STATISTICAL ANALYSIS PLAN

NCT Number: NCT03290781

Study Title: A Phase 3 Randomized, Double-blind, Placebo-controlled, Parallel-group Efficacy and Safety Study of SHP647 as Maintenance Therapy in Subjects With Moderate to Severe Ulcerative Colitis (FIGARO UC 303)

Study Number: SHP647-303

SAP Version and Date:

Version 3: 23 Jun 2021

STATISTICAL ANALYSIS PLAN

SHP647
PHASE 3

A Phase 3 Randomized, Double-blind, Placebo-controlled, Parallel-group Efficacy and Safety Study of SHP647 as Maintenance Therapy in Subjects With Moderate to Severe Ulcerative Colitis (FIGARO UC 303)

PROTOCOL IDENTIFIER: SHP647-303

Study Sponsor(s): Shire Human Genetic Therapies, Inc. ("Shire"), 300 Shire Way, Lexington, MA 02421 USA

Author: [redacted], [redacted], PPD

Protocol: V1 (10 Jul 2017)
Amendment 1 (11 Sep 2018)
Amendment 2 (11 Nov 2019)
Amendment 3 (17 Sep 2020)

SAP Version #: 3

SAP Date: 23 Jun 2021

Status: Final

REVISION HISTORY

Version	Issue Date	Summary of Changes
1.0	22 Oct 2020	Final Document
2.0	20 Nov 2020	Updated text in Section 6.1.1 and Section 6.2.1. Minor editorial and formatting changes.
2.1	13 Jan 2021	Updated Tables 6, 7 and 9 to adjust Study Day start and end dates
3	23 Jun 2021	Updated text in Section 5.3, Section 9, and Section 12.1. Fixed a hyperlink that does not work properly.

TABLE OF CONTENTS

ABBREVIATIONS

ADA	anti-drug antibody
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
CRO	contract research organization
[redacted]	[redacted]
DMC	data monitoring committee
eCRF	electronic case report form
ECG	electrocardiogram
[redacted]	[redacted]
[redacted]	[redacted]
[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
HRUA	Healthcare Resource Utilization Analyses
IBDQ	Inflammatory Bowel Disease Questionnaire
ICF	informed consent form
LLOQ	lower limit of quantification
IP	investigational product
LOCF	last observation carried forward
LTS	long-term safety extension
NAb	neutralizing antibodies
[redacted]	[redacted]
MAR	missing at random
MedDRA	Medical Dictionary for Regulatory Activities
MNT	maintenance study
PCI	potentially clinically important
[redacted]	[redacted]
PDF	probability density function

PGA	physician global assessment
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
PML	progressive multifocal leukoencephalopathy
PRO	patient-reported outcome
Q4W	once every 4 weeks
QoL	quality of life
QTcB	QT interval corrected for heart rate using Bazett's formula
QTcF	QT interval corrected for heart rate using Fridericia's formula
RB	rectal bleeding
REML	restricted maximum likelihood
SAE	serious adverse event
SAP	statistical analysis plan
SC	subcutaneous/subcutaneously
SF	stool frequency
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 and safety, as well as descriptive summaries of [redacted] data and Health Economics and Outcomes Research (HEOR) data, as described in Protocol Amendment 3 dated 17 Sep 2020 (original protocol dated 10 Jul 2017). Specifications for tables, figures, and listings are contained in a separate document. The analysis plans for [redacted] and HEOR patient-reported outcome (PRO) validation, if performed, are prepared separately.

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

2. OBJECTIVES, ESTIMANDS, AND ENDPOINTS

2.1 Objectives

2.1.1 Primary Objective

The primary objective of the study is to evaluate the efficacy of ontamalimab as maintenance treatment of remission, based on composite score of patient-reported symptoms and centrally read endoscopy, in subjects with moderate to severe UC.

2.1.2 Key Secondary Objectives

The key secondary objectives are as follows:

- To evaluate the efficacy of ontamalimab on endoscopic remission, based on centrally read endoscopy.

- To evaluate the efficacy of ontamalimab on clinical remission, based on composite score of patient-reported symptoms.

- To evaluate the efficacy of ontamalimab on maintenance of remission among subjects in remission at baseline of the SHP647-303 study, based on composite score of patient-reported symptoms and centrally read endoscopy.

- To evaluate the efficacy of ontamalimab on clinical response, based on composite score of patient-reported symptoms and centrally read endoscopy.

- To evaluate the efficacy of ontamalimab on mucosal healing, on a centrally read endoscopic and histological assessment using the Geboes Score grading system.

- To evaluate the efficacy of ontamalimab on glucocorticoid-free clinical remission.

- To evaluate the efficacy of ontamalimab on glucocorticoid-free remission.

2.1.3 Other Secondary Objectives

The other secondary objective is:

- To evaluate the effect of ontamalimab maintenance treatment on other clinical and endoscopic outcomes (Mayo-based remission, clinical remission over time, and sustained endoscopic remission).

2.2 Estimands

The primary and key 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 52 in remission.	16- to 80-year-old adolescents and adults with UC defined through inclusion and exclusion criteria as stated in the protocol	Remission at the Week 52 visit without treatment failure or rescue therapy defined in Appendix 16.5 or discontinuation	Composite: intercurrent events captured in variable definition	Difference in proportions of subjects in remission at the Week 52 visit between each active treatment group (25 mg or 75 mg ontamalimab) and the corresponding placebo (25 mg in induction placebo group or 75 mg in induction placebo group; placebo groups are not pooled)
Key Secondary	1st key secondary estimand is the effect of ontamalimab compared to placebo at Week 52 in impact on endoscopic remission.	16- to 80-year-old adolescents and adults with UC defined through inclusion and exclusion criteria as stated in the protocol	Endoscopic remission at the Week 52 visit, defined by centrally read endoscopic subscore 0 or 1 (modified, excludes friability), at the Week 52 visit without treatment failure or rescue therapy defined in Appendix 16.5 or discontinuation	Composite: intercurrent events captured in variable definition	Difference in proportions of subjects in endoscopic remission at the Week 52 visit between each active treatment group (25 mg or 75 mg ontamalimab) and the corresponding placebo (25 mg in induction placebo group or 75 mg in induction placebo group; placebo groups are not pooled)

Estimand	Definition	Attributes			
		A: Population	**B: Variable**	**C: Strategy for Addressing Intercurrent Event**	**D: Population-level Summary**
Key Secondary	2nd key secondary estimand is the effect of ontamalimab compared to placebo at Week 52 in impact on clinical remission.	16- to 80-year-old adolescents and adults with UC defined through inclusion and exclusion criteria as stated in the protocol	Clinical remission at the Week 52 visit, defined by stool frequency subscore of 0 or 1 with at least a 1-point change from induction study (SHP647-301 or SHP647-302) baseline in stool frequency subscore, and rectal bleeding subscore of 0, at the Week 52 visit without treatment failure or rescue therapy defined in Appendix 16.5 or discontinuation	Composite: intercurrent events captured in variable definition	Difference in proportions of subjects in clinical remission at the Week 52 visit between each active treatment group (25 mg or 75 mg ontamalimab) and the corresponding placebo (25 mg in induction placebo group or 75 mg in induction placebo group; placebo groups are not pooled)

Key Secondary	3rd key secondary estimand is the effect of ontamalimab compared to placebo at Week 52 in impact on sustained remission.	16- to 80-year-old adolescents and adults with UC defined through inclusion and exclusion criteria as stated in the protocol	Sustained remission defined as in remission at the SHP647-303 Week 52 visit, without treatment failure or rescue therapy defined in Appendix 16.5 or discontinuation, among subjects who were in remission at the time of baseline in Study SHP647-303. Remission is defined as a composite score of patient-reported symptoms using daily e-diary and centrally read endoscopy, with stool frequency subscore of 0 or 1 with at least a 1-point change from induction study (SHP647-301 or SHP647-302)	Composite: intercurrent events captured in variable definition	Difference in proportions of subjects with sustained remission at the Week 52 visit between each active treatment group (25 mg or 75 mg ontamalimab) and the corresponding placebo (25 mg in induction placebo group or 75 mg in induction placebo group; placebo groups are not pooled)

Estimand	Definition	Attributes			
		A: Population	**B: Variable**	**C: Strategy for Addressing Intercurrent Event**	**D: Population-level Summary**
			baseline, and rectal bleeding subscore of 0, and endoscopic subscore of 0 or 1 (modified, excludes friability).		
Key Secondary	4th key secondary estimand is the effect of ontamalimab compared to placebo at Week 52 in impact on clinical response based on composite score.	16- to 80-year-old adolescents and adults with UC defined through inclusion and exclusion criteria as stated in the protocol	Clinical response at the Week 52 visit, defined as a decrease from induction study (SHP647-301 or SHP647-302) baseline in the composite score of subject-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	Composite: intercurrent events captured in variable definition	Difference in proportions of subjects with clinical response at the Week 52 visit between each active treatment group (25 mg or 75 mg ontamalimab) and the corresponding placebo (25 mg in induction placebo group or 75 mg in induction placebo group; placebo groups are not pooled)

Estimand	Definition	Attributes			
		A: Population	**B: Variable**	**C: Strategy for Addressing Intercurrent Event**	**D: Population-level Summary**
			subscore for rectal bleeding ≥1 point or a subscore for rectal bleeding ≤1 without treatment failure or rescue therapy defined in Appendix 16.5 or discontinuation		
Key Secondary	5th key secondary estimand is the effect of ontamalimab compared to placebo at Week 52 in impact on mucosal healing.	16- to 80-year-old adolescents and adults with UC defined through inclusion and exclusion criteria as stated in the protocol	Mucosal healing at the Week 52 visit, defined as centrally read endoscopic subscore 0 or 1 (modified, excludes friability) and centrally read Geboes score of ≤2 without treatment failure or rescue therapy defined in	Composite: intercurrent events captured in variable definition	Difference in proportions of subjects with mucosal healing at the Week 52 visit between each active treatment group (25 mg or 75 mg ontamalimab) and the corresponding placebo (25 mg in induction placebo group or 75 mg in induction placebo group; placebo groups are not pooled)

Estimand	Definition	Attributes			
		A: Population	**B: Variable**	**C: Strategy for Addressing Intercurrent Event**	**D: Population-level Summary**
			Appendix 16.5 or discontinuation.		
Key Secondary	6th key secondary estimand is the effect of ontamalimab compared to placebo at Week 52 in impact on glucocorticoid-free clinical remission.	16- to 80-year-old adolescents and adults with UC defined through inclusion and exclusion criteria as stated in the protocol	Glucocorticoid-free clinical remission at the Week 52 visit, among subjects using glucocorticoids at induction study baseline, defined as clinical remission in addition to not requiring any treatment with glucocorticoids for at least 4 weeks prior to the Week 52 visit. Clinical remission is defined as stool frequency subscore of 0 or 1 with at least a 1-point change from induction study	Composite: intercurrent events captured in variable definition	Difference in proportions of subjects with glucocorticoid-free clinical remission at the Week 52 visit between each active treatment group (25 mg or 75 mg ontamalimab) and the corresponding placebo (25 mg in induction placebo group or 75 mg in induction placebo group; placebo groups are not pooled)

Estimand	Definition	Attributes			
		A: Population	B: Variable	C: Strategy for Addressing Intercurrent Event	D: Population-level Summary
			(SHP647-301 or SHP647-302) baseline in stool frequency subscore, and rectal bleeding subscore of 0, at the Week 52 visit, without treatment failure or rescue therapy defined in Appendix 16.5 or discontinuation.		

Key Secondary	7th key secondary estimand is the effect of ontamalimab compared to placebo at Week 52 in impact on glucocorticoid-free remission.	16- to 80-year-old adolescents and adults with UC defined through inclusion and exclusion criteria as stated in the protocol	Glucocorticoid-free remission at the Week 52 visit, among subjects using glucocorticoids at induction study baseline, defined as remission in addition to not requiring any treatment with glucocorticoids for at least 4 weeks prior to the Week 52 visit. Remission is defined as a composite score of subject-reported symptoms using daily e-diary and endoscopy, with stool frequency subscore of 0 or 1 with at least a 1-point change from induction study (SHP647-301 or SHP647-302) baseline, and rectal bleeding subscore of 0, and endoscopic	Composite: intercurrent events captured in variable definition	Difference in proportions of subjects with glucocorticoid-free remission at the Week 52 visit between each active treatment group (25 mg or 75 mg ontamalimab) and the corresponding placebo (25 mg in induction placebo group or 75 mg in induction placebo group; placebo groups are not pooled)

Estimand	Definition	Attributes			
		A: Population	B: Variable	C: Strategy for Addressing Intercurrent Event	D: Population-level Summary
			subscore of 0 or 1 (modified, excludes friability), without treatment failure or rescue therapy defined in Appendix 16.5 or discontinuation.		

2.3 Endpoints

2.3.1 Primary Endpoint

The primary efficacy endpoint is remission at the Week 52 visit. Remission is defined as a composite score of patient-reported symptoms using daily e-diary and centrally read endoscopy as follows:

- Stool frequency (SF) subscore of 0 or 1 with at least a 1-point change from induction (SHP647-301 or SHP647-302) baseline

  AND

- Rectal bleeding (RB) 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 52 visit.

- Clinical remission, as defined by SF subscore of 0 or 1 with at least a 1-point change from induction study (SHP647-301 or SHP647-302) baseline in SF subscore, and RB subscore of 0, at the Week 52 visit.

- Sustained remission, ie, in remission at the SHP647-303 Week 52 visit, among subjects who were in remission at the time of baseline in Study SHP647-303. Remission is defined as a composite score of patient-reported symptoms using daily e-diary and centrally read endoscopy, with SF subscore of 0 or 1 with at least a 1-point change from induction study (SHP647-301 or SHP647-302) baseline, and RB subscore of 0, and endoscopic subscore of 0 or 1 (modified, excludes friability).

- Clinical response based on composite score at the Week 52 visit. Clinical response (composite) is defined as a decrease from induction study (SHP647-301 or SHP647-302) baseline in the composite score of subject-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 RB ≥1 point or a subscore for RB ≤1.

- Mucosal healing based on endoscopic and histological assessment at the Week 52 visit. Mucosal healing is defined by centrally read endoscopic subscore 0 or 1 (modified, excludes friability) and centrally read Geboes score of ≤2.

- Glucocorticoid-free clinical remission at Week 52, among subjects using glucocorticoids at induction study (SHP647-301 or SHP647-302) baseline. Glucocorticoid-free clinical remission is defined as clinical remission in addition to not requiring any treatment with glucocorticoids for at least 4 weeks prior to the Week 52 visit. Clinical remission is defined as SF subscore of 0 or 1 with at least a 1-point change from induction study (SHP647-301 or SHP647-302) baseline in SF subscore, and RB subscore of 0, at the Week 52 visit.

- Glucocorticoid-free remission at Week 52, among subjects using glucocorticoids at induction study (SHP647-301 or SHP647-302) baseline. Glucocorticoid-free remission is defined as remission in addition to not requiring any treatment with glucocorticoids for at least 4 weeks prior to the Week 52 visit. Remission is defined as a composite score of subject-reported symptoms using daily e-diary and endoscopy, with SF subscore of 0 or 1 with at least a 1-point change from induction study (SHP647-301 or SHP647-302) baseline, and RB subscore of 0, and endoscopic subscore of 0 or 1 (modified, excludes friability).

2.3.3 Other Secondary Endpoints

- Remission defined as a total Mayo score of ≤2 with no individual subscore (SF, RB, endoscopy [modified, excludes friability], and physician's global assessment [PGA]) exceeding 1, at Week 52.

- Clinical remission over time with both RB and SF subscores of 0.

- Sustained endoscopic remission, ie, in endoscopic remission at the SHP647-303 Week 52 visit among subjects who were in remission at the time of baseline in Study SHP647-303, as defined by a centrally read endoscopic subscore of 0 or

1 (modified, excludes friability).

2.3.4 [redacted]

[redacted]

[redacted]

3. STUDY DESIGN

3.1 General Description

This is a Phase 3, randomized, double-blind, placebo-controlled, parallel-group study in subjects with moderate to severe UC who completed their participation in an induction study (either SHP647-301 or SHP647-302) and achieved a clinical response.

Clinical response is defined as:

1) A decrease from induction study (SHP647-301 or SHP647-302) baseline in the composite score of patient-reported symptoms using daily electronic diary (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 induction study (SHP647-301 or SHP647-302) baseline in total Mayo score of at least 3 points and at least 30%, with an accompanying decrease in the rectal bleeding subscore of at least 1 point or an absolute rectal bleeding subscore of 0 or 1.

This study consists of a 52-week, double-blind treatment period, followed by a 12-week safety follow-up period for subjects who either discontinue treatment early or who complete the treatment period and do not enter the long-term safety extension (LTS) study (SHP647-304).

Approximately 696 subjects will be enrolled into the study. Approximately 592 subjects from active induction treatments and approximately 104 subjects from placebo induction groups. The eligibility of a subject for the study will be assessed based on the study data collected at the Week 12 visit of the induction studies, which will be considered as the baseline visit for this maintenance study.

Eligible subjects who received active treatment in one of the induction studies and achieved a clinical response will be randomly assigned as follows: subjects who received 25 mg ontamalimab in one of the induction studies will be randomized (1:1) to receive either 25 mg ontamalimab or placebo, and subjects who received 75 mg ontamalimab in one of the induction studies will be randomized (1:1) to receive either 75 mg ontamalimab or placebo.

Eligible subjects who received placebo in one of the induction studies and achieved a clinical response will be randomly assigned in a 2:2:1 ratio to receive 1 of 3 treatments (25 mg ontamalimab, 75 mg ontamalimab, or placebo, respectively) during this maintenance study.

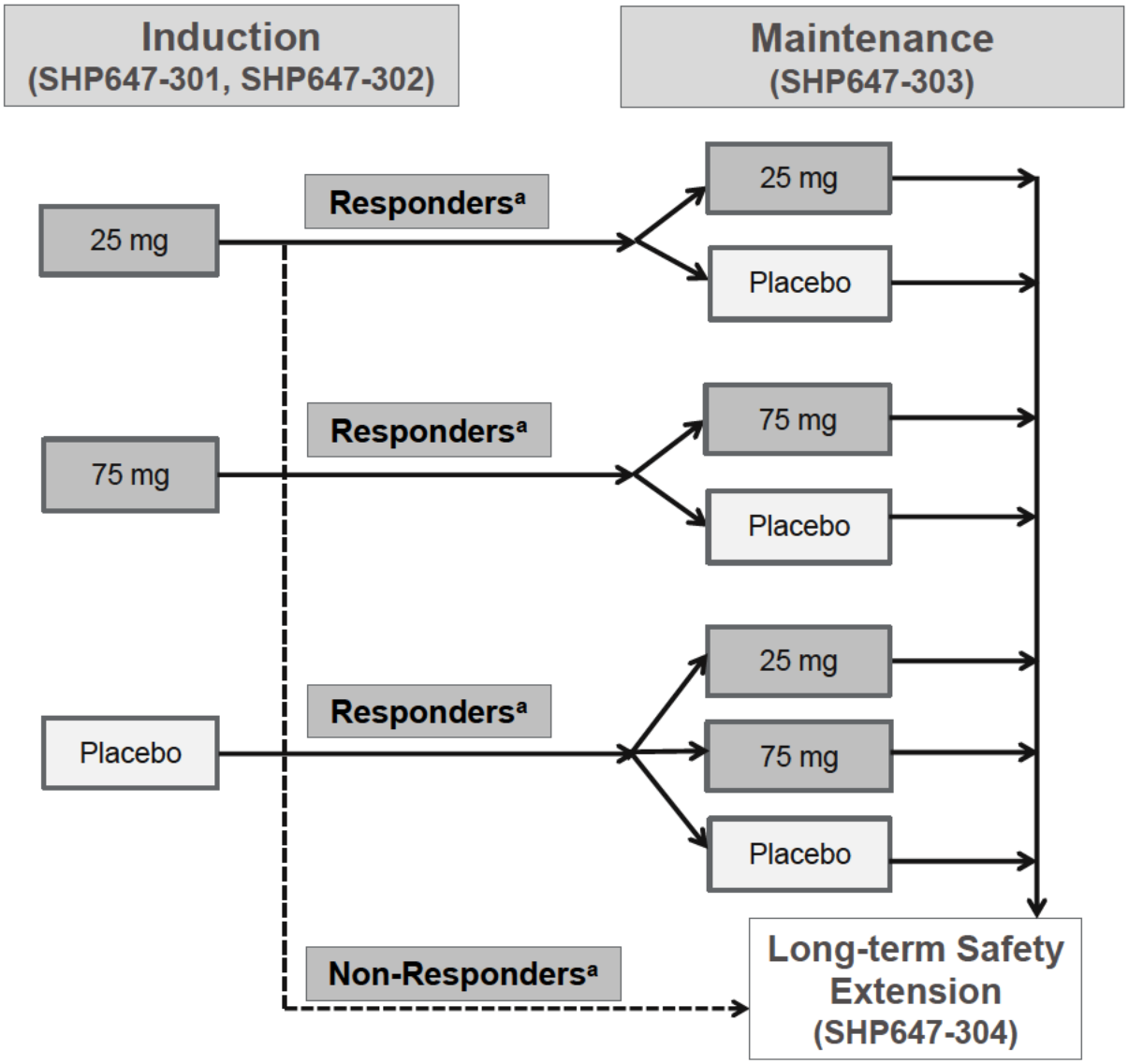
Subjects will be stratified according to glucocorticoid use at SHP647-303 baseline, the subject's status of prior anti-TNF treatment (naïve or experienced), and the degree of clinical response in the induction study (whether remission is achieved or not).

Subjects enrolled in this study (SHP647-303) will receive double-blind maintenance treatment in the form of SC injections, using a prefilled syringe (PFS), every 4 weeks for 52 weeks. Subjects will undergo efficacy, [redacted] and safety assessments as detailed in Section 16.2.

Under Protocol Amendment 3, subjects who complete the double-blind treatment period or subjects who are withdrawn from the study prior to completing the double-blind treatment period due to early closure of the study by the sponsor may be eligible to enter the LTS study (SHP647-304) provided they meet the eligibility criteria under SHP647-304 Amendment 4. Offering treatment in the LTS study after exiting this maintenance study allows subjects who benefited from active treatment, in this study or the induction study, to potentially benefit from continued treatment at the same dose of ontamalimab or at different dose of ontamalimab if, during the course of the SHP647-304 study, one of those doses (25 mg or 75 mg) is determined to be more efficacious based on emergent data from the induction studies. However, if there is no evidence of efficacy of either of the doses in comparison to placebo in the UC/CD clinical study, the entire program may be stopped, including the LTS study. Subjects who are not entering Study SHP647-304 will enter a 12-week safety follow-up period.

An overview of the ontamalimab Phase 3 UC studies is shown in Figure 1, and the overall study design is shown in Figure 2.

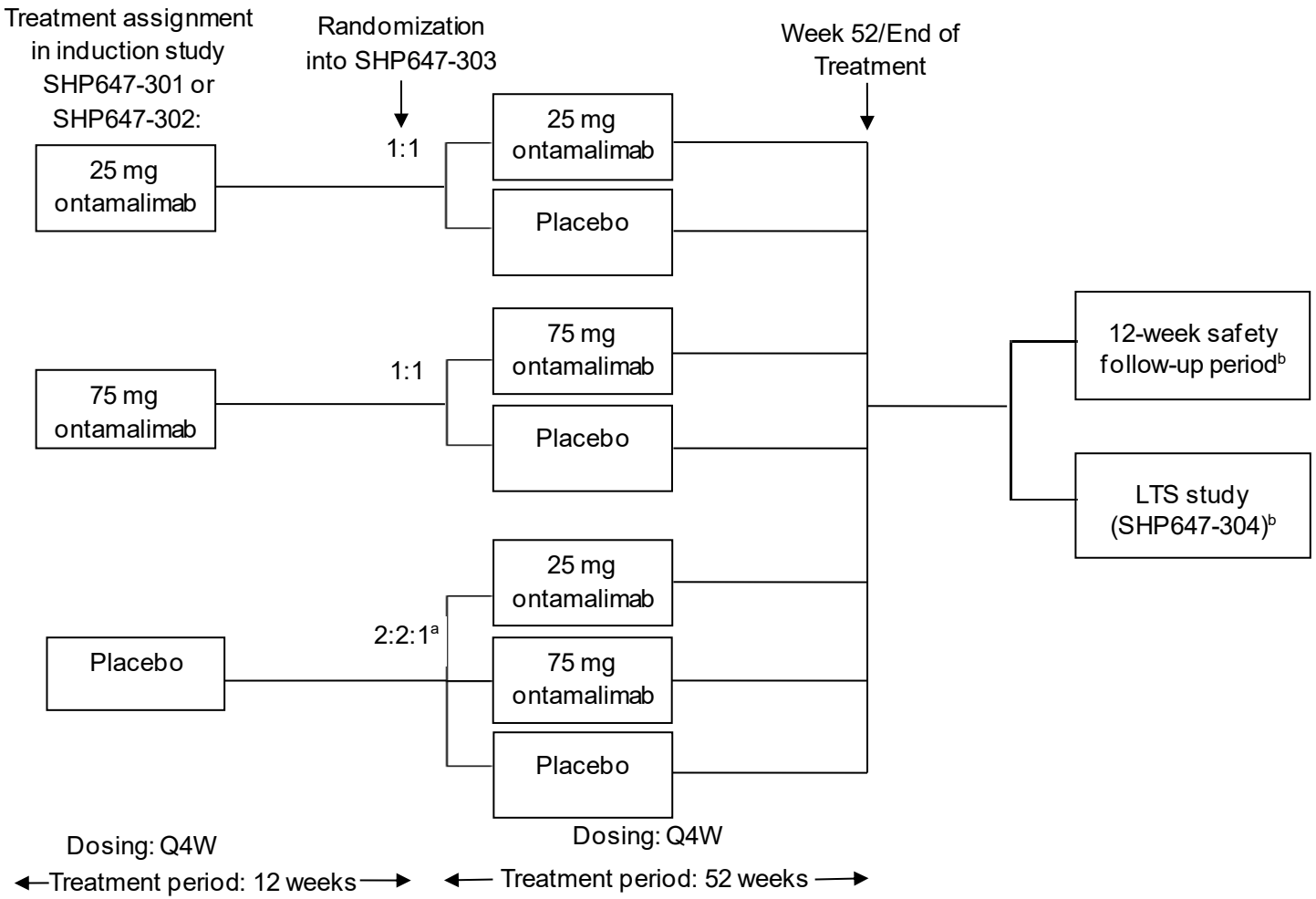
Figure 1: Overview of Ontamalimab Phase 3 Studies in Ulcerative Colitis



[a] Clinical response in the induction studies (SHP647-301 and SHP467-302) 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 in total Mayo score of at least 3 points and at least 30%, with an accompanying decrease in the rectal bleeding subscore of at least 1 point or an absolute rectal bleeding subscore of 0 or 1.

Figure 2 Study Design Flow Chart



LTS=long-term safety extension; Q4W=every 4 weeks.

[a] Eligible subjects who received placebo in one of the induction studies and achieved a clinical response will be randomly assigned in a 2:2:1 ratio to receive one of 3 treatments (25 mg SHP647, 75 mg SHP647, or placebo, respectively).

[b] With the implementation of Protocol Amendment 3, subjects who complete the double-blind treatment period, or subjects who are withdrawn from the study prior to completing the double-blind treatment period due to early closure of the study by the sponsor, may be eligible to enter the LTS study (SHP647-304) provided they meet the eligibility criteria under SHP647-304 Amendment 4. Subjects who are not entering Study SHP647-304 will enter a 12-week safety follow-up period.

3.2 Randomization

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

The composite score and total Mayo score will be calculated at baseline (Week 12/Visit 6 of induction study SHP647-301 or SHP647-302 will be used for Week 0/Day 1/Visit 1 of Study SHP647-303) before randomization. Endoscopic subscore based on the central reader's assessment will be used to determine eligibility.

Subjects who are qualified as a clinical responder, who fulfill all other eligibility criteria, and who received active treatment in the induction study (SHP647-301 or SHP647-302) will be randomized via a computer-generated randomization schedule as follows: subjects who received 25 mg ontamalimab in one of the induction studies will be randomized (1:1) to receive either 25 mg ontamalimab or placebo, and subjects who received 75 mg ontamalimab in one of the induction studies will be randomized (1:1) to receive either 75 mg ontamalimab or placebo.

Subjects who are qualified as a clinical responder, who fulfill all other eligibility criteria, and who received placebo in the induction study will be randomly assigned in a 2:2:1 ratio to receive one of 3 treatments (25 mg ontamalimab, 75 mg ontamalimab, or placebo, respectively) during this maintenance study.

Subjects will be stratified according to glucocorticoid use at SHP647-303 (maintenance; MNT) baseline defined in Appendix 16.4, the subject's status of prior anti-TNF treatment (naïve or experienced), and the degree of clinical response in the induction study (whether remission is achieved or not).

The randomization number represents a unique number corresponding to investigational product (IP) allocated to the subject, once eligibility has been determined. Individual subject treatment is 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 IP packing identification number may not be assigned to more than one 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 IP, 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.

Due to early termination of the ontamalimab program, the sponsor is providing an option for subjects who had responded to active treatment (in this maintenance study or in the induction study) to continue to receive ontamalimab in the long-term safety extension study SHP647-304. As this eligibility criterion into SHP647-304 depends on the blinded treatment assignment in this study, for these subjects, there is a potential for the treatment assignment in this study to be unblinded at the early termination visit when assessing whether a subject can be a rollover into the SHP647-304 study or proceed to the safety follow-up period. The date of the early termination visit will be recorded.

In addition, due to the early discontinuation of the SHP647 program, the SHP647-304 study is planned to unblinded prior to the database lock in this study, which would unblind the treatment assignment in this study for a significant portion of subjects. As such, this study will be considered unblinded when the SHP647-304 study is unblinded and the date of study unblinding will be recorded.

3.4 Sample Size and Power Considerations

The planned sample size for this maintenance study depends on enrollment from the induction studies (SHP647-301 and SHP647-302). Assuming that 50% of subjects receiving ontamalimab induction treatments and 35% of subjects receiving placebo induction treatment in Studies SHP647-301 and SHP647-302 will have a clinical response at Week 12, an estimated 696 subjects will be eligible to enter this maintenance study: 296 subjects from ontamalimab 25 mg induction treatment, 296 subjects from ontamalimab 75 mg induction treatment, and 104 subjects from placebo induction treatment. Expected clinical response rates at Week 12 are based on observed rates from the A7281009 study.

This study is designed to have approximately 85% power for the primary endpoint to detect an individual pairwise treatment difference at a highly statistically persuasive level (ie, a p-value $\leq$.001). However, an alpha of .05 (2-sided) will be used as the overall Type-I error rate for the purposes of demonstration of efficacy in this study and declaring study success. The power will be reported for both Type-I error thresholds (.05 and .001).

Graphical methods are used to control the global family-wise Type-I error rate (FWER) at the 0.05 level (2-sided) for the comparisons of the 2 ontamalimab treatment groups with the respective placebo group based on induction ontamalimab dose. Alpha is split equally at the .025 level (2-sided) for each of the pairwise treatment comparisons. Therefore, the power analysis and sample size estimation were calculated based on the chi-square test of proportions using nQuery Advisor Version 7.0 for an individual ontamalimab dose compared to placebo.

Power calculations are based on assuming a .025 (2-sided) significance level for each pairwise treatment comparison, 296 subjects previously treated with 25 mg ontamalimab in induction (1:1 allocation ratio: 148 subjects in 25 mg ontamalimab treatment group vs. 148 subjects in the placebo group) and 296 subjects treated with 75 mg ontamalimab in induction (1:1 allocation ratio: 148 subjects in the 75 mg ontamalimab treatment group vs. 148 subjects in the placebo group) would yield an approximately 98% power (85% for alpha = .001) to detect individual pairwise treatment difference in the primary

efficacy endpoint, remission at Week 52, of 23% (15% placebo versus 38% ontamalimab). Expected remission rates at Week 52 were based on observed rates from an analysis of Study A7281010, which was ongoing at the time of the original protocol 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 52 are imputed as failures and the above rates account for these subjects.

With the planned sample size of 296 subjects previously treated with 25 mg in induction and 296 subjects previously treated with 75 mg in induction, Table 2 provides the power for detecting a treatment difference between an ontamalimab treatment group and the placebo group for the key secondary endpoints.

With the early discontinuation of the study, the planned sample size of 696 subjects will not be attained as the final number of subjects enrolled into this study is 366. Formal statistical testing of the primary and key secondary endpoints will still be conducted despite not reaching the planned sample size and power levels.

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

Key Secondary Endpoint at Week 12	SHP647 Assumption	Placebo Assumption	Power (alpha = .05)	Power (alpha = .001)
Endoscopic remission	42.5%	15%	0.99	0.97
Clinical remission	60%	32%	0.99	0.92
Maintenance of remission[a]	65%	25%	0.96	0.68
Clinical response by composite score	50%	25%	0.98	0.84
Mucosal healing	30%	13%	0.91	0.53
Glucocorticoid-free clinical remission[b]	50%	19%	0.98	0.79
Glucocorticoid-free remission[b]	45%	14%	0.99	0.84

[a] Based on an anticipated 32% of subjects who will be in remission at baseline of this study.

[b] Based on an anticipated 58% of subjects who will be on corticosteroids at baseline of this study.

4. STATISTICAL ANALYSIS SETS

4.1 Screened Set

The Screened Set will consist of all subjects who have signed an informed consent document for the SHP647-303 study, regardless of treatment received during the induction studies (SHP647-301 and SHP647-302).

4.2 Randomized Set

The Randomized Set will consist of all subjects in the screened set for whom a SHP647-303 randomization number has been assigned, regardless of treatment received during the induction studies (SHP647-301 and SHP647-302). Analyses will be performed according to the randomized treatment regimen regardless of the treatment regimen actually received.

4.3 Safety Set

The Safety Set will consist of all subjects who have received at least 1 dose of IP in the SHP647-303 study, regardless of treatment received during the induction studies (SHP647-301 and SHP647-302). Analyses will be performed according to the treatment regimen actually received regardless of the randomized treatment regimen.

4.4 Ontamalimab Responder Full Analysis Set

The Ontamalimab Responder Full Analysis Set (FAS) will consist of all subjects in the randomized set who receive at least 1 dose of IP in the SHP647-303 study and who were previously treated with ontamalimab in the induction studies. Analyses will be performed according to the randomized treatment regimen regardless of the treatment regimen actually received.

4.5 Placebo Responder Full Analysis Set

The Placebo Responder FAS will consist of all subjects in the randomized set who receive at least 1 dose of IP in the SHP647-303 study who were previously treated with placebo in the induction studies. Analyses will be performed according to the randomized treatment regimen regardless of the treatment regimen actually received.

4.6 [redacted]

[redacted]

[redacted]

5. STUDY SUBJECTS

Outputs will be presented by treatment groups as listed in Appendix 16.6.

5.1 Disposition of Subjects

The number of screened subjects will be presented in the overall column. The number of subjects included in each analysis set (ie, Randomized, Safety, Ontamalimab Responder FAS, Placebo Responder FAS, and [redacted]) will be summarized by treatment groups as listed in Appendix 16.6. The percentage, based on the number of subjects in the Safety set, will be presented for Ontamalimab Responder FAS, Placebo Responder FAS, and [redacted] subject sets. The study analysis set classifications of each subject will be listed for the Screened Set.

The number and percentage of subjects who completed and prematurely discontinued during the treatment and follow-up periods and 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 as listed in Appendix 16.6 for the Safety Set. The number and percentage of subjects who continued to the follow-up period and who continued to SHP647-304 study will be presented by treatment group as listed in Appendix 16.6 for the Safety Set.

The number and percentage of subjects who completed and prematurely discontinued the study will be presented by treatment group as listed in Appendix 16.6 for the Safety Set. Subjects who complete 52 weeks of treatment and roll over to the SHP647-304 study or enter and complete the safety follow-up period 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 periods. 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 52 weeks of treatment and discontinued from the 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 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 treatment group as listed in Appendix 16.6 for the Safety Set, Ontamalimab Responder FAS, and Placebo Responder FAS.

Subject's age is from the induction study (SHP647-301 or SHP647-302), which is calculated as the difference between the date of birth and the date of informed consent in the induction study. 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 to <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), 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 in the induction study (SHP647-301 or SHP647-302) unless otherwise specified. Maintenance (MNT; SHP647-303) baseline is defined as the value for the assessment collected at the Week 12 visit of the induction study. The following baseline characteristics will be summarized:

- Weight,
- Height,
- Body mass index (BMI),
- UC disease duration and UC disease duration category (<1 year, ≥1 to <3 years, ≥3 to <7 years and ≥7 years)

Disease duration is calculated as the number of years from the date of UC diagnosis to the date of informed consent in the induction study (SHP647-301 or SHP647-302),

- UC disease location (Proctitis, Procto-Sigmoidits, Left-sided Colitis, Extensive Colitis/Pancolitis),
- Total Mayo severity and total Mayo severity category (<6, 6 to <9, and ≥9),
- Total Mayo severity and total Mayo severity category at MNT baseline (<6, 6 to <9 and ≥9),
- Stool frequency score (0, 1, 2, and 3),
- Stool frequency score at MNT baseline (0, 1, 2, and 3),
- Rectal bleeding score (0, 1, 2, and 3),
- Rectal bleeding score at MNT baseline (0, 1, and 2),
- Findings of endoscopy (0, 1, 2, and 3),
- Findings of endoscopy at MNT baseline (0, 1, 2, and 3),
- Physician global assessment (0, 1, 2, and 3),
- Physician global assessment at MNT baseline (0, 1, 2, and 3),
- [redacted]
- [redacted]

The following UC medication history/use will be summarized:

- Immunosuppressant Experienced (Yes, No),
- Anti-TNF Failure (Yes, No),
- Anti-TNF Failure Times (Anti-TNF Naïve, Anti-TNF Experienced without Failure, Failed 1 anti-TNF therapy, Failed 2 anti-TNF therapies, Failed 3 or more anti-TNF therapies),
- Anti-TNF Experienced (Experienced vs. Naïve) (both randomized status and actual status),
- 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(Yes, No) (actual status),

- Glucocorticoid Use at MNT Baseline (Yes, No) (both randomized status and actual status),
- Immunosuppressant Use at MNT baseline (Yes, No),
- Glucocorticoid Use at MNT baseline AND Immunosuppressant Use at MNT baseline (Both Glucocorticoid and Immunosuppressant Use, Only Glucocorticoid Use, Only Immunosuppressant Use, Neither Glucocorticoid nor Immunosuppressant Use),
- Glucocorticoid Use at MNT baseline (Systemic or Topical, Systemic only, Topical only, None),
- Systemic Glucocorticoid Dose at MNT Baseline,
- Systemic Glucocorticoid Dose at MNT Baseline Category (≤10, >10),
- 5-ASA Use at MNT baseline (Yes, No).

The following outcomes from the induction study will be summarized:

- The actual treatment received in the induction study (Placebo, Ontamalimab 25 mg, Ontamalimab 75 mg),
- Clinical Response (Yes, No),
- Remission at MNT baseline (both randomized and actual status) (Yes, No),
- Endoscopic remission at MNT baseline (Yes, No),
- Clinical remission at MNT baseline (Yes, No),
- Clinical response by composite score at Week 12 (Yes, No),
- Mucosal healing at MNT baseline (Yes, No),
- Clinical response by total Mayo score at MNT baseline (Yes, No).

Smoking history will be recorded in the eCRF at the Screening Visit (Visit 1) in the induction study (SHP647-301 or SHP647-302) and will be summarized by treatment group as listed in Appendix 16.6. Smoking history will be listed for the Safety Set. Duration of smoking will be calculated as (substance use end date – substance use start date + 1). If substance use end date is missing, then it will be imputed as the randomization date.

5.3 Medical History

Medical history will be collected at the Screening Visit (Visit 1) in the induction study (SHP647-301 or SHP647-302) and baseline visit in the SHP647-303 study, and will be listed for the Safety Set.

Medical history will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) Version 19.1 2016 or newer. The induction study medical history will be summarized by the number and percentage of subjects for each treatment group as listed in Appendix 16.6, system organ class (SOC), and preferred term.

Cardiovascular history will be collected at the Screening Visit (Visit 1) in the induction study (SHP647-301 or SHP647-302) and will be summarized by treatment group as listed in Appendix 16.6 for the Safety Set. Cardiovascular history will be listed for the Safety Set.

UC history will be collected at the Screening Visit (Visit 1) in the induction study (SHP647-301 or SHP647-302) and will be listed for the Safety Set.

5.4 Prior Medications

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

Prior medication is defined as any medication with start date prior to the date of the first dose of IP in the SHP647-303 study, and which is ongoing at the time of the baseline visit in the SHP647-303 study . Incomplete medication dates will be imputed as described in Section 12.6.3.

All prior medications will be listed for the Safety Set.

5.5 Concomitant Therapies, Procedures and Medications

Concomitant medications will be coded using the WHO Drug Dictionary dated 01 Dec 2016.

Concomitant medication is defined as any medication with a start date prior to the dates of the first dose of IP in the SHP647-303 study and continuing after the first dose of IP in the SHP647-303 study or with a start date between the dates of the first dose of IP in the SHP647-303 study and end of treatment date, inclusive. Medication that starts after the first dose of SHP647-304 IP will be collected in the SHP647-304 database and will not

be considered as concomitant medication in SHP647-303. Incomplete medication dates will be imputed as described in Section 12.6.3. Any medication with a start date between the dates of the first dose of IP and EOT date in SHP647-303, inclusive, or with a start date after the EOT date (post-treatment) in SHP647-303 will be considered a post-treatment concomitant medication.

Concomitant medication usage will be summarized by the number and percentage of subjects by treatment group as listed in Appendix 16.6, 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 concomitant medications, medical/surgical procedures, and therapies will be listed for the Safety Set.

5.6 Exposure to Investigational Product

Investigational product (ontamalimab or placebo) will be administered SC every 4 weeks (from Week 0 to Week 48). Exposure to IP in the SHP647-303 study will be summarized by presenting the number of subjects who had 1 injection, 2 injections, 3 injections, etc. Number of injections received will be summarized by treatment groups as listed in Appendix 16.6. The administration records by visit will be listed for the Safety Set.

Exposure to IP in the SHP647-303 study 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 in the SHP647-303 study to the date of the last dose of IP taken in the SHP647-303 study + 29 days. Subject years of exposure is calculated as (Date of last dose of IP in the SHP647-303 study – date of first dose of IP in the SHP647-303 study + 29)/365.25. Total subject years of exposure is calculated by summing the subject years of exposure for all subjects within each column.

Descriptive statistics (n, mean, standard deviation, minimum, median, and maximum) will be presented to describe the exposure to IP in the SHP647-303 study by treatment group as listed in Appendix 16.6.

5.7 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 in the SHP647-303 study until the end of

treatment in the SHP647-303 study divided by the number of injections expected to be taken during that time period, times 100. Percent compliance will be summarized by treatment group as listed in Appendix 16.6. Percent compliance will be listed for the Safety Set.

5.8 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, site, and treatment group as listed in Appendix 16.6, for the Randomized Set. Significant and nonsignificant protocol deviations will be listed for the Randomized Set. The protocol deviations related to COVID-19 will be listed separately for the Randomized Set.

6. EFFICACY ANALYSES

All efficacy analyses will be based on Ontamalimab Responder 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) in the induction study (SHP647-301 or SHP647-302) unless otherwise specified. Maintenance baseline for all efficacy analyses is defined as the value for the efficacy assessment collected at the Week 12 visit of the induction study.

The collected data shown in Appendix 16.1 but removed from Appendix 16.2 will be listed for the FAS including Inflammatory Bowel Disease Questionnaire questions, Short Form-36 Health Survey form, hospitalizations, [redacted]

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.

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.

Due to the early discontinuation of the study before full enrollment and the limited sample size, planned efficacy analyses have been updated.

6.1 Analyses of Primary Efficacy Endpoint

The primary efficacy endpoint is remission at the Week 52 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 induction (SHP647-301 or SHP647-302) baseline

AND

- Rectal bleeding subscore of 0

AND

- Endoscopic subscore of 0 or 1 (modified, excludes friability), based on centrally read results.

The primary efficacy endpoint will be compared for each active treatment group (25 mg or 75 mg ontamalimab) to the corresponding placebo group (25 mg in induction placebo group or 75 mg in induction placebo group; placebo groups are not pooled) using a Cochran-Mantel Haenszel (CMH) chi-square test stratified by actual status of glucocorticoid use at MNT (SHP647-303) baseline, prior anti-TNF treatment, and the degree of clinical response in the induction studies (whether remission is achieved or not). Subjects with the intercurrent events defined in Section 2.2 and/or with missing remission data at Week 52 will be considered failures and counted as nonresponders. Subjects with treatment failure will also be counted as nonresponders. Subjects with missing induction baseline value in SF subscore will be counted as remission if both SF and RB subscores are 0 and endoscopic subscore are either 0 or 1 at Week 52. 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 the corresponding placebo group comparison.

6.1.1 Sensitivity Analyses of Primary Efficacy Endpoint

A sensitivity analysis will be performed using the CMH testing described in Section 6.1, including only the subset of subjects who had the opportunity to complete 52 weeks of treatment in the SHP647-303 study prior to the implementation of Protocol Amendment 3 at their site.

The primary efficacy endpoint will be summarized separately for the placebo responder FAS by treatment group without inferential methods. The number and percentage of subjects, and the estimate of the common treatment difference along with the corresponding unstratified Newcombe 95% CI, will be summarized by treatment group at Week 52. Subjects with missing data at Week 52 will be considered failures and counted as nonresponders.

In addition, 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).

6.1.2 Supplementary Analyses of Primary Efficacy Endpoint

No supplementary analyses of the primary efficacy endpoint are planned for this study.

6.2 Analyses of Key Secondary Efficacy Endpoints

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 52 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 corresponding placebo group (25 mg in induction placebo group or 75 mg in induction placebo group; placebo groups are not pooled) comparison.

The key secondary efficacy endpoints are:

- Endoscopic remission, as defined by centrally read endoscopic subscore 0 or 1 (modified, excludes friability), at the Week 52 visit.

- Clinical remission as defined by stool frequency subscore of 0 or 1 with at least a 1-point change from induction study (SHP647-301 or SHP647-302) baseline in SF subscore, and RB subscore of 0, at the Week 52 visit. Subjects with missing induction baseline value in SF subscore will be counted as clinical remission if both SF and RB subscores are 0 at Week 52.

- Sustained remission, ie, in remission at the SHP647-303 Week 52 visit, among subjects who were in remission at the time of MNT baseline (the observed value for the efficacy assessment at Week 12 in induction study). Remission is defined in Section 6.1.

- Clinical response based on composite score at the Week 52 visit. Clinical response (composite) is defined as a decrease from induction study (SHP647-301 or SHP647-302) baseline in the composite score of subject-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. Subjects with missing induction baseline value in the composite score but in remission by composite score will be counted as clinical response.

- Mucosal healing based on endoscopic and histologic assessment, at the Week 52 visit. Mucosal healing is defined by centrally read endoscopic subscore 0 or 1 (modified, excludes friability) and centrally read Geboes score of ≤2.

- Glucocorticoid-free clinical remission at Week 52, among subjects using glucocorticoids at the MNT baseline. Glucocorticoid-free clinical remission is defined as clinical remission in addition to not requiring any treatment with glucocorticoids for at least 4 weeks prior to the Week 52 visit. Clinical remission is defined in second bullet above.

- Glucocorticoid-free remission at Week 52, among subjects using

glucocorticoids at the MNT baseline. Glucocorticoid-free remission is defined as remission in addition to not requiring any treatment with glucocorticoids for at least 4 weeks prior to the Week 52 visit. Remission is defined in Section 6.1.

6.2.1 Sensitivity Analyses of Key Secondary Efficacy Endpoints

A sensitivity analysis will be performed using the CMH testing described in Section 6.1, including only the subset of subjects who had the opportunity to complete 52 weeks of treatment in the SHP647-303 study prior to the implementation of Protocol Amendment 3 at their site.

Additionally, the key secondary endpoints will be summarized separately for the Placebo Responder FAS by treatment group without inferential methods.

In addition, 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).

6.2.2 Supplementary Analyses of Key Secondary Efficacy Endpoints

No supplementary analyses of the key secondary efficacy endpoints are planned for this study.

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 for the Ontamalimab Responder FAS. The other secondary endpoints will be analyzed using the same approach as described for the primary endpoint. Subjects with missing other secondary endpoint data at the Week 52 visit will be considered failures and counted as nonresponders. These endpoints also will be summarized for the Placebo Responder FAS by treatment group without inferential methods.

The other secondary endpoints and analyses are noted below.

- Remission, defined as a total Mayo score ≤2 with no individual subscore (SF, RB, endoscopy [modified, excludes friability], and physician's global assessment) exceeding 1, at Week 52. 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 (25 mg in induction placebo group or 75 mg in induction placebo group; placebo groups are not pooled) at Week 52.

- Clinical remission over time with both RB and SF 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 (25 mg in induction placebo group or 75 mg in induction placebo group; placebo groups are not pooled) over time.

- Sustained endoscopic remission, ie, in endoscopic remission at the SHP647-303 Week 52 visit among subjects who were in remission at the time of MNT baseline, as defined by a centrally read endoscopic subscore of 0 or 1 (modified, excludes friability). A CMH chi-square test that is described in Section 6.3 will be used to compare the proportion of subjects with sustained endoscopic remission for each active treatment group to placebo group (25 mg in induction placebo group or 75 mg in induction placebo group; placebo groups are not pooled) at Week 52.

6.3.1 Sensitivity Analyses of Other Secondary Efficacy Endpoints

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

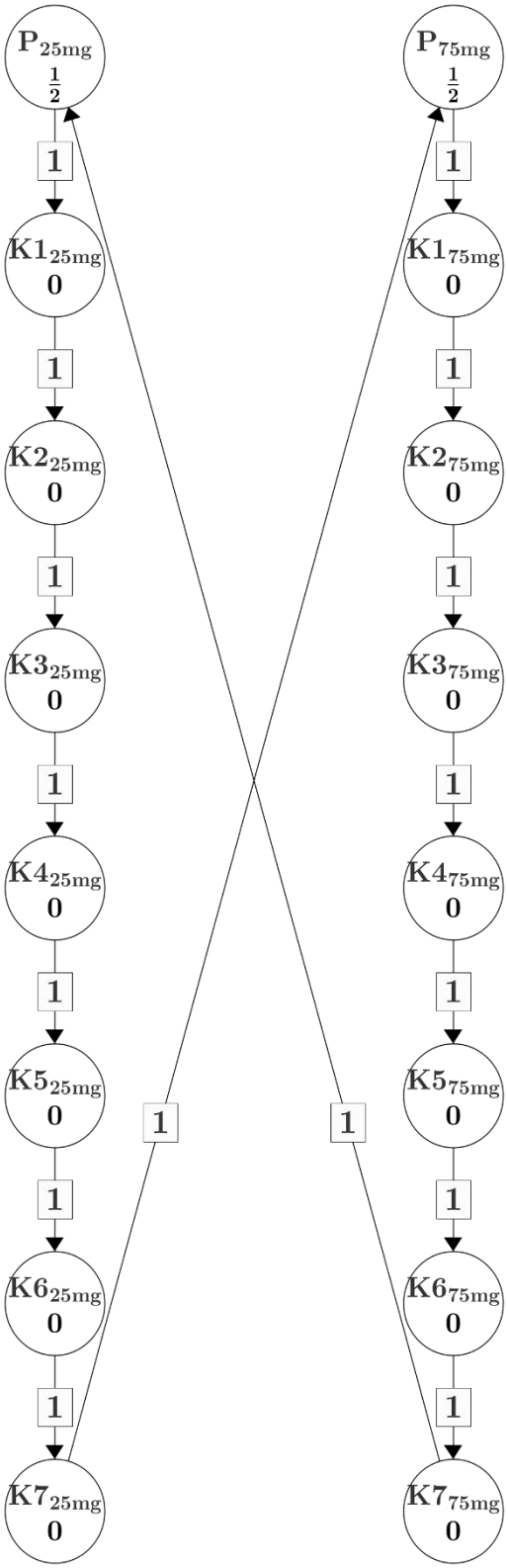
6.3.2 Supplementary Analyses of Other Secondary Efficacy Endpoints

No supplementary 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 2 ontamalimab treatment groups 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 7 key secondary endpoints (K1-K7) within a pairwise treatment comparison. A graphical visualization of the α propagation is presented in Figure 3.

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 Subgroup Analyses

No subgroup analyses are planned for this study.

7. SAFETY ANALYSIS

The safety analyses will be performed using the Safety Set. Safety variables include adverse events (AEs), clinical laboratory variables, vital signs, electrocardiogram (ECG) variables, anti-drug antibody (ADA) and neutralizing antibody (NAb) variables, and neurological variables. For each safety variable, the last value collected prior to the first dose of double-blind IP in induction study (SHP647-301 or SHP647-302) 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 MNT baseline and through the end of treatment visit.

Safety outputs will be presented by treatment group listed in Appendix 16.6. The primary set will present data by SHP647-303 treatment, regardless of the treatment in the induction study. The secondary set will present data by both the induction study treatment and the SHP647-303 study treatment.

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 IP in the SHP647-303 study.

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 IP, related serious AEs (SAEs), TEAEs leading to discontinuation of IP, TEAEs leading to study discontinuation, and TEAEs leading to death.

The number of events, incidence, and percentage of subjects reporting TEAEs will be presented by treatment group as listed in Appendix 16.6; by preferred term; by SOC and preferred term; and by SOC, preferred term, and maximum severity. Treatment-emergent AEs considered related to IP will also be summarized by SOC and preferred term. If more than one 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 IP.

The most common TEAEs (incidence ≥2% in any treatment group) will be summarized by preferred term in descending frequency by treatment group as listed in Appendix 16.6.

Serious TEAEs, TEAEs leading to discontinuation of the study or study medication, and injection site AEs will be summarized by SOC, preferred term, and treatment group as defined in Appendix 16.6. 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 6 other identified potential risks: immunotoxicity, immunogenicity, infection, vascular and thrombotic events, local tolerability, and malignant tumors. Potential risks will be summarized by treatment group as listed in Appendix 16.6. Important potential risks will be listed.

7.1.1.1 Hypersensitivity

Potential hypersensitivity reactions such as serum sickness, vasculitis, or Arthus reactions to ontamalimab will be regarded as adverse events of special interest (AESI). 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) of event it was, and recommendations of permanent discontinuation or re-challenge with IP. Reported hypersensitivity events, adjudicated hypersensitivity events, and study drug recommendation will be summarized by treatment groups as listed in Appendix 16.6.

The number of hypersensitivity reactions and percentage of subjects with hypersensitivity reactions as adjudicated will be summarized by treatment group and ontamalimab all doses; and 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 induction study (SHP647-301 or SHP647-302) baseline at each assessment time point for quantitative variables will be presented by treatment group as listed in Appendix 16.6 for the following clinical laboratory variables. The number and percentage of subjects for qualitative variables in urinalysis will be presented by treatment group as listed in Appendix 16.6 in the Safety Set.

Serum chemistry

- alkaline phosphatase
- aspartate aminotransferase (AST)
- alanine aminotransferase (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-MNT baseline (value collected after the first exposure to IP in the SHP647-303 study) PCI values will be tabulated by treatment group as listed in Appendix 16.6. The percentages will be calculated relative to the number of subjects with at least 1 post-MNT baseline assessment. The numerator is the total number of subjects with at least 1 post-MNT baseline PCI value. A supportive listing of subjects with post-MNT baseline PCI values will be provided including the subject number, site, induction study (SHP647-301 or SHP647-302) baseline, MNT baseline, and post-MNT baseline values.

Figures will be presented for hematology and chemistry to show the changes in laboratory parameters over time. Data will be presented as box-and-whisker plots by treatment group as listed in Appendix 16.6 at each visit, with 1 laboratory parameter per page.

Shifts from induction study (SHP647-301 or SHP647-302) baseline category to each visit will be presented by treatment group as listed in Appendix 16.6 for hematology, chemistry, and urinalysis. For hematology and chemistry, shifts will be categorized as Low, Normal, or High. For urinalysis, shifts will be categorized as Abnormal or Normal.

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
MCH	All		<LLN	>ULN
MCHC	All		<LLN	>ULN
MCV	All		<LLN	>ULN

Table 3: Criteria for Potentially Clinically Important Laboratory Tests

Parameter	Age Range	Sex	Outlier Criteria[a]	
			Low	High
Erythrocyte (RBC)	All		<3.0 x 10^6/µL	NA
Leukocytes (WBC)	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
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

Table 3: Criteria for Potentially Clinically Important Laboratory Tests

Parameter	Age Range	Sex	Outlier Criteria[a]	
			Low	High
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

LLN=lower limit of normal provided by the laboratory; NA=not applicable; ULN=upper limit of normal provided by the laboratory.

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

7.3 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.

7.4 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 induction study (SHP647-301 or SHP647-302) baseline at each post-baseline visit and at the end of the study will be presented by treatment group as listed in Appendix 16.6.

For pulse rate, a post-MNT baseline value is considered as a PCI value if its meets both criteria for observed value and change from induction study (SHP647-301 or SHP647-302) baseline. For systolic/diastolic blood pressure, a post-MNT baseline value is considered as a PCI value if it meets criteria for observed value or change from induction study (SHP647-301 or SHP647-302) baseline. For weight and BMI, a post-MNT baseline value is considered as a PCI value if it meets criteria for change from induction study (SHP647-301 or SHP647-302) baseline. The PCI criteria are listed in Table 4. The number and percentage of subjects with PCI post-MNT baseline values will be tabulated by treatment group as listed in Appendix 16.6. The percentages will be calculated relative to the number of subjects with induction study (SHP647-301 or SHP647-302) baseline and at least 1 post-MNT baseline assessment. The numerator is the total number of subjects with at least 1 PCI post-MNT baseline vital sign value. A supportive listing of subjects with post-MNT baseline PCI values will be provided including the subject number, site, induction study (SHP647-301 or SHP647-302) baseline, MNT baseline, and post-MNT baseline values.

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

7.5 Electrocardiogram (ECG)

Table 4: Criteria for Potentially Clinically Important Vital Signs

Vital Sign Parameter	Flag	Criteria[a]	
		Observed Value	Change from MNT 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%
Temp. (°C)		NA	NA

[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 MNT 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 MNT baseline. For weight and BMI, a post-baseline value is considered as a PCI value if it meets criteria for change from MNT baseline.

A central ECG reader will be used. Descriptive statistics for ECG variables (eg, heart rate, PR interval, QRS interval, QT interval, QTc interval, QTcB, and QTcF) and their changes from induction study (SHP647-301 or SHP647-302) baseline at each assessment time point will be presented by treatment group as listed in Appendix 16.6. Electrocardiogram interpretation will be summarized by visit. A shift table from induction study (SHP647-301 or SHP647-302) baseline to each visit for qualitative ECG results will be presented by treatment group as listed in Appendix 16.6.

Electrocardiogram 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 available induction study (SHP647-301 or SHP647-302) baseline values and post-MNT baseline PCI values will be tabulated by treatment group as listed in Appendix 16.6. The percentages will be calculated relative to the number of subjects with available induction study (SHP647-301 or SHP647-302) baseline values and at least 1 post-MNT baseline assessment. The numerator is the total number of subjects with at least 1 post-MNT 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

ECG=electrocardiogram

7.6 Other Safety Data

7.6.1 Targeted Neurological Assessment

The targeted neurological examination and neurological consult evaluation results with unexplained abnormal neurological findings will be summarized at screening and at each visit and by treatment group as listed in Appendix 16.6. The number and percentage of subjects with targeted neurological examination 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 as listed in Appendix 16.6. The number and percentage of subjects who were referred for a neurological consultation and the results (no PML, PML, no clinically significant finding, other clinically significant finding, not done, other) will also be summarized by treatment group as listed in Appendix 16.6. The neurological evaluation and consultation results will be listed for the Safety Set.

7.6.2 Immunogenicity

Presence of ADAs will be listed and summarized by visit for each treatment group as listed in Appendix 16.6.

Anti-drug antibodies 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 1 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 by treatment group as listed in Appendix 16.6.

Neutralizing antibodies will be tested on ADA-positive subjects, and samples will be defined as NAb-positive or -negative. Presence of NAbs will be listed and summarized for all ADA-positive subjects by visit for each treatment group as listed in Appendix 16.6.

Anti-drug antibody prevalence will also be calculated and summarized by treatment group as listed in Appendix 16.6. Antidrug antibody prevalence is the proportion of study population having drug-reactive antibodies (ADA) at any time point (including pre-existing antibodies) during the study.

Anti-drug antibody incidence will be calculated and summarized by treatment group as listed in Appendix 16.6. Anti-drug antibody 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.

7.6.3 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.

7.6.4 Physical Examination

Complete and targeted physical examinations will be performed at the time points specified in Appendix 16.1. Physical examination results will be listed for the Safety Set.

8. [redacted]

8.1 [redacted]

[redacted]

8.2 [redacted]

[redacted]

8.3 [redacted]

8.4 [redacted]

9. [redacted]

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. The study disruptions related to COVID-19 will be listed for the Safety 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 study. The DMC will not review efficacy data until the time of unblinding. 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 1 level of precision greater than the data collected. Standard deviation will be displayed to 2 levels of precision greater than the data collected.

Categorical and count variables will be summarized by the number of subjects (n) and the percentage of subjects in each category. Percentages will be presented to 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 Treatment Failure

Treatment failure is declared for those randomized into the SHP647-303 study prior to Protocol Amendment 2 when an endoscopic subscore that has increased by at least 1 point over baseline in this maintenance study or a value of at least 2 and an increase in clinical subscore (SF plus RB score) of at least 2 points.

Treatment failure is declared for those randomized into the SHP647-303 study under Protocol Amendment 2 when the 3 following conditions are met:

- E-diary documented sustained* symptomatic worsening that includes RB
- No other etiology (ie, infection) can be identified
- Endoscopy confirms disease activity.

*Sustained: for at least 1 week if the RB score has increased by ≥2 points, or at least 2 weeks if the RB score has increased by <2 points.

If treatment failure is not confirmed, subjects may continue with their next scheduled visit.

12.3 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 to Table 9.

Should there be more than 1 assessment mapped into a given study visit with nonmissing 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, Partial Mayo, Vital Sign, Weight, and Endoscopy Assessments

Visit	Planned Study Day	Start Day of Window	End Day of Window
Week 4	28	2	42
Week 8	56	43	70

Week 12	84	71	98
Week 16	112	99	126
Week 20	140	127	154
Week 24	168	155	182
Week 28	196	183	210
Week 32	224	211	238
Week 36	252	239	266
Week 40	280	267	294
Week 44	308	295	322
Week 48	336	323	350
Week 52	364	351	EOT Date
Follow-up	EOT Date + 84	EOT Date +1	EOF Date

EOT=end of treatment; PRO=patient-reported outcomes; RB=rectal bleeding; SF=stool frequency
*Endoscopy is performed only as part of the Week 52 assessments. To accommodate the scheduling of the endoscopy prior to the Week 52 visit, the start day of the analysis window will be extended to >342 for this parameter only if there is a valid assessment of SF and RB mapped to Week 52.
*Follow-up visit applies to vital sign and weight assessments only.

Table 7: Visit Windows (Study Day Based) – ADA Assessments

Visit	Planned Study Day	Start Day of Window	End Day of Window
Week 12	84	2	126
Week 24	168	127	210
Week 36	252	211	EOT Date

ADA=anti-drug antibody; EOT=end of treatment
*Following the implementation of Protocol Amendment 3, ADA assessments are not performed at Week 52.

Table 8: Visit Windows (Study Day Based) – [redacted] and ECG Assessments

Visit	Planned Study Day	Start Day of Window	End Day of Window
Week 12	84	2	126
Week 24	168	127	266
Week 52	364	267	EOT Date

ECG=electrocardiogram; EOT=end of treatment; [redacted]

Table 9: Visit Windows (Study Day Based) – Safety Lab, Physical Examination, and Neurological Testing Assessments

Visit	Planned Study Day	Start Day of Window	End Day of Window
Week 12	84	2	126
Week 24	168	127	210
Week 36	252	211	308
Week 52	364	309	EOT Date
Follow-up	EOT Date + 84	EOT Date + 1	EOF Date
EOF=end of follow-up; EOT=end of treatment			

12.4 Derived Efficacy Endpoints

12.4.1 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 on endoscopy (0-3)
- Physician's global assessment (0-3).

The calculation of the total score requires SF subscore and RB subscores, reported by subjects in the PRO UC daily e-diary, and PGA and centrally read endoscopic subscores.

The Mayo SF and RB 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 visit 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. Physician's global assessment and centrally read endoscopic subscores will be collected at MNT baseline

and the Week 52 visit. The total Mayo score is the sum of the 4 subscores, and will be calculated at MNT baseline and at the Week 52 visit. The total Mayo score from the induction baseline is calculated in an identical fashion in the induction studies (SHP647-301 and SHP647-302) and will be carried over for use in the SHP647-303 study.

12.4.2 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.

12.4.3 Mucosal Healing

Mucosal healing is based on the endoscopic and histological assessment at the Week 52 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 MNT baseline and Week 52 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.

12.4.4 Total Sign/Symptom Score

Patient-reported UC signs and symptoms data will be collected using a daily e-diary throughout the study. Subjects will be required to enter e-diary data daily. 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, <8 out of 10 diary entries) when compared to 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 the same scale as shown in Table 10. No conversion will be applied for average scores of worst abdominal pain over the past 24 hours.

Table 10 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.

12.5 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 to Table 9.

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-MNT baseline assessments will be used for PCI value determination.

12.6 Handling of Missing, Unused, and Spurious Data

12.6.1 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.

12.6.2 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).

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

The induction study (SHP647-301 or SHP647-302) imputed dates will be used for prior medications for those collected in induction study.

For prior or concomitant medications collected in SHP647-303, 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.

12.6.3.1 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.

12.6.3.1.1 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 IP, then the day and month of the date of the first dose of IP in the SHP647-303 study 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 IP in the SHP647-303 study, 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 IP in the SHP647-303 study, then 01 January will be assigned to the missing fields.

12.6.3.1.2 Missing Month Only

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

12.6.3.1.3 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 IP, then the day of the date of the first dose of IP in the SHP647-303 study 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 IP in the SHP647-303 study, 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 IP in the SHP647-303 study, then the first day of the month will be assigned to the missing day.

12.6.3.2 Incomplete Stop Date

The following rules will be applied to impute the missing numerical fields. If the date of the last dose of IP in the SHP647-303 study 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.

12.6.3.2.1 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 IP, then the day and month of the date of the last dose of IP in the SHP647-303 study 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 IP in the SHP647-303 study, 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 IP in the SHP647-303 study, then 01 January will be assigned to the missing fields.

12.6.3.2.2 Missing Month Only

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

12.6.3.2.3 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 IP, then the day of the date of the last dose of IP in the SHP647-303 study 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 IP in the SHP647-303 study, 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 IP in the

SHP647-303 study, then the first day of the month will be assigned to the missing day.

12.6.4 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.

12.6.4.1 Incomplete Start Date

Follow the same rules as in Section 12.6.3.1.

12.6.4.2 Incomplete Stop Date

Not applicable.

12.6.5 Missing Severity Assessment for Adverse Events

If severity is missing for an AE starting prior to the date of the first dose of IP, 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 IP in the SHP647-303 study, 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.

12.6.6 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 IP in the SHP647-303 study, 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.

12.6.7 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.

14. CHANGES TO ANALYSIS SPECIFIED IN PROTOCOL

The following changes to the analysis specified in Protocol Amendment 3 dated 17 Sep 2020 (final version 1.0 dated 10 Jul 2017, Amendment 1 dated 11 Sep 2018, Amendment 2 dated 11 Nov 2019) have been made.

Changed the other secondary endpoint total sign/symptom score calculation from sum of daily e-diary entries to average of daily e-diary entries.

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.

16. APPENDICES

16.1 Schedule of Activities Prior to Amendment 3

	Baseline[a]	Treatment													Follow-Up[b]	
Study Week	0	4	8	12	16	20	24	28	32	36	40	44	48	52[c]	60[c]	68[c]
Visit Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14 (Pt 1)[d]	14 (Pt 2)	15	16
Study Day	1	28	56	84	112	140	168	196	224	252	280	308	336		364	420	476
Visit Window	None	±7 days													± 7 days	±7 days	
Informed consent/assent	X																
Eligibility assessment[e]	X																
Medical history[f]	IDT																
Complete physical examination[g]	IDT														X		X
Targeted physical examination[g]				X			X			X							
Targeted neurological assessment[h]	IDT			X			X			X					X		X
Vital signs[i]	IDT	X	X	X	X	X	X	X	X	X	X	X	X		X		X
Weight	IDT	X	X	X	X	X	X	X	X	X	X	X	X		X		X
12-lead ECG[i]	IDT			X			X								X		
Contraception check[j]	IDT	X	X	X	X	X	X	X	X	X	X	X	X		X		X
Laboratory Assessments																	
Hematology	IDT			X			X			X					X		X
Serum chemistry	IDT			X			X			X					X		X
Urinalysis	IDT			X			X			X					X		X
FSH[k]	X[j]																
Urine β–hCG[l]	IDT	X	X	X	X	X	X	X	X	X	X	X	X		X		X
[redacted]	IDT			X			X								X		
[redacted]	IDT			X			X								X		
[redacted]	IDT			X			X								X		
[redacted]	IDT			X			X								X		
[redacted]	IDT			X			X			X					X		
ADA and NAb sampling[m]	IDT			X			X			X					X		
Endoscopic Procedure																	
Endoscopy (including biopsy)[n]	IDT													X			
UC Assessments[o]																	
Total Mayo score	IDT														X[p]		
Partial Mayo score	IDT	X	X	X	X	X	X	X	X	X	X	X	X				
PRO-UC e-diary data instruction	X																

	Baseline[a]	Treatment														Follow-Up[b]	
Study Week	0	4	8	12	16	20	24	28	32	36	40	44	48	52[c]		60[c]	68[c]
Visit Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14 (Pt 1)[d]	14 (Pt 2)	15	16
Study Day	1	28	56	84	112	140	168	196	224	252	280	308	336		364	420	476
Visit Window	None	±7 days													± 7 days	±7 days	
PRO-UC daily e-diary data[q]	IDT	X	X	X	X	X	X	X	X	X	X	X	X	X			
Health Assessment[r]																	
IBDQ	IDT			X			X								X		
[redacted]	IDT			X			X								X		
Hospitalizations, inpatient days, [redacted] (HRUA)	IDT	X	X	X	X	X	X	X	X	X	X	X	X		X		X
[redacted]	IDT			X			X								X		
[redacted]	IDT			X			X								X		
[redacted]	IDT			X			X								X		
SF-36-v.2, acute	IDT			X			X								X		
[redacted]	IDT														X		
Treatment Procedures																	
Randomization[s]	X																
Administration of SHP647 or placebo[s]	X	X	X	X	X	X	X	X	X	X	X	X	X				
Adverse events	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Concomitant medication and procedures	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Dispense stool collection kit for stool sample[t]			X			X							X				

	Baseline[a]	Treatment														Follow-Up[b]	
Study Week	0	4	8	12	16	20	24	28	32	36	40	44	48	52[c]		60[c]	68[c]
Visit Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14 (Pt 1)[d]	14 (Pt 2)	15	16
Study Day	1	28	56	84	112	140	168	196	224	252	280	308	336		364	420	476
Visit Window	None	±7 days													± 7 days	±7 days	

ADA=antidrug antibodies; β-hCG=beta-human chorionic gonadotropin; [redacted]; ECG=electrocardiogram; [redacted]; ET=early termination; FSH=follicle-stimulating hormone; HEENT=head, eyes, ears, nose, and throat; HRUA=Healthcare Resource Utilization Assessment; IBDQ=Inflammatory Bowel Disease Questionnaire; IDT=induction study (SHP647-301 or SHP647-302); LTS=long-term safety extension; [redacted]; NAb=neutralizing antibody; [redacted]; PRO=patient-reported outcomes; Pt=part; SF-36=Short Form-36 health survey; [redacted]; UC=ulcerative colitis; [redacted].

[a] The Week 12 assessment from the induction study (SHP647-301 or SHP647-302) will be used as the baseline (Day 1/Week 0) assessments for Study SHP647-303.

[b] No follow-up is required if subject is entering the LTS study (SHP647-304) at the Week 52 visit (Visit 14).

[c] Any subject who prematurely withdraws from the study should return for an early termination visit and then enter into the safety follow-up period if not entering the LTS study (SHP647-304). For subjects participating in the 16-week safety follow-up period (not entering the LTS study), the Week 60 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 68 visit will take place at the study site.

[d] Part 1 of Visit 14 should be scheduled 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] Eligibility will be assessed after the consent form is signed and after induction study SHP647-301 or SHP647-302 Week 12, Visit 6 (Part 2) procedures are completed.

[f] Medical history for induction study SHP647-301 or SHP647-302 will be used as the baseline medical history data for study SHP647-303.

[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. A targeted physical examination only 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] Subject will be evaluated to reveal any potential abnormalities in the following neurologic domains: vision, motor, tactile sensation, coordination/cerebellar function, speech, verbal comprehension, cognition/behavior.

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

[j] Contraception check should be performed for female subjects of childbearing potential and male subjects who are with a partner of childbearing potential.

[k] For confirmation of postmenopausal status in females who have had 12 consecutive months of spontaneous amenorrhea and are ≥51 years of age. This does not need to be performed if postmenopausal status was confirmed by FSH in induction study SHP647-301 or SHP647-302. If a female subject's status has changed to postmenopausal since the induction studies (SHP647-301 or SHP647-302), the FSH confirmation test will be done at baseline (Visit 1) of SHP647-303. Once FSH results are received, and if postmenopausal status is uncertain, the routine pregnancy testing will continue as planned for the remainder of the study. If postmenopausal status is confirmed, routine pregnancy testing is no longer required.

[l] For females of childbearing potential that are not surgically sterile, do not have confirmed ovarian failure, or do not meet the definition of postmenopausal.

	Baseline[a]	Treatment													Follow-Up[b]	
Study Week	0	4	8	12	16	20	24	28	32	36	40	44	48	52[c]	60[c]	68[c]
Visit Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14 (Pt 1)[d]	14 (Pt 2)	15	16
Study Day	1	28	56	84	112	140	168	196	224	252	280	308	336		364	420	476
Visit Window	None	±7 days													± 7 days	±7 days	

[m] Samples must be collected before administration of investigational product.

[n] Flexible sigmoidoscopy or colonoscopy (if preferred). Biopsy samples will be collected for histological evaluation using the Geboes Score classification.

[o] Mayo and PRO-UC assessments will be based on subject daily e-diary entries.

[p] The total Mayo score at Week 52 (Visit 14, Part 2) will be calculated based on the centrally read endoscopic subscore for the endoscopy performed at Week 52 (Visit 14, Part 1).

[q] PRO UC daily e-diary will be collected daily for 10 days before each visit. See Section 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, <8 out of 10 diary entries) when compared with the previous visit.

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

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

[t] 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.

16.2 Schedule of Activities Under Amendment 3

	Baseline[a]	Treatment														Follow-Up[b]
Study Week	0	4	8	12	16	20	24	28	32	36	40	44	48	52/ET[c]		64[c]
Visit Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14 (Pt 1)[d]	14 (Pt 2)[d]	15
Study Day	1	28	56	84	112	140	168	196	224	252	280	308	336		364	448
Visit Window	None	±10 days													±10 days	±10 days
Informed consent/assent	X															
Eligibility assessment[e]	X															
Medical history[f]	IDT															
Complete physical examination[g]	IDT														X	X
Targeted physical examination[g]				X			X			X						
Targeted neurological assessment[h,i]	IDT			X			X			X					X	X
Vital signs[j]	IDT	X	X	X	X	X	X	X	X	X	X	X	X		X	X
Weight	IDT	X	X	X	X	X	X	X	X	X	X	X	X		X	X
12-lead ECG[j]	IDT			X			X								X	
Contraception check[k]	IDT	X	X	X	X	X	X	X	X	X	X	X	X		X	X
Laboratory Assessments[l]																
Hematology	IDT			X			X			X					X	X
Serum chemistry[m]	IDT			X			X			X					X	X
Urinalysis	IDT			X			X			X					X	X
Stool microbiology[n]																
FSH[o]	X[j]															
Urine β–hCG[p]	IDT	X	X	X	X	X	X	X	X	X	X	X	X		X	X
[redacted]	IDT			X			X								X	
[redacted]	IDT			X			X								X	
ADA and NAb sampling[q]	IDT			X			X			X						
Endoscopic Procedure																
Endoscopy (including biopsy)[r]	IDT													X[r]		
UC Assessments[s]																
Total Mayo score	IDT														X[t]	
Partial Mayo score	IDT	X	X	X	X	X	X	X	X	X	X	X	X			
PRO-UC e-diary data instruction	X															
PRO-UC daily e-diary data[u]	IDT	X	X	X	X	X	X	X	X	X	X	X	X	X	X	

Study Week	Baseline[a]	Treatment												52/ET[c]		Follow-Up[b]
	0	4	8	12	16	20	24	28	32	36	40	44	48			64[c]
Visit Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14 (Pt 1)[d]	14 (Pt 2)[d]	15
Study Day	1	28	56	84	112	140	168	196	224	252	280	308	336		364	448
Visit Window	None	±10 days													±10 days	±10 days
Treatment Procedures																
Randomization[v]	X															
Administration of ontamalimab or placebo[v,w,x]	X	X	X	X	X	X	X	X	X	X	X	X	X			
Hypersensitivity monitoring[y]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Adverse events	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Concomitant medication and procedures	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Dispense stool collection kit for stool sample[z]			X			X							X			

ADA=antidrug antibody; β-hCG=beta-human chorionic gonadotropin; [redacted]; ECG=electrocardiogram; e-diary=electronic diary; ET=early termination; FSH=follicle-stimulating hormone; HEENT=head, eyes, ears, nose, and throat; IDT=induction study (SHP647-301 or SHP647-302); LTS=long-term safety extension; NAb=neutralizing antibody; PRO=patient-reported outcome; Pt=part; UC=ulcerative colitis

Note: As soon as Amendment 3 is implemented, the subject's next scheduled visit will be the ET visit, which should be conducted no later than 4 weeks (±10 days) from the subject's last study visit prior to the implementation of SHP647-304 Amendment 4.

[a] The Week 12 assessment from the induction study (SHP647-301 or SHP647-302) will be used as the baseline (Day 1/Week 0) assessments for Study SHP647-303.

[b] No follow-up is required if subject is entering the LTS study (SHP647-304) at the Week 52/ET visit (Visit 14).

[c] Any subject who is prematurely withdrawn from the study (including for treatment failure) should return for the ET visit and then enter into the safety follow-up period if not entering the LTS study (SHP647-304). Subjects who enter the safety follow-up period will have a final visit at 12 weeks following the Week 52/ET visit. Both the Week 52/ET and 12-week safety follow-up visits are preferred to be on-site visits; however, due to the COVID-19 situation, these may also be done at a subject's home provided a qualified site staff member performs these evaluations following DTP guidance.

[d] Part 1 of Visit 14 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. For subjects who meet the criteria for treatment failure (as defined in Protocol Section 4.5.1.1) or are discontinuing prior to completing the 52-week treatment period due to early study closure and will be entering the LTS study (SHP647-304), Part 2 of Visit 14 should be scheduled at least 2 weeks after the last dose of investigational product, to allow a sufficient time interval prior the first dose in the LTS study.

[e] Eligibility will be assessed after the consent form is signed and after induction study SHP647-301 or SHP647-302 Week 12, Visit 6 (Part 2) procedures are completed.

[f] Medical history for induction study SHP647-301 or SHP647-302 will be used as the baseline medical history data for Study SHP647-303.

[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. A targeted physical examination only 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.

	Baseline[a]	Treatment														Follow-Up[b]
Study Week	0	4	8	12	16	20	24	28	32	36	40	44	48	52/ET[c]		64[c]
Visit Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14 (Pt 1)[d]	14 (Pt 2)[d]	15
Study Day	1	28	56	84	112	140	168	196	224	252	280	308	336		364	448
Visit Window	None	±10 days													±10 days	±10 days

[h] Subject 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. See Section 7.2.3.3 of the protocol for further details.

[i] In case of a DTP situation, to be performed by remote visits via virtual communications (eg, TeleHealth application).

[j] Vital signs (including blood pressure, pulse, respiratory rate, and temperature) and 12-lead ECG should be performed prior to the collection of blood samples for laboratory assessments.

[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] Clinical laboratory assays (LFT) can also be performed by the local laboratory in case of issues related to the COVID-19 pandemic and if deemed necessary by the investigator to confirm the subject's safety. In such a case, the investigative site must obtain the local laboratory's normal ranges as well as CLIA (Clinical Laboratory Improvement Amendments) certificate, and the investigator must add the local laboratory as appropriate.

[m] Subjects performing home administrations consecutively for 3 months will need to perform LFT per FDA requirement. It may be done locally if it is not possible to collect samples at the central laboratory.

[n] 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 *Clostridium difficile* as described in Appendix 4 of the protocol.

[o] For confirmation of postmenopausal status in females who have had 12 consecutive months of spontaneous amenorrhea and are ≥51 years of age. This does not need to be performed if postmenopausal status was confirmed by FSH in induction study SHP647-301 or SHP647-302. If a female subject's status has changed to postmenopausal since the induction studies (SHP647-301 or SHP647-302), the FSH confirmation test will be done at baseline (Visit 1) of SHP647-303. Once FSH results are received, and if postmenopausal status is uncertain, the routine pregnancy testing will continue as planned for the remainder of the study. If postmenopausal status is confirmed, routine pregnancy testing is no longer required.

[p] For females of childbearing potential that 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 7.3.3.7 of the protocol.

[q] Samples must be collected before administration of investigational product.

[r] Flexible sigmoidoscopy or colonoscopy (if preferred). Biopsy samples will be collected for histological evaluation using the Geboes Score classification and [redacted]. Note: For subjects discontinued from the study under Amendment 3 and prior to the completion of 52 weeks of treatment, endoscopy is **not** required (optional) at the ET visit.

[s] Mayo and PRO-UC assessments will be based on the subject's daily e-diary entries.

[t] The total Mayo score at Week 52/ET (Visit 14, Part 2) will be calculated based on the centrally read endoscopic subscore for the endoscopy performed at Week 52 (Visit 14, Part 1).

	Baseline[a]	Treatment														Follow-Up[b]
Study Week	0	4	8	12	16	20	24	28	32	36	40	44	48	52/ET[c]		64[c]
Visit Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14 (Pt 1)[d]	14 (Pt 2)[d]	15
Study Day	1	28	56	84	112	140	168	196	224	252	280	308	336		364	448
Visit Window	None	±10 days													±10 days	±10 days

[u] PRO-UC daily e-diary will be available throughout the study. Subjects will be required to enter e-diary data daily. See Section 7.3.2.3 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, <8 out of 10 diary entries) when compared with the previous visit.

[v] All patient-reported questionnaires should be completed before completing any other visit assessments.

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

[x] In a situation in which a subject is not able to visit the study site due to the COVID-19 public health emergency, DTP investigational product administration options may be available, in accordance with local regulations (refer to DTP guidance document).

[y] 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.

[z] 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.

Note: See Section 7.3 of the protocol for the order in which assessments should be performed. Timing of visits is relative to SHP647-303 baseline (Visit 1).

16.3 Geographic Regions

Table 11 Geographic Regions

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

Country	Region
Turkey	ROW
Canada	North America
United States	North America

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

16.4 UC Medication History or Use Subgroup Identification

The following UC medications/therapies/procedures are recorded on the concomitant medication indication specific form.

- **Immunosuppressant Experienced at Baseline**
  - Rule
    - Yes: At least one of the listed medications is on or before the date of the first dose of study medication in the induction study
    - No: Otherwise
- **Immunosuppressant Use at MNT Baseline**
  - Rule
    - Yes: At least one of the listed medications is used 6 days prior to the dose of study medication at the Week 12 visit in the induction study
    - No: Otherwise
- **Anti-TNF Experienced**
  - Recorded in induction study eCRF (Ulcerative Colitis History); Experienced (Yes)/Naïve (No)
- **Anti-TNF Failure**
  - Rule
    - Yes: At least one of the anti-TNF medications is on or before the date of the first dose of study medication in the induction study and "Inadequate response" or "Loss of response" was the reason for the discontinuation.
    - No: Otherwise
- **Anti-TNF Failure Times**
  - Rule
    - Naïve: None of the anti-TNF medications is on or before the date of the first dose of study medications in the induction study
    - Experienced without failure: At least one of the anti-TNF medications is on or before the date of the first dose of study medication in the induction study but "Inadequate response" or "Loss of response" is NOT the reason for discontinuation
    - Number of failures: Count (1, 2, or 3 or more): Count the number of unique records on or before the date of the first dose of study medication in the induction study where anti-TNF medication is in the list below and "Inadequate response" or "Loss of response" is the reason for the discontinuation.
- **Glucocorticoid Use at Baseline**
  - Rule
    - Yes:
      - At least one systemic glucocorticoids in the 6 days prior to the first dose of study medication in the induction study given by "oral" route

        OR

- At least one topical glucocorticoids in the 6 days prior to the first dose of study medication in the induction study given by "oral" route
- No: Otherwise

- **Glucocorticoid Use at MNT Baseline**
  - Rule
    - Yes:
      - At least one systemic glucocorticoids in the 6 days prior to the dose of study medication at the Week 12 visit in the induction study given by "oral" route

        OR

      - At least one topical glucocorticoids in the 6 days prior to the dose of study medication at the Week 12 visit in the induction study given by "oral" route
    - No: Otherwise

- **Maximum Prior Treatment Experience**
  - Immunosuppressant experienced:
    - Rule
      - Yes: At least one of the listed medications in or before the date of the first dose of study medication in the induction study
      - No: Otherwise
  - Biologic Failure
    - Rule
      - Yes: At least one of the listed medications on or before the date of first dose of study medication in the induction study and "Inadequate response" or "Loss of response" was the reason for the discontinuation.
      - No: Otherwise
        -
  - Systemic glucocorticoid experienced
    - Rule
      - Yes: least one of the systemic glucocorticoids on or before the date of first dose of study medication in the induction study given by "Oral", "Intravenous", or "Intramuscular"
      - No: Otherwise
  - Topical glucocorticoid experienced
    - Rule
      - Yes:
        - At least one of the topical glucocorticoids on or before the date of first dose of study medication in the induction study given by "oral" or "rectal" route.

        OR

- At least one of the systemic glucocorticoids on or before the date of first dose of study medication in the induction study given by "rectal" route.

OR

- At least one of the following medications on or before the date of first dose of study medication in the induction study regardless of route.
    - PROCTOFOAM
    - PROCTOSEDYL /03159801/"
    - STEROIDS
    - CORTICOSTEROIDS ACTING LOCALLY
    - ENTOCORT ENEMA
    - ENEMAS

- No: Otherwise

- Amino salicylates experienced
    - Rule
        - Yes: At least one of the listed medications on or before the date of first dose of study medication in the induction study
        - No: Otherwise

- **5-ASA Use at MNT Baseline**
    - Rule
        - Yes: At least one of the listed medications in the 6 days prior to the dose of study medication at the Week 12 visit in the induction study given by "oral" or "rectal" route
        - No: Otherwise

16.5 Rescue Therapy for Ulcerative Colitis

In order for a medication to be a rescue therapy for UC, it should be reported on the concomitant medication indication specific form.

If the exposure happens in the follow-up period (beginning on the EOT date), then it would not satisfy the criteria, as we are concerned here only about exposure with regards to the estimand, which is during the treatment period.

- **Biologics with immunomodulatory properties**
  - Any exposure after first dose
- **Non-biologics with immunomodulatory properties**
  - **Immunosuppressants**
    - Rule
      - After first dose and up to and including Week 48: any increase from MNT baseline for more than 5 days
      - After Week 48: Any dose above MNT baseline for greater than 1 day
    - MNT baseline
      - Subjects who have at least one dose in the 6 days prior to the dose of study medication at the Week 12 visit in the induction study will be considered to have a MNT baseline dose >0. The MNT baseline dose to be used for these subjects is the median nonmissing daily dose from these 6 days. Subjects without at least one dose in the 6 days prior to the dose at the Week 12 visit in the induction study will be considered to take a 0 for the dose for the purposes of assessing an increase from MNT baseline
  - **5-ASA**
    - Rule
      - After first dose and up to and including Week 48: any increase from MNT baseline for more than 5 days
      - After Week 48: Any dose above MNT baseline for greater than 1 day
    - MNT baseline
      - Subjects who have at least one dose in the 6 days prior to the dose of study medication at the Week 12 visit in the induction study will be considered to have a MNT baseline dose >0. The MNT baseline dose to be used for these subjects is the median nonmissing daily dose from these 6 days. Subjects without at least one dose in the 6 days prior to the dose at the Week 12 visit in the induction study will be considered to take a 0 for the dose for the purposes of assessing an increase from MNT baseline
    - Note

- Subjects cannot start the study on the combination of Beclometasone and mesalazine so there would be no switching. It would be only according to the rules.

- **Other small molecule immunodulatory active agents**
  - Rule
    - Any exposure after first dose

- **Leukocyte apheresis, other apheresis, and plasma exchange**
  - Rule
    - Any exposure after first dose
- **Systemic glucocorticoids**
  - **Systemic glucocorticoids given via oral or rectal routes of administration**
    - Rule
      - After first dose and up to and including Week 48: any increase from MNT baseline for more than 7 days
      - After Week 48: Any dose above MNT baseline for greater than 1 day
      -
    - Equivalency
      - Because of multiple types of systemic glucocorticoids that could be given to the same subject, we need to apply systemic glucocorticoid equivalency rules. Below is the table to follow.
    - MNT baseline
      - Subjects who have at least one dose in the 6 days prior to the dose of study medication at the Week 12 visit in the induction study will be considered to have a baseline dose >0. The MNT baseline dose to be used for these subjects is the median nonmissing daily dose from these 6 days. Subjects without at least one dose in the 6 days prior to the dose at the Week 12 visit in the induction study will be considered to take a 0 for the dose for the purposes of assessing an increase from MNT baseline
  - **Systemic glucocorticoids given via parenteral routes of administration**
    - Rule
      - Any exposure after first dose

- **Topical glucocorticoids**
  - **Budesonide**
    - Rule
      - Subjects not taking Budesonide at MNT baseline:

- After first dose and up to and including Week 48: more than 9 mg/day for one day or any exposure for more than 5 days
- After Week 48: Any dose above MNT baseline for greater than 1 day

- Subjects taking oral Budesonide at baseline:
  - After first dose and up to and including Week 48: more than 9 mg/day for one day or any increase for more than 5 days
  - After Week 8: Any dose above MNT baseline for greater than 1 day

- MNT Baseline
  - Subjects who have at least one dose in the 6 days prior to the dose of study medication at the Week 12 visit in the induction study will be considered to have a baseline dose >0. The MNT baseline dose to be used for these subjects is the median nonmissing daily dose from these 6 days. Subjects without at least one dose in the 6 days prior to the dose at the Week 12 visit in the induction study will be considered to take a 0 for the dose for the purposes of assessing an increase from MNT baseline

- **Beclomethasone**
  - Rule
    - Subjects not taking Budesonide at MNT baseline:
      - After first dose and up to and including Week 48: more than 5 mg/day for one day or any exposure for more than 5 days
      - After Week 48: Any dose above MNT baseline for greater than 1 day
    - Subjects taking oral Budesonide at MNT baseline:
      - After first dose and up to and including Week 48: more than 5 mg/day for one day or any increase for more than 5 days
      - After Week 8: Any dose above MNT baseline for greater than 1 day
  - MNT Baseline
    - Subjects who have at least one dose in the 6 days prior to the dose of study medication at the Week 12 visit in the induction study will be considered to have a MNT baseline dose >0. The MNT baseline dose to be used for these subjects is the median nonmissing daily dose from these 6 days. Subjects without at least one dose in the 6 days prior to the dose at the Week 12 visit

in the induction study will be considered to take a 0 for the dose for the purposes of assessing an increase from MNT baseline

16.6 Treatment Groups

Data will be presented by treatment group into 3 pages as follows:

Page 1 (SHP647-301 or SHP647-302 Study Treatment/SHP647-303 Study Treatment)

- Ontamalimab 25 mg/Placebo
- Ontamalimab 25 mg/Ontamalimab 25 mg
- Ontamalimab 75 mg/Placebo
- Ontamalimab 75 mg/Ontamalimab 75 mg

Page 2 (SHP647-301 or SHP647-302 Study Treatment/SHP647-303 Study Treatment)

- Placebo/Placebo
- Placebo/Ontamalimab 25 mg
- Placebo/Ontamalimab 75 mg

Page 3 (SHP647-303 Study Treatment)

- Placebo
- Ontamalimab 25 mg
- Ontamalimab 75 mg
- Ontamalimab All Doses
- Total (for disposition, demographics, and baseline characteristics tables)