remission is defined as both endoscopic and rectal bleeding subscores of 0, and stool frequency subscore ≤1 and a centrally read Geboes score of ≤2.
- Change from baseline at the Week 12 visit in abdominal pain, diarrhea, and urgency item scores, absolute stool frequency, absolute rectal bleeding and total sign/symptom score based on subject daily e-diary entries (average of rectal bleeding, stool frequency, abdominal pain, diarrhea, and urgency).
- Change from baseline in IBDQ domain and total (absolute) scores (time frame: Week 0, Week 8, up to Week 12, or ET).
- Change from baseline in SF-36, version 2, acute (physical and mental component summary scores and individual domain scores) to the Week 12/ET visit.
- Incidence of all-cause hospitalizations and total inpatient days.

2.3.4 Exploratory Endpoints

[redacted]

3. STUDY DESIGN

3.1 General Description

This is a Phase 3, randomized, double-blind, placebo-controlled, parallel-group study designed to evaluate the efficacy and safety of ontamalimab in inducing remission in subjects with moderate to severe UC.

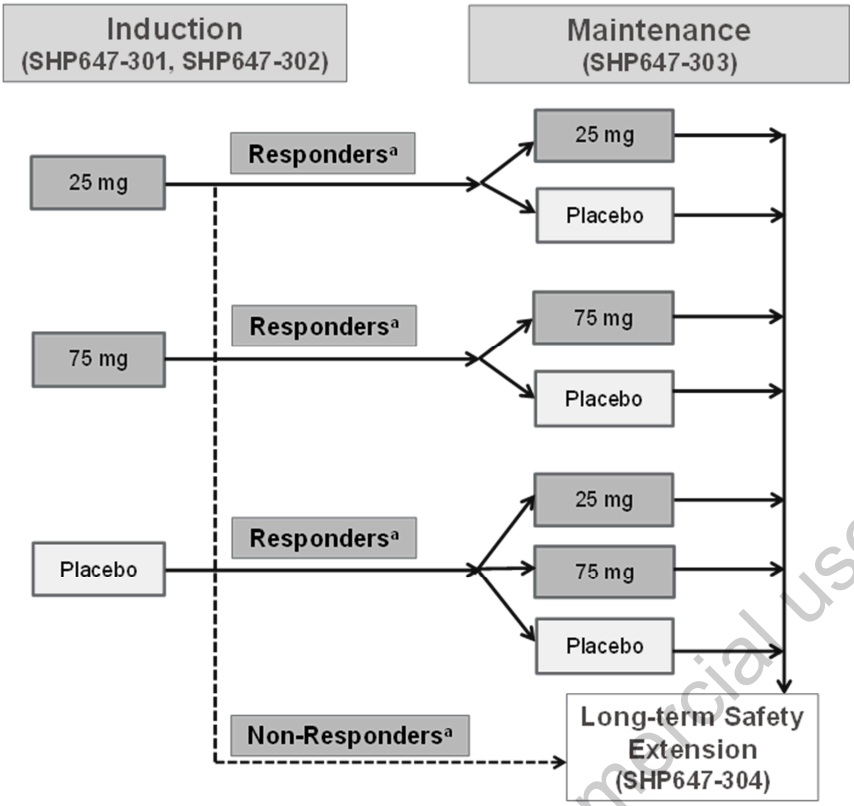
A total of 740 subjects (296 subjects at 25 mg ontamalimab, 296 subjects at 75 mg ontamalimab, and 148 subjects on placebo) are planned for enrollment into the study (Figure 1). Subjects must be at least 16 years of age and no more than 80 years of age at the time of signing the informed consent/assent form.

The study consists of a screening period up to 6 weeks and a 12-week treatment period. After the screening period, eligible subjects will be randomly assigned to receive 1 of 3 treatments (25 mg ontamalimab, 75 mg ontamalimab, or placebo) in a 2:2:1 ratio. Randomization will be stratified based upon the subject's status of prior anti- TNF treatment (naive or experienced) and glucocorticoid use at baseline (on glucocorticoids at baseline versus not on glucocorticoids at baseline). Subjects will receive SC injections of ontamalimab or placebo, using a prefilled syringe (PFS), on Week 0/Day 1 (Visit 2), Week 4 (Visit 4), and Week 8 (Visit 5). Subjects will undergo efficacy, [redacted], [redacted], safety, and health outcome assessments at these visits as detailed in Table A1.

At the end of the 12-week treatment period, eligible subjects will be offered the opportunity to participate in a double-blind maintenance study (SHP647-303; for subjects who achieve clinical response) or an LTS study (SHP647-304; for subjects who do not achieve a clinical response) as shown in Figure 1. Subjects who withdraw early from the 12-week treatment period or who do not wish to enter the maintenance study (SHP647-303) or LTS study (SHP647-304) will continue into a 16-week safety follow-up period. Only those subjects who complete the full course of investigational product treatment in the induction studies (SHP647-301 or SHP647-302) will be eligible to continue in the maintenance study or LTS study.

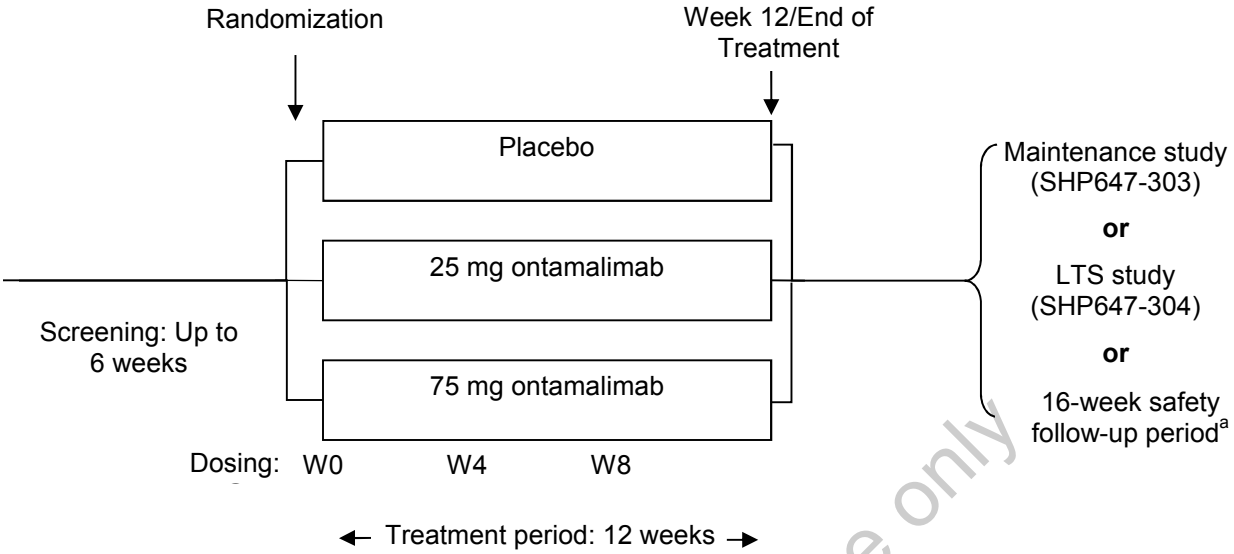
The overall study design is shown in Figure 2.

Figure 1 Overview of ontamalimab Phase 3 Studies in Ulcerative Colitis



[a] Clinical response is defined as:
1. A decrease from the induction study (SHP647-301 or SHP647-302) baseline in the composite score of patient-reported symptoms using daily e-diary and centrally read endoscopy of at least 2 points and at least 30%, with an accompanying decrease in the subscore for rectal bleeding ≥1 point or a subscore for rectal bleeding ≤1
OR
2. A decrease from the induction study (SHP647-301 or SHP647-302) baseline total Mayo score of at least 3 points and at least 30%, with an accompanying decrease in the rectal bleeding subscore of ≥1 point or an absolute rectal bleeding subscore of ≤1.

Figure 2 SHP647-301 Study Design Flow Chart



LTS=long-term safety extension; W=week
[a] Subjects who withdraw early from the 12-week treatment period or who do not wish to enter the maintenance study (SHP647-303) or LTS study (SHP647-304) will continue into a 16-week safety follow-up period.

3.2 Randomization

The actual treatment given to individual subjects is determined by a randomization schedule.

The total Mayo score will be calculated at baseline (Visit 2) before randomization. Endoscopic subscore based on the central reader's assessment will be used to determine eligibility. Subjects with a total Mayo score of ≥6, including a centrally read endoscopic subscore ≥2, rectal bleeding subscore ≥1, and stool frequency subscore ≥1, and who fulfill all other eligibility criteria, will be randomized in a ratio of 2:2:1 via a computer-generated randomization schedule to receive SC injections of 25 mg ontamalimab, 75 mg ontamalimab, or placebo, respectively. The randomization will be performed centrally and stratified by status of prior anti-TNF therapy (2 strata: Naive versus experienced) and glucocorticoid use at baseline (2 strata: on glucocorticoids at baseline versus not on glucocorticoids at baseline).

To ensure that the allocation of subjects with prior anti-TNF therapy exposure is similar to that observed in Study A7281009, the percentage of subjects with prior anti-TNF therapy exposure will be capped at 60% of the sample population. Additionally, in Japan only, enrollment of Japanese subjects with prior anti-TNF therapy exposure will be capped at 60%, to ensure comparability to the rest of the global study population. There will be no cap on the number of anti-TNF Naive subjects randomized.

Subject numbers are assigned to all eligible subjects as they consent to take part in the study. Within each site (numbered uniquely within a protocol), the subject number will be assigned to subjects according to the sequence of presentation for study participation.

The randomization number represents a unique number corresponding to investigational product allocated to the subject, once eligibility has been determined. Individual subject treatment will be automatically assigned by the IRT system.

Investigational product packaging identification numbers, separate from randomization numbers/unique identifiers, may also be assigned to subjects for specific treatment assignment as dictated by the study. In these cases, the same investigational product packing identification number may not be assigned to more than 1 subject.

3.3 Blinding

This is a double-blind, placebo-controlled study. All investigational and reference product (ontamalimab 25 mg, ontamalimab 75 mg, or placebo) will appear identical to protect the study blind.

Data that may potentially unblind the treatment assignment (eg, IP serum concentrations, antibodies to investigational product, treatment allocation, and IP preparation/accountability data) will be handled with special care during the data cleaning and review process. These data will be handled in such a way that, before unblinding, any data that may unblind study team personnel will be presented as blinded information or otherwise will not be made available. If applicable, unblinded data may be made available to quality assurance representatives for the purposes of conducting independent drug audits.

Whenever possible, the investigator or sub-investigator should contact the Shire physician and/or assigned medical monitor before breaking the blind. It is understood that in an emergency situation it may not be possible to communicate with the study team before breaking the blind. The safety of the subject should be of primary concern. When the blinding code is broken the reasons must be fully documented.

In the event that the treatment assignment is broken, the date, the signature of the person who broke the code, and the reason for breaking the code are recorded on the IRT and the source documents. Upon breaking the blind, the subject is withdrawn from the study, but should be followed up for safety purposes. The IRT will notify the relevant personnel in the event of any code break. Code-break information is held by the pharmacist/designated person at the site.

3.4 Sample Size and Power Considerations

Graphical methods are used to control the global family-wise Type I error rate (FWER) at the .05 level (2-sided) for the comparisons of the 2 ontamalimab treatment groups with the placebo group. Alpha is initially split equally at the .025 level (2-sided) for each of the pairwise treatment comparisons for the primary endpoint. Therefore, the power analysis and sample size estimation were calculated based on the chi-square test of proportions using nQuery Advisor Version 7.0 (Statistical Solutions Ltd, Cork, Ireland) for an individual ontamalimab dose compared with placebo.

Power calculations are made based on assuming a .025 (2-sided) significance level for each pairwise treatment comparison. Approximately 1346 subjects will be screened to randomize

740 subjects (2:2:1 allocation ratio: 296 subjects in the 25 mg ontamalimab treatment group, 296 subjects in the 75 mg ontamalimab treatment group, and 148 subjects in the placebo group) which would yield an approximately 90% power to detect individual pairwise treatment difference in the primary efficacy endpoint, remission at Week 12, of 11% (5% placebo versus 16% ontamalimab).

Expected remission rates at Week 12 are based on observed rates from the A7281009 study and placebo remission rates from literature (Feagan et al., 2013; Sandborn et al., 2017). No adjustment for missing data is required in these sample size calculations as subjects with missing data for remission at Week 12 are imputed as failures and the above rates account for these subjects.

With the sample size of 740, Table 2 provides the power for detecting a treatment difference between a ontamalimab treatment group and the placebo group for the key secondary endpoints.

Table 2 Power to Detect the Corresponding Treatment Effect for Key Secondary Endpoints

Key Secondary Endpoint at Week 12	ontamalimab Premise	Placebo Premise	Power
Endoscopic remission	24%	8%	0.98
Clinical remission	30%	16%	0.85
Clinical response by composite score	50%	35%	0.78
Mucosal healing	15%	5%	0.84

4. STATISTICAL ANALYSIS SETS

4.1 Screened Set

The Screened Set will consist of all subjects who have signed informed consent.

4.2 Randomized Set

The Randomized Set will consist of all subjects in the screened set for whom a randomization number has been assigned.

4.3 Safety Set

The Safety Set will consist of all subjects who have received at least 1 dose of investigational product. Analysis will be performed according to the treatment regimen actually received regardless of the randomized treatment regimen.

4.4 Full Analysis Set

The Full Analysis Set (FAS) will consist of all subjects in the randomized set who have received at least 1 dose of investigational product. Analysis will be performed according to the randomized treatment regimen regardless of the treatment regimen actually received.

5. STUDY SUBJECTS

5.1 Disposition of Subjects

The number of screened subjects and the number of screen failures will be presented in the overall column. The number of subjects included in each analysis set (ie, Randomized, Safety, FAS, [redacted],) will be summarized by treatment group, ontamalimab all doses and overall. The percentage, based on the number of subjects in the Safety Set, will be presented for FAS, [redacted] Sets. The study analysis set classifications of each subject will be listed for the Screened Set. Subjects excluded from efficacy analysis will be listed for the Randomized Set.

The number and percentage of subjects who completed and prematurely discontinued during the treatment and follow-up periods will be presented for each treatment group, ontamalimab all doses and overall for the Safety Set. Reasons for premature discontinuation from the treatment and follow-up periods as recorded on the termination page of the electronic case report form (eCRF) will be summarized (number and percentage) by treatment group, ontamalimab all doses, and overall for the Safety Set. The number and percentage of subjects who continued to the follow-up period and who continued to SHP647-303 or 304 studies will be presented for each treatment group, ontamalimab all doses and overall for the Safety Set.

The number and percentage of subjects who completed and prematurely discontinued the study will be presented for each treatment group, ontamalimab all doses and overall for the Safety Set. Subjects who complete 12 weeks treatment and rollover 303/304 or enter and complete safety follow-up will be considered completed the study. Reasons for premature discontinuation from the study are derived from reasons for premature discontinuation from the treatment and follow-up period. For subjects who discontinued from treatment, the reasons for discontinuation from the treatment will be presented regardless the status of safety follow up period. For subjects who completed the 12 weeks treatment and discontinued from safety follow up period, the reasons for discontinuation from the safety follow-up period will be presented. All subjects who prematurely discontinued during the treatment period, follow up period and study will be listed with their primary reason for discontinuation reason and duration of exposure for the Safety Set.

In addition, number of subjects screened, randomized and completed will be summarized for each site. The duration of enrollment, in days, will be summarized for each site, and overall. Duration of enrollment will be calculated as (last date of contact for any subject at that site - the first date of informed consent for any subject at that site +1).

5.2 Demographic and Other Baseline Characteristics

Descriptive summaries of demographic and baseline characteristics will be presented by each treatment group, ontamalimab all doses and overall for the Safety Set and FAS.

Subject's age is calculated as the difference between the date of birth and the date of informed consent. If day of birth is missing then the day will be imputed as 1, if both the day and month of birth are missing then the day will be imputed as 1 and the month will be imputed as 1 (January). The following demographic characteristics will be summarized in the following order in the tables: age, age category (<18, 18-<65 and >=65; <35 and >=35), sex (Male, Female), ethnicity

(Hispanic or Latino, Not Hispanic or Latino, Not Reported, Unknown), region (North America, Western Europe, Eastern Europe, Asia (Japan/South Korea), ROW (Africa/Australia/Latin America/Middle East) in Appendix 2, race (American Indian or Alaska Native, Asian (Japanese, Korean, Other), Black or African American, White, Native Hawaiian or Other Pacific Islander, and Other), Japanese Ancestry (Currently living in Japan, Born in Japan and currently living outside of Japan for less than 5 years, and Other) and Korean Ancestry (Currently living in Korea, Born in Korea and currently living outside of Korea for less than 5 years, and Other).

Baseline is defined as the last assessment prior to the first administration of the IP unless otherwise specified. The following baseline characteristics will be summarized:

- Weight,
- Height,
- Body mass index (BMI),
- UC disease duration and UC disease duration category (<1 years, >=1 - <3 years, >=3 - <7 years and >=7 years),
- Total Mayo severity and total Mayo severity category (< 6, 6 - <9, and >=9),
- Stool frequency score (0,1,2 and 3),
- Rectal bleeding score (0,1,2 and 3),
- Findings of endoscopy (0,1,2 and 3),
- Physician global assessment (0,1,2 and 3),
- [redacted]
- UC disease location (Proctitis, Procto-Sigmoidits, Left-sided Colitis, Extensive Colitis/Pancolitis)

Disease duration is the number of years from the date of UC diagnosis to the date of informed consent.

The following UC medication history/ use will be summarized.

- Immunosuppressant Experienced (Yes vs. No),
- Immunosuppressant Use at Baseline (Yes vs. No),
- 5-ASA Use at Baseline (Yes vs. No),
- Anti-TNF Experienced (Experienced vs. Naive) (both randomized status and actual status),
- Anti-TNF Failure (Yes vs. No),
- Anti-TNF Failure Times (Anti-TNF Naive, Anti-TNF Experienced without Failure, Failed 1 Anti-TNF Therapy, Failed 2 Anti-TNF Therapies, Failed 3 or more Anti-TNF Therapies),
- Glucocorticoid Use at Baseline (Yes vs. No) (both randomized status and actual status),

- Glucocorticoid Use at Baseline (Systemic and Topical, Systemic Only, Topical Only, None),
- Systemic Glucocorticoid Dose at Baseline,
- Systemic Glucocorticoid Dose at Baseline Category (<=10 mg, >10 mg),
- Maximum Prior Treatment Experience (Aminosalicylates experienced, Glucocorticoid experienced (further broken down into topical glucocorticoid experienced and systemic glucocorticoid experienced), Immunosuppressant experienced or Biologic failure, Immunosuppressant experienced and Biologic failure),
- Glucocorticoid Use at Baseline AND Immunosuppressant Use at Baseline (Both Glucocorticoid and Immunosuppressant Use, Only Glucocorticoid Use, Only Immunosuppressant Use, Neither Glucocorticoid nor Immunosuppressant Use).

5.3 Smoking history

Smoking history will be summarized for each treatment group and ontamalimab all doses.

5.4 Medical History

Medical history will be collected at the Screening Visit (Visit 1), including UC history and cardiac history. Medical history will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) Version 19.1 2016. The number and percentage of subjects who have general medical history will be summarized by each treatment group and ontamalimab all doses, system organ class (SOC), and preferred term. General medical history will be listed for the Safety Set.

Cardiovascular history information will be summarized by each treatment group and ontamalimab all doses. Cardiovascular history will be listed for the Safety Set.

UC history will be listed for the Safety Set.

5.5 Prior Procedures and Medication

Prior medications will be coded using the World Health Organization (WHO) Drug Dictionary dated December 01, 2016.

Prior medication is defined as any medication with the start date prior to the date of the first dose of investigational product. Incomplete medication dates will be imputed as described in Section 12.5.3.

Prior medication for Indication Under Study usage will be summarized by the number and percentage of subjects in each treatment group and ontamalimab all doses, for subjects receiving each medication within each therapeutic class and preferred term for the Safety Set. Multiple medication usage by a subject in the same category will be counted only once.

All prior medications and medical/surgical procedures will be listed for the Safety Set.

5.6 Concomitant Therapies, Procedures and Medications

Concomitant medications will be coded using the World Health Organization (WHO) Drug Dictionary dated 01 Dec 2016. Concomitant medication is defined as any medication with a start date prior to the date of the first dose of IP and continuing after the first dose of IP or with a start date between the dates of the first dose of IP and end of treatment date, inclusive. Medication that starts after the first dose of SHP647-303/304 IP will be collected in SHP647-303/304 database and will not be considered as concomitant medication in SHP647-301. Incomplete medication dates will be imputed as described in Section 12.5.3. Any medication with a start date between the dates of the first dose of IP and end of treatment date, inclusive, or with a start date after the end of treatment date (post-treatment) will be considered a post-treatment concomitant medication.

Concomitant medication usage will be summarized by the number and percentage of subjects in each treatment group and ontamalimab all doses, for subjects receiving each medication within each therapeutic class and preferred term for the Safety Set. Multiple medication usage by a subject in the same category will be counted only once. Summaries are presented separately for "Indication Under Study" and "not for Indication Under Study".

All medications, medical/surgical procedures and therapies will be listed for the Safety Set.

5.7 Exposure to Investigational Product

Investigational product (ontamalimab or placebo) will be administered SC every four weeks (Weeks 0, 4, and 8). Exposure to IP will be summarized by presenting the number of subjects who had 1 injection, 2 injections or 3 injections. Number of injections received will be summarized in each treatment group and ontamalimab all doses.

Exposure to IP for the Safety Set will be summarized in terms of treatment duration, which is calculated as the number of days from the date of first dose of IP taken to the date of the last dose of IP taken +29 days. Subject years of exposure is calculated as (Date of Last Dose of investigational product in this study – Date of First Dose of investigational product in this study +29)/365.25. Total subject years of exposure is calculated by summing of each subject years of exposure within each column.

Descriptive statistics (n, mean, standard deviation, minimum, median, and maximum) will be presented to describe the exposure to IP by each treatment group and ontamalimab all doses.

5.8 Measurements of Treatment Compliance

Compliance for the treatment period is defined as the total number of SC injections administered from the start of the treatment until the end of the treatment divided by the number of injections expected to be taken during that time period, times 100. Percentage compliance will be summarized by treatment group.

5.9 Protocol Deviations

Protocol deviations will be recorded by Pharmaceutical Product Development (PPD) separately from the clinical database. PPD/Shire will classify significant and nonsignificant protocol deviations per the agreed protocol deviation management plan. The Shire study team will review the protocol deviations and their classification throughout the study and before treatment unblinding and database lock.

For any criteria for protocol deviations that can be completely implemented by a computer program, the detailed algorithm will be agreed upon. Details of such algorithms will be included in the derived dataset specifications and finalized before treatment unblinding.

Confirmed significant and nonsignificant protocol deviations will be documented in the Protocol Deviation tracker for the study. Significant and nonsignificant protocol deviations will be summarized by category and site for each treatment group, ontamalimab all doses and overall, for the Randomized Set. Significant and nonsignificant protocol deviations will be listed for the Randomized Set.

6. EFFICACY ANALYSES

All efficacy analyses will be based on the FAS unless stated otherwise. Baseline for all efficacy analyses is defined as the last observed value for the efficacy assessment prior to taking the first dose of IP (based on dates or date/times) unless otherwise specified.

All efficacy analyses will be conducted according to the randomized treatment, regardless of the treatment actually received.

All confidence intervals (CIs) will be 2-sided 95% CIs, unless stated otherwise. Control of Type I error is discussed in Section 6.4 in detail.

For continuous endpoints, descriptive summary statistics will be presented by treatment group at each scheduled visit and will include the following: n, mean, median, standard deviation, minimum and maximum. For binary endpoints, number and percentage of subjects in each category will be summarized by treatment group at each scheduled visit.

6.1 Analyses of Primary Efficacy Endpoint(s)

The primary efficacy endpoint is remission at the Week 12 visit. Remission is defined as a composite score of patient-reported symptoms using daily e-diary and centrally read endoscopy as follows:

- stool frequency subscore of 0 or 1 with at least a 1-point change from baseline

AND

- rectal bleeding subscore of 0

AND

- endoscopic subscore of 0 or 1 (modified, excludes friability).

The primary efficacy endpoint will be compared for each active treatment group (25 mg or 75 mg ontamalimab) to the placebo group using a Cochran-Mantel Haenszel (CMH) chi-square test stratified by actual status of prior anti-TNF treatment and glucocorticoid use at baseline. Subjects with the intercurrent events defined in Section 2.2 and/or with missing remission data at the Week 12 will be considered failures and counted as non-remission. The endoscopy score will be based on centrally read results.

The primary endpoint will be tested by the following hypothesis:

$$H_0: \delta = 0$$

$$H_1: \delta \neq 0$$

where δ is the common treatment difference across strata, j=1 to m. The common treatment difference is a weighted average of the stratum-specific treatment differences.

The estimate of the common treatment difference along with the corresponding stratified Newcombe 95% CI using the method of Yan and Su (2010) and CMH p-value will be presented for each active treatment group to placebo comparison.

The proportion of subjects in remission at Week 12 will be plotted by each treatment group; by actual status of prior anti-TNF treatment and each treatment group; by actual glucocorticoid use at baseline and each treatment group.

6.1.1 Sensitivity Analyses of Primary Efficacy Endpoint

If significant noncompliance with regulatory requirements during the course of the study is detected or reported at any clinical site, additional sensitivity analyses may be conducted on the primary efficacy endpoint by using the same approach but excluding all subjects from the noncompliant site(s).

6.2 Analyses of Key Secondary Efficacy Endpoint(s)

Similarly to the primary endpoint, each secondary endpoint will be summarized by treatment group. Subjects with the intercurrent events defined in Section 2.2 and/or missing key secondary endpoint data at the Week 12 visit are considered failures. Each of the key secondary endpoints will be analyzed using the same approach as described for the primary efficacy endpoint using the following hypothesis:

$$H_0: \delta = 0$$

$$H_1: \delta \neq 0$$

Where δ is the common treatment difference across strata, j=1 to m. The common treatment difference is a weighted average of the stratum-specific treatment differences.

The estimate of the common treatment difference along with the corresponding stratified Newcombe 95% CI using the method of Yan and Su (2010) and CMH p-value will be presented for each active treatment group to placebo comparison.

Proportion of subjects achieve secondary endpoint at Week 12 will be plotted by each treatment group; by actual status of prior anti-TNF treatment and each treatment group; by actual glucocorticoid use at baseline and each treatment group.

The key secondary efficacy endpoints are:

- Endoscopic remission, as defined by centrally read endoscopic subscore 0 or 1 (modified, excludes friability), at the Week 12 visit.
- Clinical remission, as defined by stool frequency subscore of 0 or 1 with at least a 1-point change from baseline in stool frequency subscore, and rectal bleeding subscore of 0, at the Week 12 visit.

- Clinical response based on composite score at the Week 12 visit. Clinical response (composite) is defined as a decrease from baseline in the composite score of patient-reported symptoms using daily e-diary and centrally read endoscopy of at least 2 points and at least 30%, with an accompanying decrease in the subscore for rectal bleeding ≥1 point or a subscore for rectal bleeding ≤1.
- Mucosal healing based on endoscopic and histological assessment at the Week 12 visit. Mucosal healing is defined by centrally read endoscopic subscore 0 or 1 (modified, excludes friability) and centrally read Geboes score of ≤2.

6.2.1 Sensitivity Analyses of Key Secondary Efficacy Endpoints

If significant noncompliance with regulatory requirements during the course of the study is detected or reported at any clinical site, additional sensitivity analyses may be conducted on the key secondary endpoints using the same approach but excluding all subjects from the noncompliant site(s).

6.3 Analyses of Other Secondary Efficacy Endpoints

Other secondary endpoints will be summarized by descriptive statistics at each visit the endpoint is assessed and presented by treatment group. In general, the following analysis methods, where specified, will be used to analyze the other secondary endpoints.

Binary endpoints will be compared for each active treatment group (25 mg or 75 mg ontamalimab) to the placebo group using a CMH chi-square test stratified by actual status of prior anti-TNF treatment and glucocorticoid use at baseline. Subjects with the intercurrent events defined in Section 2.2 and/or missing binary endpoint data at a visit will be considered failures (Yan and Su, 2010).

The estimate of the common treatment difference along with the corresponding stratified Newcombe 95% CI using the method of Yan and Su (2010) and CMH p-value will be presented for each active treatment group to placebo comparison.

Continuous endpoints that are measured repeatedly over time will be analyzed using a linear repeated measures mixed model with restricted maximum likelihood (REML) estimation. The model will include fixed effects for treatment group (categorical), visit (categorical), treatment group by visit interaction, actual status of prior anti-TNF treatment (categorical), and glucocorticoid use at baseline (categorical); baseline value as a continuous covariate; and repeated measures across visit for subject. From this model, estimates of least squares means, treatment differences, standard errors, p-values, size effect and 95% CIs for least squares mean treatment differences for each visit will be provided. The REML estimation, which is default in SAS, will be utilized along with the Kenward-Roger method for estimating the covariance matrix and degrees of freedom. The model will use an unstructured covariance type. If the fit of the unstructured covariance matrix fails to converge, the following covariance structures will be tried in order until convergence is reached: Heterogeneous Toeplitz, Toeplitz, Autoregressive (1) and Compound Symmetry.

The other secondary endpoints and analyses are noted below.

- Remission, defined as a total Mayo score ≤2 with no individual subscore (stool frequency, rectal bleeding, endoscopy [modified, excludes friability], and physician's global assessment) exceeding 1. A CMH chi-square test that is described in Section 6.3 will be used to compare the proportion of subjects with remission for each active treatment group to placebo group in remission at Week 12.
- Clinical response based on total Mayo score at the Week 12 visit. Clinical response (Mayo) is defined as a decrease from baseline in the total Mayo score of at least 3 points and at least 30%, with an accompanying decrease in the subscore for rectal bleeding ≥1 point or a subscore for rectal bleeding ≤1. A CMH chi-square test that is described in Section 6.3 will be used to compare the proportion of subjects with clinical response for each active treatment group to placebo group with clinical response at Week 12.
- Partial Mayo remission, defined as partial mayo score ≤2 with no individual subscore >1 at the Weeks 4, 8, and 12 visits. A CMH chi-square test that are described in Section 6.3 will be used to compared for the proportion of subjects with partial mayo remission for each active treatment group to placebo group in partial mayo remission at Weeks 4, 8, and 12.
- Clinical remission as defined by stool frequency subscore of 0 or 1 with at least a 1-point change from baseline in stool frequency subscore, and rectal bleeding subscore of 0, at Week 4 and 8. A CMH chi-square test that is described in Section 6.3 will be used to compare the proportion of subjects with clinical remission for each active treatment group to placebo group in clinical remission at Week 4 and Week 8.
- Endoscopic remission at the Week 12 visit with endoscopic subscore of 0. A CMH chi-square test that is described in Section 6.3 will be used to compare the proportion of subjects with endoscopic remission for each active treatment group to placebo group in endoscopic remission at Week 12.
- Clinical remission at the Week 4, 8, and 12 visits with both rectal bleeding and stool frequency subscores of 0. A CMH chi-square test that is described in Section 6.3 will be used to compare the proportion of subjects with clinical remission for each active treatment group to placebo group in clinical remission at Week 4 and Week 8.
- Deep remission at the Week 12 visit. Deep remission is defined as both endoscopic and rectal bleeding subscores of 0, and stool frequency subscore ≤1 and a centrally read Geboes score of ≤2. A CMH chi-square test that is described in Section 6.3 will be used to compare the proportion of subjects with deep remission for each active treatment group to placebo group in deep remission at Week 12.
- The absolute and change from baseline in abdominal pain, urgency, diarrhea, absolute stool frequency, and absolute rectal bleeding raw and categorized scores as well as the total sign/symptom score based on subject daily e-diary entries (average of rectal bleeding, stool frequency, abdominal pain, diarrhea, and urgency) will be summarized by descriptive statistics and presented by visits and each treatment group. The linear repeated measures mixed model that is described in Section 6.3 will be used to analyze change from baseline in the raw, categorized, and total scores with the change from baseline at each timepoint as the outcome variable. The cumulative distribution function (CDF) and probability density function (PDF) of change from baseline in the above scores at Week 12 will be plotted. The

proportion of subjects with remission at Week 4, Week 8 and Week 12 will be summarized by treatment group and analyzed using a using a logistic regression with the effects of treatment group, actual status of prior anti-TNF treatment and glucocorticoid use at baseline. The CDF of subjects in remission at Week 4, Week 8 and Week 12 will be plotted by each treatment group. The categorized scoring systems as well as the threshold for remission based on the total score will be confirmed after database lock according to the UC Patient-Reported Outcome (PRO) SAP and are subject to change after the psychometric assessment has been completed.

- IBDQ total score and domain scores will be summarized by treatment group at baseline, Week 8, and Week 12. Change from baseline will also be summarized by treatment at Week 8 and Week 12. IBDQ individual scores, total score, and domain scores will be listed at all the visits. The linear repeated measures mixed model that is described in Section 6.3 will be used to analyze change from baseline in IBDQ total score and domain scores at Week 8 and Week 12.
- Subjects with an improvement of >=16 points in the IBDQ total score will be designated as having an improvement in IBDQ total score. The proportion of subjects with IBDQ total score improvement at Week 8 and Week 12 will be summarized by treatment group and analyzed using logistic regression with the effects of treatment group, actual status of prior anti-TNF treatment and glucocorticoid use at baseline. The odds ratio for subjects with IBDQ total score improvement for the 25 mg vs placebo group and 75 mg vs placebo group, and the associated 95% CIs will be estimated from the models. The CMH chi-square test that is described in Section 6.3 will also be used to compare each active treatment group to placebo group in IBDQ total score improvement at Week 8 or Week 12. The CDF and PDF of IBDQ change from baseline at Week 12 will be plotted.
- Subjects with a total score >=170 will be classified as achieving IBDQ clinical remission. The proportion of subjects with IBDQ clinical remission at Week 8 and Week 12 will be analyzed using a logistic regression with the effects of treatment group, actual status of prior anti-TNF treatment and glucocorticoid use at baseline. The odds ratio for subjects in remission for the 25 mg vs placebo group and 75 mg vs placebo group, and the associated 95% CIs will be estimated from the models. The CMH chi-square test that is described in Section 6.3 will also be used to compare each active treatment group to placebo group in IBDQ clinical remission at Week 8 or Week 12.
- Each SF-36 domain score, the physical component summary score, and the mental component summary score will be summarized by treatment group at baseline, Week 8, and Week 12. Change from baseline will also be summarized by treatment at Week 8 and Week 12. The linear repeated measures mixed model that is described in Section 6.3 will be used to analyze change from baseline in SF-36 domain and component summary scores at Week 8 and Week 12. Each SF-36 domain score, the physical component summary score, and the mental component summary score will be listed at all the visits.
- Descriptive summary statistics of all-cause hospitalizations, gastrointestinal related hospital admission, other illness/problem related hospital admission, gastrointestinal related surgical procedures during hospital admission will be presented by treatment group during the entire treatment period and during the follow-up period, and will include the following: Number of Events and Number of Subjects with Events.

- Descriptive summary statistics of total inpatient days will be presented by treatment group during the entire treatment period and during the follow-up period, and will include the following: n, mean, median, standard deviation, minimum and maximum.

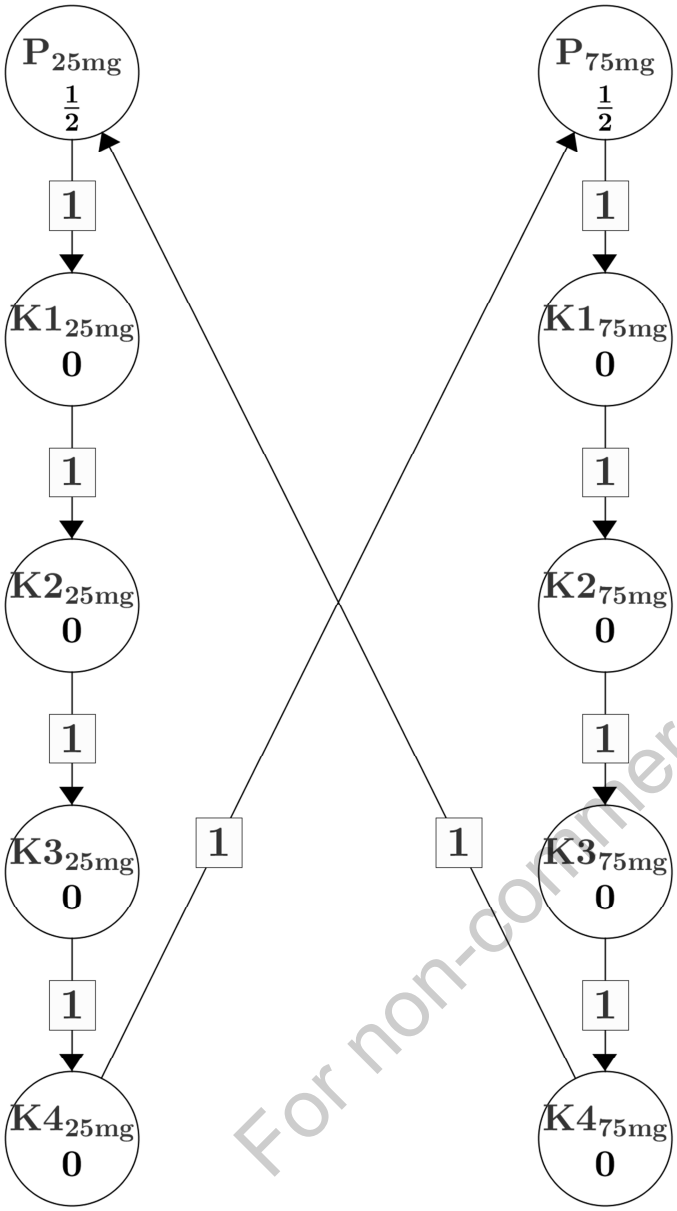
6.3.1 Sensitivity Analyses of Other Secondary Efficacy Endpoint(s)

No sensitivity analyses of other secondary efficacy endpoints are planned for this study.

6.4 Multiplicity Adjustment

The global FWER for the statistical tests of the primary and key secondary endpoints will be strongly controlled at .05 (2-sided). To control the FWER, graphical methods discussed in Bretz et al (2009) will be utilized to propagate α from primary to key secondary endpoints and between the two ontamalimab treatment group and placebo comparisons. Alpha is initially split equally at the .025 level (2-sided) for each of the pairwise treatment comparisons for the primary endpoint (P) and alpha is propagated in a hierarchical manner to each of the 4 key secondary endpoints (K1-K4) within a pairwise treatment comparison. A graphical visualization of the α propagation is presented in Figure 3. Statistical results with the p-value and significance for the primary and each key secondary endpoint will be presented.

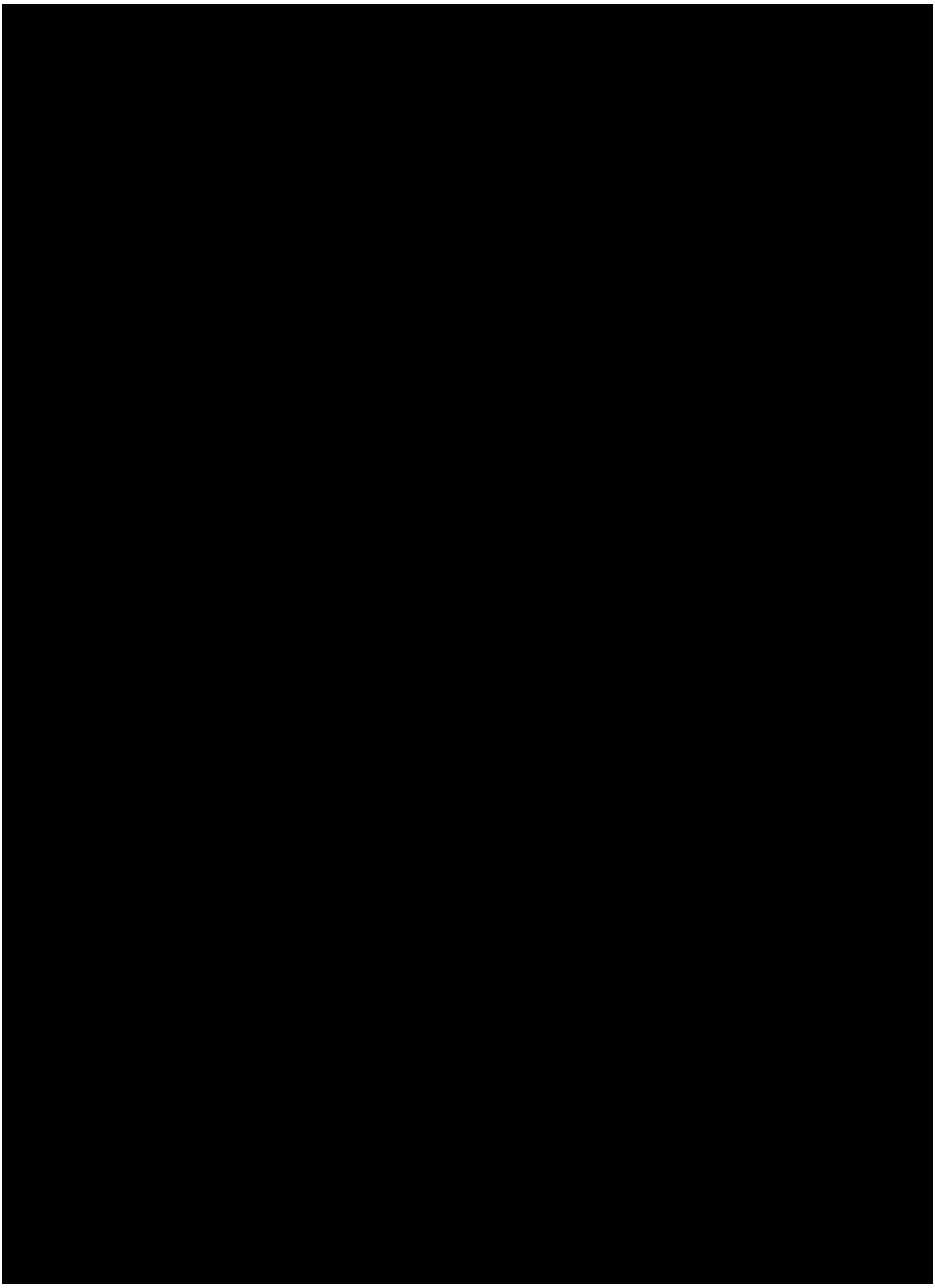
Figure 3 Visualization of Alpha Propagation



Only p-values that are significant according to this graphical approach are inferential and statistically significant. All other p-values are descriptive.

6.5 Analyses of Exploratory Endpoint(s)

[redacted]



6.6 Subgroup Analyses

Subgroup analyses are planned for the primary endpoint, the key secondary endpoints and the first other secondary endpoint (remission using total Mayo score). These endpoints will be derived by using the same strategy for handling intercurrent events and missing data as was done for endpoints in the overall population.

Within subgroups, each endpoint will be summarized by treatment group and will be compared for each active treatment group (ontamalimab 25 mg and ontamalimab 75 mg) with the placebo group using a Chi-square test. The 95% CI for difference in proportion is based on unstratified Wilson CI.

Subgroup variables are:

- Actual Anti-TNF Experienced (Experienced vs. Naive)
- Actual Glucocorticoid Use at Baseline (Yes vs. No)

7. SAFETY ANALYSIS

The safety analyses will be performed using the Safety Set. Safety variables include AEs, clinical laboratory variables, vital signs, electrocardiogram (ECG) variables, Anti-Drug Antibodies (ADA) and Neutralizing Antibodies (NAB) variables, and neurological variables. For each safety variable, the last value collected prior to the first dose of double-blind IP will be used as baseline for all analyses of that safety variable. A Final on-Treatment Assessment will be defined as the last valid assessment obtained after baseline and through the end of the treatment visit.

7.1 Adverse Events

Adverse events will be coded using MedDRA Version 19.1 2016.

Treatment-emergent AEs (TEAEs) are defined as AEs with start dates at the time of or following the first exposure to investigational product.

An overall summary of the number of subjects with TEAEs will be presented, including the number and percentage of subjects with any TEAEs, serious TEAEs, TEAEs related to investigational product, related serious AEs (SAE), and TEAEs leading to discontinuation of investigational product, and TEAEs leading to death.

The number of events, incidence, and percentage of subjects reporting TEAEs in each treatment group will be presented by treatment group and for ontamalimab all doses; by preferred term; by SOC and preferred term; by SOC, preferred term, and maximum severity. TEAEs considered related to IP will also be summarized by SOC and preferred term. If more than 1 AE occurs with the same preferred term for the same subject, then the subject will be counted only once for that preferred term using the most severe and most related occurrence for the summarization by severity and by relationship to investigational product.

Most common TEAE (incidence ≥2% in any treatment group) will be summarized by preferred term, in each treatment group and for ontamalimab all doses.

Serious TEAEs, TEAEs leading to discontinuation of the study or study medication and injection site adverse events will be summarized by SOC, preferred term, in each treatment group and ontamalimab all doses. Adverse events leading to withdrawal, serious AEs, and deaths will be similarly summarized or listed.

7.1.1 Adverse Events of Special Interest and other Potential Risks

There is one identified important potential risk of Progressive Multifocal Leukoencephalopathy. There are six other identified potential risks: immunotoxicity, immunogenicity, infection, vascular and thrombotic events, local tolerability and malignant tumours. Potential risks will be summarized by treatment group and for ontamalimab all doses. Important potential risk will be listed.

7.1.1.1 Hypersensitivity

Potential hypersensitivity, serum sickness, vasculitis, and arthus reactions to ontamalimab will be regarded as AESIs.

An external hypersensitivity adjudication committee is established to review reported hypersensitivity events and adjudicate whether the event was a hypersensitivity event, which Type (Type I or Type III) and recommendations of permanent discontinuation or re-challenge with investigational product. Reported hypersensitivity events, adjudicated hypersensitivity events and study drug recommendation will be summarized by treatment group and for ontamalimab all doses.

The number of events, incidence, and percentage of subjects reporting Treatment-emergent hypersensitivity events in each treatment group will be presented by treatment group and for ontamalimab all doses; by SOC, preferred term and hypersensitivity type.

Reported hypersensitivity events and adjudicated hypersensitivity events will be listed for the safety set.

7.2 Clinical Laboratory Data

Descriptive statistics for clinical laboratory values (in conventional units) and changes from baseline at each assessment time point for quantitative variables will be presented by treatment group and for ontamalimab all doses for the following clinical laboratory variables.

Serum chemistry

- alkaline phosphatase
- AST
- ALT
- total bilirubin
- total protein
- albumin
- glucose
- blood urea nitrogen
- creatinine
- sodium
- potassium
- chloride
- calcium
- carbon dioxide

Hematology

- hemoglobin
- hematocrit
- mean corpuscular hemoglobin
- mean corpuscular hemoglobin concentration
- mean corpuscular volume
- erythrocyte (red blood cell) count
- leukocyte (white blood cell) count
- neutrophils
- lymphocytes
- monocytes
- eosinophils
- basophils
- platelet count

Urinalysis

- glucose
- protein
- specific gravity
- pH
- nitrite
- bilirubin
- ketones
- hemoglobin
- urobilinogen
- leukocyte esterase

Clinical laboratory test values are potentially clinically important (PCI) if they meet either the low or high PCI criteria listed in Table 3. The number and percentage of subjects with post-baseline PCI values will be tabulated by treatment group. The percentages will be calculated relative to the number of subjects with at least 1 post-baseline assessment. The numerator is the total number of subjects with at least 1 post-baseline PCI value. A supportive listing of subjects with post-baseline PCI values will be provided including the subject number, site, baseline, and post-baseline values.

Figures will be presented for hematology and chemistry to show the changes in lab parameters over time. Data will be presented as box-and-whisker plots for each treatment group (placebo, ontamalimab 25 mg, and ontamalimab 75 mg) at each visit, with one lab parameter per page.

Shifts (Low, Normal and High) from baseline to each visit will be presented by treatment group and for ontamalimab all doses for hematology, chemistry and urinalysis.

All laboratory data will be listed for the Safety Set.

Table 3 Criteria for Potentially Clinically Important Laboratory Tests

Parameter	Age Range	Sex	Outlier Criteria[a]	
			Low	High
Hematology				
Hemoglobin	All		<8 g/dL	NA
Hematocrit	All		<32%	NA
Mean Corpuscular Hemoglobin	All		<LLN	>ULN
Mean Corpuscular Hemoglobin Concentration	All		<LLN	>ULN
Mean Corpuscular Volume	All		<LLN	>ULN
Erythrocyte (red blood cell)	All		<3.0 x 10^6/µL	NA
Leukocytes (white blood cell)	All		<3.0 x 10^3/µL	>20 x 10^3/µL
Neutrophils (Abs)	All		<1.5 x 10^3/µL	>15 x 10^3/µL
Neutrophils (%)	All		<40%	NA
Lymphocytes (Abs)	All		NA	NA
Lymphocytes (%)	All		<10%	>50%
Monocytes (Abs)	All		NA	NA
Monocytes (%)	All		NA	>25%
Eosinophils (Abs)	All		NA	NA
Eosinophils (%)	All		NA	>10%
Basophils (Abs)	All		NA	NA
Basophils (%)	All		NA	>10%
Platelets	All		<75 x 10^3/µL	>1,000 x 10^3/µL
Chemistry				
Alkaline Phosphatase	All		NA	>2.5 x ULN (or alternatively >400 U/L)
Aspartate Aminotransferase (AST)	All		NA	>2.5 x ULN
Alanine Aminotransferase (ALT)	All		NA	>2.5 x ULN
Total Bilirubin	All		NA	>1.5 x ULN
Total Protein, plasma or serum	All		<5 g/dL	>9 g/dL
Albumin	All		<3 g/dL	NA
Glucose (fasting)	All		<55 mg/dL	>160 mg/dL
Blood Urea Nitrogen (BUN)	All		NA	>2.5 x ULN (or alternatively >29.4 mg/dL)
Creatinine, serum	All		NA	>1.5 x ULN (or alternatively >1.98 mg/dL)
Sodium	All		<130 mEq/L	>150 mEq/L

Table 3 Criteria for Potentially Clinically Important Laboratory Tests

Parameter	Age Range	Sex	Outlier Criteria[a]	
			Low	High
Potassium, plasma or serum	All		<3 mEq/L	>5.5 mEq/L
Chloride	All		<90 mEq/L	>115 mEq/L
Calcium	All		<8.0 mg/dL	>11.2 mg/dL
Carbon dioxide (NCI uses bicarb)	All		NA	NA
DILI Screen (ongoing safety monitoring)	All		NA	AST or ALT >3 x ULN and TBL >2 x ULN
Urinalysis				
Bilirubin	All		NA	NA
Leukocyte esterase	All		NA	NA
Protein	All		NA	>=2+
Glucose	All		NA	NA
Blood	All		NA	NA
Ketones	All		NA	NA
Nitrite	All		NA	NA
pH	All		NA	NA
Specific gravity	All		NA	NA
Urobilinogen	All		NA	NA

NA=Not Applicable; LLN=Lower limit of normal provided by the laboratory; ULN=Upper limit of normal provided by the laboratory, Abs=Absolute, NCI = National Cancer Institute, DILI = Drug Induced Liver Injury

[a] If criteria in both directions are shown for a single parameter, then abnormalities in each direction are summarized separately.

7.3 Vital Signs

Descriptive statistics for vital signs (eg, systolic and diastolic blood pressure, pulse rate, body weight, respiratory rate, and temperature) and their changes from baseline at each post-baseline visit and at the end of study will be presented by treatment group and for ontamalimab all doses.

For pulse rate, a post-baseline value is considered as a PCI value if its meets both criteria for observed value and change from baseline. For systolic/diastolic blood pressure, a post-baseline value is considered as a PCI value if it meets criteria for observed value or change from baseline. For weight and BMI, a post-baseline value is considered as a PCI value if it meets criteria for change from baseline. The PCI criteria are listed in Table 4. The number and percentage of subjects with PCI post-baseline values will be tabulated by each treatment group and ontamalimab all doses treatment group. The percentages will be calculated relative to the number of subjects with at least 1 post-baseline assessment. The numerator is the total number of subjects with at least 1 PCI post-baseline vital sign value. A supportive listing of subjects with post-baseline PCI values will be provided including the subject number, site, baseline, and post-baseline values.

All vital signs data will be listed for the Safety Set.

Table 4 Criteria for Potentially Clinically Important Vital Signs

Vital Sign Parameter	Flag	Criteria[a]	
		Observed Value	Change from Baseline
Systolic blood pressure (mmHg)	High	≥180	Increase of ≥20
	Low	≤90	Decrease of ≥20
Diastolic blood pressure (mmHg)	High	≥105	Increase of ≥15
	Low	≤50	Decrease of ≥15
Pulse rate (beats per minute)	High	≥120	Increase of ≥15
	Low	≤50	Decrease of ≥15
Weight (kg)	High	-	Increase of ≥7%
	Low	-	Decrease of ≥7%
BMI (kg/m^2)	High	-	Increase of ≥10%
	Low	<18	Decrease of ≥10%
Temperature (°C)		NA	NA

NA=Not Applicable

[a] For pulse rate, a post-baseline value is considered as a PCI value if its meets both criteria for observed value and change from baseline. For systolic/diastolic blood pressure, a post-baseline value of is considered as a PCI value if its meets criteria for observed value or change from baseline. For weight and BMI, a post-baseline value is considered as a PCI value if it meets criteria for change from baseline.

7.4 Electrocardiogram (ECG)

A central ECG reader will be used. Descriptive statistics for ECG variables (eg, heart rate, PR interval, QRS interval, QT interval, and QTc interval using both Bazett and Fridericia corrections) and their changes from baseline at each assessment time point will be presented by treatment group. ECG interpretation will be summarized by visit. A shift table from baseline to Week 12 for ECG Interpretation results will be presented by treatment group and for ontamalimab all doses.

ECG variable values will be considered PCI if they meet or exceed the upper limit values listed in Table 5. The number and percentage of subjects with post-baseline PCI values will be tabulated by treatment group and for ontamalimab all doses. The percentages will be calculated relative to the number of subjects with at least 1 post-baseline assessment. The numerator is the total number of subjects with at least 1 post-baseline PCI value.

Listings of ECG data including the central reader's assessment and investigator's interpretation by individual subject will be produced. Separate listings will be produced for subjects with ECG results meeting the PCI criteria. Data from unscheduled visits will be listed, but not summarized.

Table 5 Criteria for Potentially Clinically Important ECG Values

ECG Parameter	Unit	Higher Limit
QRS Interval	msec	≥150
PR Interval	msec	≥250
QTc Interval	msec	≥500

7.5 Other Safety Data

7.5.1 Targeted Neurological Assessment

The targeted neurological evaluation results and neurological consult evaluation results with unexplained abnormal neurological findings will be summarized at screening and at each visit and by treatment group and for ontamalimab all doses. The number and percentage of subjects with targeted neurological exam in each of the neurological domains (vision, motor, tactile sensation, coordination/cerebellar function, speech, verbal comprehension, and cognition/behavior) will be summarized by the result category (abnormal, normal, not done) at each visit and by treatment group and for ontamalimab all doses. The number and percentage of subjects who were referred for a neurological consult and the results (no PML, leukoencephalopathy, no clinically significant finding, other clinically significant finding, not done) will also be summarized by treatment group and for ontamalimab all doses.

7.5.2 Immunogenicity

Presence of Anti-Drug Antibodies (ADAs) will be listed and summarized by visit for each treatment group and for ontamalimab all doses.

ADAs will be classified into pre-existing, treatment-induced responses and treatment-boosted responses. Pre-existing is defined as a signal detected prior to treatment. Treatment-induced responses are defined as a negative pretreatment sample with at least one positive sample at a subsequent time point. Treatment-boosted responses are defined as positive pretreatment samples that are boosted to a higher level following drug administration. Those categories will be listed and summarized for each treatment group and for ontamalimab all doses.

Neutralizing antibodies will be tested on ADA-positive subjects and samples will be defined as Nab-Positive or Negative. Presence of neutralizing antibodies will be listed and summarized for all ADA-positive subjects by visit for each treatment group and for ontamalimab all doses.

ADA prevalence will also be calculated and summarized for each treatment group and for ontamalimab all doses. ADA prevalence is the proportion of study population having drug-reactive antibodies (ADA) at any time point (including pre-existing antibodies) during the study.

ADA incidence will be calculated and summarized for each treatment group and for ontamalimab all doses. ADA incidence is the proportion of study population found to have

developed ADA or boosted their ADA (including pre-existing ADA) at any point during the study period

Listings of positive immunogenicity results and Individual subject immunogenicity data will be presented.

7.5.3 Pregnancy Test and Follicle-stimulating Hormone Test

Pregnancy tests are not required for females of nonchildbearing potential. All pregnancy tests data will be listed for the Safety Set.

7.5.4 Contraception Check

Contraception Check should be performed for female subjects of childbearing potential and male subjects who are with a partner of childbearing potential. Contraception check results will be listed for the Safety Set.

7.5.5 Stool Microbiology

Stool microbiology will be performed at the screening visit (Visit 1) or at any time a subject experience an increase in Gastrointestinal symptoms. Stool microbiology data will be listed for the Safety Set.

7.5.6 Physical Examination

Complete and targeted physical examinations will be performed at the time points specified in Table A1. physical examinations results will be listed for the Safety Set.

10. OTHER ANALYSES

10.1 Coronavirus Pandemic

The Coronavirus (COVID-19) pandemic of 2019-20 particularly poses risks to the safety of subjects enrolled in clinical trials, and the availability and interpretability of data from those trials. COVID-19 impacts on individual subjects collected on COVID-19 CRF pages will be listed for the Randomized Set. Protocol deviations related to COVID-19 will be listed for the Randomized Set.

11. INTERIM ANALYSIS/ DATA MONITORING (REVIEW) COMMITTEE

A data monitoring committee (DMC) was set up to review the safety during the course of the trial. The DMC will not review efficacy data.

No interim analyses are planned for this study.

12. DATA HANDLING CONVENTIONS

12.1 General Data Reporting Conventions

Continuous variables will be summarized using the following descriptive statistics: n, mean, median, standard deviation, minimum, and maximum. For the summary statistics of all numerical variables, unless otherwise specified, minimum and maximum will be displayed to the same level of precision as reported. Mean and median will be displayed to one level of precision greater than the data collected. Standard deviation will be displayed to two levels of precision greater than the data collected.

Categorical and count variables will be summarized by the number of subjects (n) and the percent of subjects in each category. Percentages will be presented with 1 decimal place. When count data are presented, the percentage will be suppressed when the count is zero. A row denoted "Missing" will be included in count tabulations only if there are missing values. The denominator for all percentages will be the number of subjects in that treatment within the population of interest, unless otherwise specified.

P-values will generally be presented to 3 decimal places; values less than 0.001 will be presented as <0.001.

'ONTA' is the acronym of ontamalimab which will be used in output treatment presentation.

12.2 Definition of Visit Windows

Assessments will be assigned to visits based upon the date the assessment took place regardless of the completed CRF page. Assessments will be mapped to visits as outlined in Table 6, Table 7, Table 8, Table 9, and Table 10.

Should there be more than 1 assessment mapped into a given study visit with non-missing results, the assessment closest to the planned visit will be used for analysis (referred to as analysis visit); in case of ties between observations the later assessment will be used.

Study day will be calculated as follows

- If the assessment date is on or after the date of first dose of IP:
  Study day = assessment date – first dosing date + 1
- If the assessment date is before the date of first dose of IP:
  Study day = assessment date – first dosing date

Table 6 Visit Windows (Study Day Based) – PRO, Mayo, [redacted], and Endoscopy Assessments

Visit	Planned Study Day	Start Day of Window	End Day of Window
Baseline	1	ICF Date	<=1
Week 4	28	2	42
Week 8	56	43	70
Week 12	84	>70	EOT Date

*To accommodate the scheduling of the endoscopy prior to the Week 12 visit, the start day of the analysis window will be extended to >61 for this parameter only if there is a valid assessment of stool frequency and rectal bleeding mapped to Week 12.

[redacted]

Table 7 Visit Windows (Study Day Based) – ADA and [redacted] Assessments

Visit	Planned Study Day	Start Day of Window	End Day of Window
Baseline	1	ICF Date	<=1
Week 2	14	2	20
Week 4	28	21	42
Week 8	56	43	70
Week 12	84	>70	EOT Date
Follow-up	EOT Date +112	EOT Date+1	EOF Date

[redacted]

Table 8 Visit Windows (Study Day Based) – [redacted], IBDQ, [redacted], SF-36 Assessments

Visit	Planned Study Day	Start Day of Window	End Day of Window
Baseline	1	ICF Date	<=1
Week 8	56	2	70
Week 12	84	>70	EOT Date

Table 9 Visit Windows (Study Day Based) – Safety Lab and Vital Sign Assessments

Visit	Planned Study Day	Start Day of Window	End Day of Window
Baseline	1	ICF Date	<=1
Week 4	28	2	42
Week 8	56	43	70
Week 12	84	>70	EOT Date
Follow-up	EOT Date +112	EOT Date+1	EOF Date

Table 10 Visit Windows (Study Day Based) – Physical Examination, ECG, Neurological Testing, [redacted] Assessments

Visit	Planned Study Day	Start Day of Window	End Day of Window
Baseline	1	ICF Date	<=1
Week 12	84	2	EOT Date
Follow-up	EOT Date +112	EOT Date+1	EOF Date

*Note, ECG, [redacted] do not have scheduled assessments in Follow-up.

12.3 Derived Efficacy Endpoints

12.3.1 Total Mayo Score

The Mayo score is a measure of UC disease activity. The total Mayo score ranges from 0 to 12 points and consists of the following 4 subscores, each graded from 0 to 3 with higher scores indicating more severe disease:

- Stool frequency (0-3)
- Rectal bleeding (0-3)
- Findings of endoscopy (0-3)
- Physician global assessment (PGA, 0-3).

The calculation of the total score requires stool frequency subscore and rectal bleeding subscore, reported by subjects in the PRO UC daily e-diary and PGA and the centrally read endoscopic subscore.

The Mayo stool frequency and rectal bleeding subscores will be calculated based on each subject's daily e-diary data recorded over the most recent 3 days (consecutive or nonconsecutive) of the last 10 days prior to the scheduled visit start date excluding the following days: day of any bowel preparation, day of endoscopy, any days between day of bowel preparation and day of endoscopy, and the 2 days after the day of endoscopy. The most recent 3 days subscores will be

averaged and rounded to the nearest score according to standard rounding rules. PGA and centrally read endoscopic subscore will be collected at baseline and the Week 12 visit. The total mayo score is the sum of the 4 subscores, and will be calculated at baseline and at the Week 12 visit.

12.3.2 Partial Mayo Score

The partial Mayo score consists of the Mayo score without the endoscopic subscore and ranges from 0 to 9 points.

The calculation of the partial Mayo score requires the stool frequency subscore and the rectal bleeding subscore, reported by subjects in the PRO UC daily e-diary and PGA.

The partial mayo score will be calculated at baseline, and the Week 4, Week 8 and Week 12 visits.

12.3.3 Composite Score

The composite score is a regulatory-authority recommended measure derived from the Mayo score omitting the PGA subscore and ranges from 0 to 9 points.

12.3.4 Mucosal Healing

Mucosal healing is based on the endoscopic and histological assessment at the Week 12 visit. Mucosal healing is defined by the centrally read endoscopic subscore 0 or 1 (modified, excludes friability) and the centrally read Geboes score of ≤2.

Mucosal healing based on endoscopic and histological assessment at the Week 12 visit. Mucosal healing is defined by centrally read endoscopic subscore 0 or 1 (modified, excludes friability) and [redacted], with lamina propria neutrophils, neutrophils in epithelium, and erosion or ulceration scores of 0.

12.3.5 Total Sign/Symptom Score

Patient-reported UC signs and symptom data will be collected using a daily e-diary starting during the screening period. Collection of the daily e-diary data must begin at least 10 days before the baseline visit. Subjects will enter data on UC signs and symptoms items using an electronic handheld device that will be provided to subjects at the start of the study. Compliance is assessed by site staff at each visit. The site staff will instruct the subject on the appropriate use of the e-diary, particularly when compliance is below 80% (eg, <23 out of 28 e-diary entries) when compared with the previous visit.

Subjects will be asked to record the following signs and symptom data, as experienced over the previous 24 hours, in the e diary:

- Stool frequency: Number of Bowel Movements (0-99)
- Rectal bleeding severity and frequency: Rectal Bleeding Worst Experience (0-3) and Number of Bowel Movements with Blood (0-99)

- Diarrhea frequency: Number of Loose Bowel Movements (0-99)
- Urgency frequency: Number of Bowel Movements with Urgency (0-99)
- Abdominal pain worst severity: Worst Abdominal Pain Over the Past 24 Hours (0-10)

Note: In the instrument, if the Number of Bowel Movements is entered as 0, then the rest of questions except Worst Abdominal Pain Over the Past 24 Hours question are skipped as they are further questions about the bowel movements. These skipped items will be considered as 0 for analysis purposes.

Subject's signs and symptom average scores at each scheduled visit will be calculated based on data recorded over the most recent 3 days (consecutive or nonconsecutive) of the last 10 days prior to the scheduled visit start date excluding the following days: day of any bowel preparation, day of endoscopy, any days between day of bowel preparation and day of endoscopy, and the 2 days after the day of endoscopy. Daily signs and symptom records will be assigned to visits based on visit window in Table 6. The assessment closest to the planned visit day will be used for analysis at that visit.

Number of bowel movements and rectal bleeding worst experience will be used to determine the Mayo stool frequency and rectal bleeding subscores, which will be used to calculate the total and partial Mayo scores and the composite score. Subject's signs and symptom average scores of number bowel movements with blood, number bowel movements with urgency, number of bowel movements and number of loose bowel movements will be converted to same scale as shown in Table 11. No conversion will be applied for average scores of worst abdominal pain over the past 24 hours.

Table 11 Proposed Categorized Sign/Symptom Score Conversion Scale

Response	Conversion (applied to each item)
0-2	0
3-5	2.5
6-8	5
9-11	7.5
>=12	10

Total sign/symptom score is the average of the average scores of worst abdominal pain over the past 24 hours and the conversion scale values for number of bowel movements blood, number of bowel movements with urgency, number of bowel movements and number of loose bowel movements, with scale range of 0-10. The categorized and total score scoring systems will be confirmed after database lock according to the UC PRO SAP and are subject to change after the psychometric assessment has been completed.

12.3.6 IBDQ Total and Domain Scores

The IBDQ is a psychometrically validated patient-reported outcome instrument for measuring the disease-specific HRQL in subjects with inflammatory bowel disease, including UC. The IBDQ consists of 32 items (scale: 1-7), which are grouped into 4 domains: bowel function, emotional status, systemic symptoms, and social function. Each domain score is calculated as the sum of the questions scores within the domain as follows:

- Bowel function – sum of scores for questions (1, 5, 9, 13, 17, 20, 22, 24, 26, 29).
- Emotional status - sum of scores for questions (3, 7, 11, 15, 19, 21, 23, 25, 27, 30, 31, 32).
- Systemic symptoms - sum of scores for questions (2, 6, 10, 14, 18).
- Social function - sum of scores for questions (4, 8, 12, 16, 28).

The IBDQ total score is the sum of scores for all 32 questions and ranges from 32 to 224. For the total score and for each domain, a higher score indicates better HRQL. Domain score is missing if one or more item scores are missing. Total score is missing if one or more domain scores are missing.

[redacted]

[redacted]

[redacted]

[redacted]

12.3.10 SF-36

The SF-36 (version 2) consists of 36 items that are aggregated into 8 multi-item scales (physical functioning, role – physical, bodily pain, general health, vitality, social functioning, role – emotional, and mental health), with scores ranging from 0 to 100. Higher scores indicate better HRQL.

These 8 subscales (domains) will be scored using norm-based methods that have standardized the scores to a mean of 50 and a standard deviation of 10 in the general population. The scores range from 0 to 100, with a higher score indicating better quality of life. Physical component

summary and mental component summary scores will be calculated by taking a weighted linear combination of the 8 individual subscales.

The physical component summary of the SF-36 V2.0 consists of these 4 subscales:

- Physical functioning
- Role-physical
- Bodily pain
- General health

The mental component summary of the SF-36 V2.0 consists of these 4 subscales:

- Vitality
- Social functioning
- Role-emotional
- Mental health

12.4 Repeated or Unscheduled Assessments of Safety Parameters

Assessments will be assigned to visits based on the date the assessment took place regardless of the completed CRF page. Assessments will be mapped to visits as outlined in Table 6, Table 7, Table 8, Table 9, and Table 10.

If a subject has more than 1 assessment mapped into a given study visit with non-missing results, the assessment closest to the planned visit will be used for analysis. However, all post-baseline assessments will be used for PCI value determination.

12.5 Handling of Missing, Unused, and Spurious Data

12.5.1 Missing Date of End of Treatment

When the date of the end of treatment is missing for a subject, all efforts should be made to obtain the date from the investigator. If it is still missing after all efforts, then the last dose date +28 days will be used in the calculation of treatment duration.

12.5.2 Missing Date of Ulcerative Colitis Diagnoses

If day of diagnosis date is missing, then the day will be imputed as 1; if both the day and month of diagnosis date are missing, then the day will be imputed as 1 and the month will be imputed as 1 (January).

12.5.3 Missing Date Information for Prior or Concomitant Medications (Therapies/Procedures)

For prior or concomitant medications, including rescue medications, incomplete (ie, partially missing) start date and/or stop date will be imputed. When the start date and the stop date are both incomplete for a subject, impute the start date first.

12.5.3.1 Incomplete Start Date

The following rules will be applied to impute the missing numerical fields. If the stop date is complete and the imputed start date is after the stop date, then the start date will be imputed using the stop date.

MISSING DAY AND MONTH

- If the year of the incomplete start date is the same as the year of the date of the first dose of investigational product, then the day and month of the date of the first dose of IP will be assigned to the missing fields
- If the year of the incomplete start date is before the year of the date of the first dose of investigational product, then December 31 will be assigned to the missing fields
- If the year of the incomplete start date is after the year of the date of the first dose of investigational product, then 01 January will be assigned to the missing fields.

MISSING MONTH ONLY

- The day will be treated as missing and both month and day will be replaced according to the above procedure.

MISSING DAY ONLY

- If the month and year of the incomplete start date are the same as the month and year of the date of the first dose of investigational product, then the day of the date of the first dose of IP will be assigned to the missing day
- If either the year is before the year of the date of the first dose of IP or if both years are the same but the month is before the month of the date of the first dose of investigational product, then the last day of the month will be assigned to the missing day
- If either the year is after the year of the date of the first dose of IP or if both years are the same but the month is after the month of the date of the first dose of investigational product, then the first day of the month will be assigned to the missing day.

12.5.3.2 Incomplete Stop Date

The following rules will be applied to impute the missing numerical fields. If the date of the last dose of IP is missing, then replace it with the last visit date. If the imputed stop date is before the start date (imputed or non-imputed start date), then the imputed stop date will be equal to the start date.

MISSING DAY AND MONTH

- If the year of the incomplete stop date is the same as the year as of the date of the last dose of investigational product, then the day and month of the date of the last dose of IP will be assigned to the missing fields
- If the year of the incomplete stop date is before the year of the date of the last dose of investigational product, then 31 December will be assigned to the missing fields
- If the year of the incomplete stop date is after the year of the date of the last dose of investigational product, then 01 January will be assigned to the missing fields.

MISSING MONTH ONLY

- The day will be treated as missing and both month and day will be replaced according to the above procedure.

MISSING DAY ONLY

- If the month and year of the incomplete stop date are the same as the month and year of the date of the last dose of investigational product, then the day of the date of the last dose of IP will be assigned to the missing day
- If either the year is before the year of the date of the last dose of IP or if both years are the same but the month is before the month of the date of the last dose of investigational product, then the last day of the month will be assigned to the missing day
- If either the year is after the year of the last dose of IP or if both years are the same but the month is after the month of the date of the last dose of investigational product, then the first day of the month will be assigned to the missing day.

12.5.4 Missing Date Information for Adverse Events

For AEs, the default is to only impute incomplete (ie, partially missing) start dates. If start date is missing, no imputation will be performed.

12.5.4.1 Incomplete Start Date

Follow the same rules as in Section 12.5.3.1.

12.5.4.2 Incomplete Stop Date

NA

12.5.5 Missing Severity Assessment for Adverse Events

If severity is missing for an AE starting prior to the date of the first dose of investigational product, then a severity of "Mild" will be assigned. If the severity is missing for an AE starting on or after the date of the first dose of investigational product, then a severity of "Severe" will be assigned. The imputed values for severity assessment will be used for incidence summaries, while the actual values will be used in data listings.

12.5.6 Missing Relationship to IP for Adverse Events

If the relationship to IP is missing for an AE starting on or after the date of the first dose of investigational product, a causality of "Related" will be assigned. The imputed values for relationship to double-blind IP will be used for incidence summaries, while the actual values will be presented in data listings.

12.5.7 Character Values of Clinical Laboratory Variables

If the reported value of a clinical laboratory variable cannot be used in a statistical analysis due to, for example, that a character string is reported for a numerical variable, the appropriately determined coded value will be used in the statistical analysis. However, the actual values as reported in the database will be presented in data listings.

13. ANALYSIS SOFTWARE

Statistical analyses will be performed using Version 9.4 (or newer) of SAS® on a suitably qualified environment.

To score generic SF-36 health survey outcomes instruments, the QualityMetric Health Outcomes™ Scoring Software 5.0 will be used.

14. CHANGES TO ANALYSIS SPECIFIED IN PROTOCOL

The following changes to the analysis specified in Protocol Amendment 2 dated 11 Nov 2019 have been made.

Concomitant medications in the protocol are defined as any medications between first dose of investigational product and the end of safety follow-up. For the purposes of analysis, these medications have been separated into concomitant medication that include medications taken during the treatment period and concomitant post-treatment medications that include medications taken after the treatment period ends.

Due to the early discontinuation of the study before full enrollment and the resulting reduction in sample size, previously planned supplementary and subgroup analyses were removed.

- Removed Subgroup for efficacy endpoints beyond the actual randomization factors.
- Removed supplementary analyses for primary and key secondary efficacy endpoints.
- Removed Per-protocol Set and Completers Set.

Added Coronavirus Pandemic Section.

15. REFERENCES

Bretz, F., Maurer, W., Brannath, W.and Posch, M. 2009. A graphical approach to sequentially rejective multiple test procedures. Stat Med, 28, 586-604.

Feagan, B. G., Rutgeerts, P., Sands, B. E., Hanauer, S., Colombel, J. F., Sandborn, W. J., Van Assche, G., Axler, J., Kim, H. J., Danese, S., Fox, I., Milch, C., Sankoh, S., Wyant, T., Xu, J.and Parikh, A. 2013. Vedolizumab as induction and maintenance therapy for ulcerative colitis. N Engl J Med, 369, 699-710.

Sandborn, W. J., Su, C., Sands, B. E., D'Haens, G. R., Vermeire, S., Schreiber, S., Danese, S., Feagan, B. G., Reinisch, W., Niezychowski, W., Friedman, G., Lawendy, N., Yu, D., Woodworth, D., Mukherjee, A., Zhang, H., Healey, P.and Panes, J. 2017. Tofacitinib as Induction and Maintenance Therapy for Ulcerative Colitis. N Engl J Med, 376, 1723-1736.

Yan, X.and Su, X. 2010. Stratified Wilson and Newcombe Confidence Intervals for Multiple Binomial Proportions. Stat Biopharm Res, 2, 329-35.

APPENDIX 1. SCHEDULE OF ACTIVITIES

Table A1 Schedule of Assessments

Study Procedure	Screening[a]	Baseline	Treatment					Follow-up	
	Weeks -6 to -1	Week 0/ Day 1	Week 2	Week 4	Week 8	Week 12/ET[b]		Week 20[c]	Week 28[c]
Visit Number	**1**	**2**	**3**	**4**	**5**	**6 (Part 1)[d]**	**6 (Part 2)**	**7**	**8**
Study Day	**-42 to 0**	**1**	**14 ±3**	**28 ±3**	**56 ±3**		**84 ±3**	**140 ±7**	**196 ±7**
Informed consent/assent	X								
Eligibility assessment	X	X					X[e]		
Demographics and medical history[f]	X								
Complete physical examination[g]	X						X		X
Targeted physical examination[g]		X							
Targeted neurological assessment[h]	X						X		X
Vital signs[i]	X	X		X	X		X		X
Height	X								
Weight	X	X		X	X		X		X
12-lead ECG[i]	X	X					X		
Chest x-ray[j]	X								
Contraception check[k]	X	X	X	X	X		X		X
Laboratory Assessments									
Hematology	X[l]	X		X	X		X		X
Serum chemistry	X[l]	X		X	X		X		X
Urinalysis	X[l]	X		X	X		X		X
Stool microbiology[m]	X[l]								
HBsAg, HBcAb, HCVAb[n]	X[l]								
HIV testing per local regulation[o]	X								
FSH[p]	X[l]								
Serum β–hCG[q]	X[l]								
Urine β–hCG[q]		X		X	X		X		X
TB test (PPD or QuantiFERON TB Gold Plus)[r]	X[l]								
JCV antibody banked sample[s]		X							
[redacted]		X			X		X		
[redacted]		X[t]			X[t]		X		
[redacted]		X[t]			X[t]		X		
[redacted]		X[t]			X[t]		X		
[redacted]		X[u]	X[u]	X[u]	X[u]		X		

Table A1 Schedule of Assessments

Study Procedure	Screening[a]	Baseline	Treatment					Follow-up	
	Weeks -6 to -1	Week 0/ Day 1	Week 2	Week 4	Week 8	Week 12/ET[b]		Week 20[c]	Week 28[c]
Visit Number	1	2	3	4	5	6 (Part 1)[d]	6 (Part 2)	7	8
Study Day	-42 to 0	1	14 ±3	28 ±3	56 ±3		84 ±3	140 ±7	196 ±7
ADA and NAb sampling		X[t]	X	X[t]	X[t]		X		X
Endoscopic Procedure									
Endoscopy (including biopsy)[v]	X[v]					X			
UC Assessments									
Total Mayo score		X[w]					X[w]		
Partial Mayo score				X	X				
Remission stool frequency and pre-UC stool frequency	X								
PRO-UC daily e-diary data instruction	X								
PRO-UC daily e-diary data[x]	X	X	X	X	X	X	X		
Health Assessment[y]									
IBDQ		X			X		X		
[redacted]		X			X		X		
Hospitalizations, inpatient days, ED visits (HRUA)				X	X		X		X
[redacted]				X	X		X		
[redacted]		X		X	X		X		
[redacted]		X					X		
SF-36, v2, acute form		X			X		X		
[redacted]		X					X		
Treatment Procedures									
Randomization[z]		X							
Administration of ontamalimab or placebo[z]		X		X	X				
Hypersensitivity monitoring[aa]		X	X	X	X	X	X	X	X
Adverse events	X	X	X	X	X	X	X	X	X
Concomitant medications and procedures	X	X	X	X	X	X	X	X	X
Dispense stool collection kit for stool sample[ab]	X			X	X				

ADA=antidrug antibody; β–hCG=beta-human chorionic gonadotropin; [redacted]; ECG=electrocardiogram; ED=emergency department; e-diary=electronic diary; [redacted]; ET=early termination; FSH=follicle-stimulating hormone; HBcAb=hepatitis B core antibody; HBsAg=hepatitis B surface antigen; HBV=hepatitis B virus; HCVAb=hepatitis C virus antibody; HEENT=head, eyes, ears, nose, and throat; HIV=human immunodeficiency virus; HRUA=Healthcare Resource Utilization Assessment; IBDQ=Inflammatory Bowel Disease Questionnaire; IGRA=interferon-gamma release assay; JCV=John Cunningham virus; LTS=long-term safety extension; [redacted]; NAb=neutralizing antibody; PGA=physician global

Table A1 Schedule of Assessments

Study Procedure	Screening[a]	Baseline	Treatment					Follow-up	
	Weeks -6 to -1	Week 0/ Day 1	Week 2	Week 4	Week 8	Week 12/ET[b]		Week 20[c]	Week 28[c]
Visit Number	1	2	3	4	5	6 (Part 1)[d]	6 (Part 2)	7	8
Study Day	-42 to 0	1	14 ±3	28 ±3	56 ±3		84 ±3	140 ±7	196 ±7

assessment; [redacted]; PML=progressive multifocal leukoencephalopathy; PPD=purified protein derivative; PRO=patient-reported outcome; SF-36 v2=Short Form-36 Health Survey, version 2; TB=tuberculosis; [redacted]; UC=ulcerative colitis; [redacted]

[a] At least 2 visits will be necessary to complete the screening procedures, including endoscopy.

[b] Subjects who withdraw early during the treatment period should return for the ET visit and then enter into the safety follow-up period.

[c] Participation in the safety follow-up period is not required if subject is entering the maintenance study (SHP647-303) or LTS (SHP647-304) at the completion of the Week 12 visit. For subjects participating in the 16-week safety follow-up period (not entering the maintenance study or LTS study), the Week 20 (Visit 7) visit will routinely be conducted by telephone; however, as an exception the visit can be performed as a study site visit if preferred. The Week 28 (Visit 8) visit will be at the study site.

[d] Part 1 of Visit 6 must be completed within 10 days (preferably, within 5 to 7 days) before Part 2; this will allow sufficient time for data from the centrally read endoscopy to be available at Part 2 of the visit.

[e] The outcome of Visit 6, Part 2 is used to assess eligibility to enroll in the maintenance (SHP647-303) or LTS (SHP647-304) studies. Eligibility for SHP647-303 and SHP647-304 will be assessed under those respective protocols.

[f] Medical history will include UC history, cardiac history, and smoking history.

[g] Complete physical examination includes the review of the following body systems: general appearance, skin, HEENT, heart, lungs, confrontational visual fields (eyes), breast (optional), abdomen, external genitalia (optional), extremities, neurologic function, back, and lymph nodes. Targeted physical examination includes the review of the following body systems: skin, heart, lungs, confrontational visual fields (eyes), abdomen, and examination of body systems where there are symptom complaints by the subject.

[h] Subjects will be evaluated to reveal any potential abnormalities in the following neurologic domains: vision, motor, tactile sensation, coordination/cerebellar function, speech, verbal comprehension, and cognition/behavior. Subjects with any unexplained positive item at screening that is suggestive of PML should be excluded. See Section 7.2.7.3 of the protocol for further details.

[i] Vital signs (including blood pressure, pulse, respiratory rate, and temperature) and 12-lead ECG should be performed before collection of blood samples for laboratory assessments and before endoscopic procedure.

[j] A chest x-ray performed up to 12 weeks before screening (Visit 1) may be used if available; the official reading must be located in the subject's source documentation.

[k] Contraception check should be performed for female subjects of childbearing potential and male subjects who are with a partner of childbearing potential. See Section 4.4 of the protocol for further details.

[l] Screening laboratory test results, if considered by the investigator to be transient and inconsistent with the subject's clinical condition, may be repeated once during the screening period for confirmation. Results of repeated tests must be reviewed for eligibility prior to the screening endoscopy procedure.

[m] Diagnosis of *Clostridium difficile* infection should be made using the central laboratory. If for any reason the central laboratory is not available, see in Appendix 5 of the protocol for guidance regarding alternate diagnostic algorithms. When a subject experiences an increase in gastrointestinal symptoms, which could be an exacerbation of disease, an infectious etiology must be evaluated including testing for *C. difficile* as described in Appendix 5 of the protocol.

[n] Subjects who test negative for HBsAg but positive for HBcAb without HBV DNA may be considered eligible. For subjects who test positive for HBcAb and negative for HBsAg, a blood sample should be taken for HBV DNA. Blood for HBV DNA reflex testing is collected for required subjects only. If HBV DNA is positive, these subjects will not be eligible.

[o] Documentation of a negative HIV test result within 6 months prior to screening will be accepted.

[p] For confirmation of postmenopausal status in females who have had 12 consecutive months of spontaneous amenorrhea and are ≥51 years of age.

Table A1 Schedule of Assessments

Study Procedure	Screening[a]	Baseline	Treatment					Follow-up	
	Weeks -6 to -1	Week 0/ Day 1	Week 2	Week 4	Week 8	Week 12/ET[b]		Week 20[c]	Week 28[c]
Visit Number	1	2	3	4	5	6 (Part 1)[d]	6 (Part 2)	7	8
Study Day	-42 to 0	1	14 ±3	28 ±3	56 ±3		84 ±3	140 ±7	196 ±7

[q] For females of childbearing potential who are not surgically sterile, do not have confirmed ovarian failure, or do not meet the definition of postmenopausal as described in Section 4.4.1 and Section 7.2.7.3 of the protocol.

[r] A documented negative IGRA or PPD test within 12 weeks before screening (Visit 1) is acceptable provided that an IGRA result or PPD official reading and method is located in the source documentation.

[s] A serum sample will be collected and banked. It may be analyzed if a subject shows neurologic symptoms suggestive of PML.

[t] The sample must be collected before administration of investigational product.

[u] [redacted]

[v] Flexible sigmoidoscopy or colonoscopy (if preferred) at screening must be performed on all subjects after the majority of other eligibility criteria (eg, laboratory values) are met. Endoscopy must be performed during the screening period within 10 days before baseline (Visit 2), preferably within 5 to 7 days before the baseline visit, to obtain the centrally read endoscopic subscore. Biopsy samples will be collected for histological evaluation using the Geboes Score classification for the key secondary efficacy endpoint and [redacted]. Subjects at risk of colorectal cancer, as defined in exclusion criterion 5 (see Section 4.2 of the protocol.), must have a colonoscopy performed at screening, unless the subject has had a surveillance colonoscopy performed within 1 year prior to screening, and any adenomatous polyps found at that examination have been excised. Colonoscopy report and pathology report (if biopsies are obtained) from the colonoscopy performed during screening or in the prior year confirming no evidence of dysplasia and colon cancer must be available in the source documents.

[w] Mayo score will be based on subject daily e-diary entries. The Mayo score at baseline will be calculated based on the stool frequency subscore and rectal bleeding subscore, reported by subjects in the PRO-UC daily e-diary and PGA obtained at baseline and the centrally read endoscopic subscore for the endoscopy performed during the screening period. The Mayo score at the Week 12 visit will be calculated based on the stool frequency subscore and rectal bleeding subscore reported by subjects in the PRO-UC daily e-diary and PGA and the centrally read endoscopic subscore for the endoscopy performed at Week 12, Visit 6 (Part 1). See Section 7.2.2.2 of the protocol for further details.

[x] PRO-UC daily e-diary will be collected using a daily e-diary during the treatment period. Collection of the daily e-diary data must begin at least 10 days before the baseline visit (Visit 2). The subject should be provided with the handheld device to take home on their first visit. See Section 7.2.7.3 of the protocol for further details. Compliance is assessed by site staff at each visit. The site staff will instruct the subject on the appropriate use of the e-diary, particularly when compliance is below 80% (eg, <23 out of 28 e-diary entries) when compared with the previous visit.

[y] All health outcome or patient-reported questionnaires should be completed before completing any other visit assessments.

[z] Interactive response technology will be used for randomization and dispensation of study treatment.

[aa] Beginning at Visit 2, at each visit the subject will be assessed for the presence of Type I and Type III hypersensitivity reactions since the prior visit. If a suspected hypersensitivity reaction has occurred, the next dose of investigational product should be withheld if necessary until the precise etiology (investigational product related or not) has been determined. If a Type III reaction is suspected, appropriate samples for testing will be collected and stored until the adjudication committee determines whether testing is appropriate.

[ab] Stool sample collection kit will be dispensed to the subject to take home at the visit prior to the visit at which testing will be done.

APPENDIX 2. GEOGRAPHIC REGION

Table A2 Geographic Region

Country	Region
Japan	Asia
Korea, Republic of	Asia
Bosnia and Herzegovina	Eastern Europe
Bulgaria	Eastern Europe
Croatia	Eastern Europe
Czech Republic	Eastern Europe
Estonia	Eastern Europe
Hungary	Eastern Europe
Lithuania	Eastern Europe
Poland	Eastern Europe
Romania	Eastern Europe
Russia	Eastern Europe
Serbia	Eastern Europe
Slovakia	Eastern Europe
Ukraine	Eastern Europe
Austria	Western Europe
Belgium	Western Europe
Germany	Western Europe
Greece	Western Europe
Ireland	Western Europe
Italy	Western Europe
Netherlands	Western Europe
Portugal	Western Europe
Spain	Western Europe
Switzerland	Western Europe
United Kingdom	Western Europe
Argentina	ROW
Australia	ROW
Brazil	ROW
Colombia	ROW
Israel	ROW
Lebanon	ROW
Mexico	ROW
New Zealand	ROW
South Africa	ROW
Turkey	ROW
Canada	North America
United States	North America

ROW (Africa/Australia/Latin America/Middle East). Asia (Japan/South Korea)

APPENDIX 3. RESCUE THERAPY FOR UC

- **Biologics with immunomodulatory properties**
  - Any exposure after first dose
- **Non-biologics with immunomodulatory properties**
  - **Immunosuppressants**
    - After first dose and up to and including Week 8: any increase from baseline for more than 5 days
    - After Week 8: Any dose above baseline
  - **5-ASA**
    - After first dose and up to and including Week 8: any increase from baseline for more than 5 days
    - After Week 8: Any dose above baseline
  - **Other small molecule immunomodulatory active agents**
    - Any exposure after first dose
- **Leukocyte apheresis, other apheresis, and plasma exchange**
  - Any exposure after first dose
- **Systemic glucocorticoids**
  - **Systemic glucocorticoids given via oral or rectal routes of administration**
    - After first dose and up to and including Week 8: any increase from baseline for more than 7 days
    - After Week 8: Any dose above baseline
  - **Systemic glucocorticoids given via parenteral routes of administration**
    - Any exposure after first dose
- **Topical Glucocorticoids**
  - **Budesonide**
    - Subjects not taking Budesonide at baseline:
      - After first dose and up to and including Week 8: more than 9 mg/day for one day or any exposure for more than 5 days
      - After Week 8: Any dose above baseline
    - Subjects taking Budesonide at baseline:
      - After first dose and up to and including Week 8: more than 9 mg/day for one day or any increase for more than 5 days
      - After Week 8: Any dose above baseline
  - **Beclomethasone**
    - Subjects not taking Beclomethasone at baseline:
      - After first dose and up to and including Week 8: more than 5 mg/day for one day or any exposure for more than 5 days
      - After Week 8: Any dose above baseline
    - Subjects taking Beclomethasone at baseline:

- After first dose and up to and including Week 8: more than 5 mg/day for one day or any increase for more than 5 days
- After Week 8: Any dose above baseline