

**Table 1: Schedule of Assessments**

Study Procedure	Screening <sup>a</sup>		Baseline	Treatment						Follow-up			
	Weeks -6 to -1		Week 0/ Day 1	Week 2	Week 4	Week 8	Week 12	Week 16/ET <sup>b</sup>			Week 24 <sup>c</sup>	Week 32 <sup>c</sup>	
Visit Number	1 (Part 1) <sup>a</sup>	1 (Part 2) <sup>a</sup>	2	3	4	5	6	7 (Part 1) <sup>b</sup>	7 (Part 2) <sup>b</sup>	7 (Part 3) <sup>b</sup>	8	9	
Study Day	<b>-42 to 0</b>		1	14 ±3	28 ±3	56 ±3	84 ±3				112 ±3	168 ±7	224 ±7
JCV antibody banked sample <sup>q</sup>			X										
			X					X	X				
			X					X	X				
			X					X	X				
			X	X	X	X	X	X	X				
ADA and NAb sampling			X	X	X	X	X	X	X				X
Endoscopic Procedure											X		
Colonoscopy (including biopsy) <sup>r</sup>		X											
CD Assessments													
CDAI <sup>s</sup>		X				X	X	X			X		
PRO-CD daily e-diary data instruction	X												
PRO-CD daily e-diary data <sup>t</sup>	X	X	X	X	X	X	X	X	X	X			
SES-CD <sup>u</sup>			X								X		
Health Assessment <sup>v</sup>													
IBDQ			X			X	X				X		
			X			X	X				X		
Hospitalizations, inpatient days, [REDACTED] (HRUA)					X	X	X				X		X
					X	X	X				X		
			X		X	X	X				X		
			X								X		
SF-36, version 2, acute			X			X	X				X		
			X								X		

### 9.8.1 Coprimary Efficacy Endpoints

The coprimary efficacy endpoints are:

- Clinical remission at the Week 16 visit as defined by the following: 2-item PRO subscores of average worst daily abdominal pain  $\leq 3$  (based on 11-point NRS) over the 7 most recent days and average daily stool frequency  $\leq 2$  of type 6/7 (very soft stools/liquid stools) as shown in the BSFS over the 7 most recent days. The 7 most recent days may or may not be contiguous during the 10 days of data collection before colonoscopy preparation, depending on days to be excluded because of missing data. If fewer than 7 days are available, the endpoint will be calculated on all available most recent 6 or 5 days. If fewer than 5 days are available, the endpoint will be treated as missing.
- Endoscopic response at Week 16 as measured by a decrease in SES-CD of at least 25% from baseline.

The coprimary efficacy endpoints, clinical remission at the Week 16 visit and endoscopic response at the Week 16 visit, will each be compared for each active treatment group (25 mg or 75 mg SHP647) to the placebo group using a Cochran-Mantel-Haenszel (CMH) chi-square test stratified by status of prior anti-TNF treatment, glucocorticoid use, and SES-CD at baseline for each of the stages of the study (stage 1 includes subjects whose primary efficacy data are used in the interim analysis and stage 2 includes all other subjects. Note: classification of stage 1 and stage 2 is based on the time of randomization rather than the time of study completion or termination). Subjects with missing data at the Week 16 visit will be considered failures and counted as nonresponders.

Weighted inverse normal p-value combination methods are used to combine the p-values from stage 1 and stage 2 through the following formula:

$$C(p_1, p_2) = 1 - \Phi[w_1 \Phi^{-1}(1-p_1) + w_2 \Phi^{-1}(1-p_2)]$$

Where  $p_1, p_2$  are the p-values computed from the CMH chi-square test for each stage,  $w_i^2 = n_i/(n_1 + n_2)$ ,  $n_1$  and  $n_2$  are the preplanned stage-wise sample sizes that are fixed at the time of the interim analysis based on an original total sample size, and  $\Phi$  denotes the cumulative distribution function of the standard normal distribution (Bretz et al., 2009a). Given that there is no possibility of stopping early for efficacy, that any potential stopping for futility of either or both doses of SHP647 is nonbinding, and that weights are prespecified, the test statistic  $C(p_1, p_2)$  can be compared against the nominal alpha level to assess statistical significance (Chang and Chow, 2008).

The coprimary endpoints will each be tested by the following hypothesis:

$$H_0: \delta = 0$$

$$H_1: \delta \neq 0$$

Where  $\delta$  is the common treatment difference across strata. The common treatment difference is a weighted average of the stratum-specific treatment differences.

This definition is further supported by a study conducted in a similar condition (irritable bowel syndrome) that examined the minimal clinically important difference on the 11-point NRS for abdominal pain ([Spiegel et al., 2009](#)) as well as post hoc analyses of the Phase 2 data from the SHP647 program (Study A7281006, OPERA).

### Rationale for Very Soft Stool/Liquid Stool Frequency

Diarrhea is the most common sign in the presentation of CD, affecting approximately 85% of patients with a diagnosis of CD. In the CDAI, the number of liquid or soft stools (each day for 7 days) is used with a multiplier of 2. The coprimary endpoints for clinical remission in studies SHP647-305 and SHP647-306 requires the use of a definition without any such multiplying factor and will use the BSFS for defining the very soft or liquid stools according to types 6 and 7, respectively. A retrospective study of PROs in CD based on data from randomized controlled studies using rifaximin and methotrexate showed that a mean daily stool frequency score of  $\leq 1.5$  had an area under the receiving operating characteristic curve of 0.79 ([Khanna et al., 2015](#)) and provided a potential cutoff for defining remission as measured by CDAI. In a recent study to select the attributes determining overall disease severity and to rank the importance of and to score these individual attributes for both CD and UC based on specialist opinion, a sample of at least 10 loose stools per week was considered as an attribute contributing to overall disease severity in CD ([Siegel et al., 2016](#)). Based on post hoc analyses of the Phase 2 data in the SHP647 program (Study A7281006, OPERA) and by choosing the population of subjects satisfying the moderate to severe CD inclusion criteria, various cutoffs were explored and a stool frequency  $\leq 2.0$  was found to be optimal in terms of treatment separation while still allowing for a reasonable threshold for remission. Based on these and other recent data that support this cutoff, an average daily stool frequency  $\leq 2$  of type 6/7 (very soft stools/liquid stools) has been chosen as the stool frequency criterion for clinical remission.

### Endoscopic Response

Endoscopic response is defined in 2 ways:

- 25% reduction in SES-CD score (“endoscopic response”)
- 50% reduction in SES-CD score (“enhanced endoscopic response”).

“Endoscopic response” will be used as a coprimary endpoint and “enhanced endoscopic response” will be used as a key secondary endpoint in this study as these magnitudes of changes are likely to be clinically relevant.

Mucosal healing or “endoscopic healing” is considered to be a pivotal long-term target in the treatment of CD; however, partial healing or endoscopic response may also provide benefits. Endoscopic response can be an important indicator that the mucosal inflammation has decreased as an effect of the investigational product. Some treatments may result in a partial initial response, even though at a later stage a complete response may occur. Median duration of remission after 1 year treatment with infliximab was similar in subjects achieving complete absence of mucosal ulcer to subjects who achieved significant but incomplete mucosal healing ([D'Haens et al., 2002](#)).

It is important to note that assessments based on neurological evaluations are collected and evaluated in a different manner than observed or volunteered AEs. Given these differences, no attempt will be made to reconcile any apparent discrepancies between observed or volunteered AEs and data from neurological assessment collected from subjects. Investigators may determine if any finding on neurological testing constitutes an AE. Adverse event incidence rates will not be calculated from these neurological evaluation data but rather from the AE information recorded by the investigator.

#### **7.2.3.4 Adverse Event Collection**

At each study visit, subjects will be questioned in a general way to ascertain if AEs have occurred since the previous visit (eg, “Have you had any health problems since your last visit?”). Adverse events are collected from the time informed consent and/or assent is signed until the end of the study or the defined follow-up period stated in Section 7.1.3 (Section 8, Adverse and Serious Adverse Events Assessment.)

#### **7.2.3.5 Vital Signs**

Vital signs will be measured at the time points specified in Table 1. Additional collection times or changes to collection times will be permitted, as necessary to ensure appropriate collection of safety data. Vital signs include blood pressure, pulse, respiratory rate, and temperature.

Single measurements of sitting blood pressure will be recorded at each time point. Blood pressure should be determined by cuff with the subject’s arm supported at the level of the heart and recorded to the nearest mm Hg using the same method, the same arm (preferably the dominant arm), and the same position throughout the study.

Respiratory rate will be measured with the subject in a comfortable position. The observer should hold the extremity of the subject as a distraction for the subject (ie, pretending he/she is taking the subject’s radial pulse) and count the respiration for 1 minute.

Oral temperature should be taken by placing a digital thermometer under the tongue for at least 30 seconds and the temperature reported in degrees Celsius or Fahrenheit. Tympanic temperature may also be used.

Any deviations from baseline (Visit 2) vital signs that are deemed clinically significant in the opinion of the investigator are to be recorded as an AE, unless documented in the subject’s medical history as a pre-existing medical condition.

#### **7.2.3.6 Clinical Laboratory Evaluations**

All clinical laboratory assays will be performed according to the central laboratory’s normal procedures. Reference ranges are to be supplied by the central laboratory and will be used to assess the clinical laboratory data for clinical significance and out-of-range pathological changes. The investigator should assess out-of-range clinical laboratory values for clinical significance, indicating if the value(s) is/are not clinically significant or clinically significant. Abnormal clinical laboratory values, which are unexpected or not explained by the subject’s clinical condition, may, at the discretion of the investigator or sponsor, be repeated as soon as possible until confirmed, explained, or resolved.

FSH	follicle-stimulating hormone
FWER	family-wise Type I error rate
GCP	Good Clinical Practice
GI	gastrointestinal
HBcAb	hepatitis B core antibody
HBsAg	hepatitis B surface antigen
HBV	hepatitis B virus
HCV	hepatitis C virus
HCVAb	hepatitis C virus antibody
HIPAA	Health Insurance Portability and Accountability Act
HIV	human immunodeficiency virus
HRQL	health-related quality of life
hsCRP	high-sensitivity C-reactive protein
IB	investigator's brochure
IBD	inflammatory bowel disease
IBDQ	Inflammatory Bowel Disease Questionnaire
ICH	International Council for Harmonisation
IgG <sub>2κ</sub>	immunoglobulin G2 kappa
██████████	██████████
IGRA	interferon-gamma release assay
IRB	institutional review board
IRT	interactive response technology
LP	lumbar puncture
LTS	long-term safety extension
██████████	██████████
MTX	methotrexate
NAb	neutralizing antibody
NRS	numerical rating scale
PCR	polymerase chain reaction
████	████
PFS	prefilled syringe
██████████	██████████
██████████	██████████
PML	progressive multifocal leukoencephalopathy

**Rationale:**

This study is designed to evaluate the efficacy and safety of ontamalimab in inducing clinical remission and endoscopic response in subjects with moderate to severe CD.

The CD clinical development program includes 3 completed studies: 1 Phase 1 study (A7281008) and 2 Phase 2 studies (A7281006 and A7281007). The ontamalimab dose selection (25 mg and 75 mg) for this study is based on data from these 3 previous studies, which evaluated the activity of ontamalimab in adult patients with moderately to severely active CD based on CDAI scores between 220 and 450. The Phase 1 study (A7281008, TOSCA) and Phase 2 studies (A7281006, OPERA; and A7281007, OPERA II [long-term safety study]) that investigated the safety, tolerance, pharmacokinetics (PK), and pharmacodynamic (PD) properties of ontamalimab support further clinical development of ontamalimab using subcutaneous (SC) administration in subjects with moderate to severe CD.

**Investigational product, dose, and mode of administration:**

The test product is ontamalimab (SHP647), which will be provided as a sterile aqueous buffered solution for SC administration in a glass prefilled syringe (PFS) with a fixed needle. Each PFS contains 1 mL of ontamalimab solution for injection at an appropriate concentration to provide the intended dose of drug (25 mg or 75 mg). Additional information is provided in the current ontamalimab investigator's brochure.

The reference product is placebo, which will be provided in a PFS with a fixed needle containing 1 mL of placebo solution for SC administration. The placebo solution will contain the same sterile aqueous buffered solution as the test product but will not contain ontamalimab.

**Methodology:**

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

A total of 1032 subjects (387 subjects in the 25 mg ontamalimab treatment group, 387 subjects in the 75 mg ontamalimab treatment group, and 258 subjects in the placebo group) are planned for enrollment into the study. Subjects must be at least 16 years of age and no more than 80 years of age at the time of signing the informed consent/assent form.

The study consists of a screening period up to 6 weeks and a 16-week treatment period. After the screening period, eligible subjects will be randomized to receive 1 of 3 treatments (25 mg ontamalimab, 75 mg ontamalimab, or placebo) in a 3:3:2 ratio. Randomization will be stratified based upon the subject's status of prior antitumor necrosis factor (TNF) treatment (naïve or experienced), glucocorticoid use at baseline (on glucocorticoids at baseline versus not on glucocorticoids at baseline), and Simple Endoscopic Score for CD (SES-CD) at baseline (SES-CD  $\geq$ 17 or SES-CD <17). Subjects will receive SC injections of ontamalimab or placebo, using a PFS, on Week 0/Day 1 (Visit 2), Week 4 (Visit 4), Week 8 (Visit 5), and Week 12 (Visit 6). Subjects will undergo efficacy, [REDACTED], [REDACTED], safety, and health outcome assessments at these visits.

At the end of the 16-week treatment period, subjects will be offered the opportunity to participate in either a double-blind maintenance study (SHP647-307; for subjects who fulfill the entry criteria) or a long-term safety extension (LTS) study (SHP647-304; for subjects who do not fulfill the entry criteria for Study SHP647-307). Subjects who withdraw early from the 16-week treatment period or who do not wish to enter the maintenance study (SHP647-307) or LTS study (SHP647-304) will continue into a 16-week safety follow-up period. Only those subjects who complete the full course of investigational product treatment in the induction studies (SHP647-305 or SHP647-306) will be eligible to continue in the maintenance study or LTS study.

A planned interim analysis for the coprimary endpoints will take place after approximately the first 50% of all randomized subjects in both the SHP647-305 and SHP647-306 studies have either completed the studies or have prematurely withdrawn from the studies. The sample size will be reassessed as part of this interim analysis.

**Inclusion and exclusion criteria:**

**Inclusion criteria:**

Subjects must meet all of the following inclusion criteria to be eligible for enrollment into the study.

1. Subjects and/or their parent or legally authorized representative must have an understanding, ability, and willingness to fully comply with study procedures and restrictions.

**Table 1 Schedule of Assessments**

Study Procedure	Screening <sup>a</sup>		Baseline	Treatment						Follow-up		
	Weeks -6 to -1		Week 0/ Day 1	Week 2	Week 4	Week 8	Week 12	Week 16/ET <sup>b</sup>			Week 24 <sup>c</sup>	Week 32 <sup>c</sup>
Visit Number	1 (Part 1) <sup>a</sup>	1 (Part 2) <sup>a</sup>	2	3	4	5	6	7 (Part 1) <sup>b</sup>	7 (Part 2) <sup>b</sup>	7 (Part 3) <sup>b</sup>	8	9
Study Day	<b>-42 to 0</b>		1	14 ±3	28 ±3	56 ±3	84 ±3			112 ±3	168 ±7	224 ±7
TB test (PPD or QuantiFERON-TB Gold Plus) <sup>q</sup>	X											
JCV antibody banked sample <sup>r</sup>			X									
			X					X	X			
			X					X	X			
			X					X	X			
			X					X	X			
			X	X	X	X	X	X	X			
ADA and NAb sampling			X	X	X	X	X	X				X
<b>Endoscopic Procedure</b>												
Colonoscopy (including biopsy) <sup>s</sup>		X								X		
<b>CD Assessments</b>												
CDAI <sup>t</sup>		X			X	X	X				X	
PRO-CD daily e-diary data instruction	X											
PRO-CD daily e-diary data <sup>u</sup>	X	X	X	X	X	X	X	X	X	X		
SES-CD <sup>v</sup>			X								X	
<b>Health Assessment<sup>w</sup></b>												
IBDQ			X		X	X				X		
			X		X	X				X		
Hospitalizations, inpatient days, [REDACTED] (HRUA)					X	X	X			X		X
					X	X	X			X		
			X		X	X	X			X		
			X							X		
SF-36, version 2, acute			X			X	X			X		
			X							X		

Subjects with inflammatory bowel disease (IBD) (including CD), with or without stoma, who failed or were intolerant to both anti-TNF and immunosuppressant therapy and who had moderate to severe active disease underwent a lumbar puncture (LP), completed induction therapy with 3 doses of 225 mg SHP647 4 weeks apart, and then underwent a second LP 2 ( $\pm 1$ ) weeks later. The primary endpoint was the percent change from baseline (pretreatment) in absolute lymphocyte count in cerebrospinal fluid (CSF) in subjects with IBD after receiving 3 doses of 225 mg SHP647. The mean percentage change from baseline in absolute lymphocytes in CSF was 61.76% with a median change of 35.2% (range: -70.2% to 267.8%). The post-treatment LP/pretreatment LP geometric mean ratio for CSF lymphocytes was 1.33 with the lower bound of the 80% confidence interval (CI)=1.13, which was greater than 0.5, supporting rejection of the null hypothesis (ie, that the percent decrease in total lymphocytes counts after treatment would be  $\geq 50\%$  (equivalent to the geometric mean ratio in total lymphocyte counts being  $\leq 0.5$ ). This result supports the hypothesis that SHP647 does not impair trafficking of lymphocytes into the CNS and thus should not impair CNS immune surveillance.

Study A7281007 (OPERA II) was a Phase 2 open-label extension study to provide additional long-term safety data on subjects with moderate to severe CD who completed Study A7281006 or Study A7281008 and wished to continue to receive SHP647. SHP647 75 mg (with potential dose escalation to 225 mg) SC given every 4 weeks for 72 weeks was generally well tolerated in subjects with CD over the treatment period evaluated in this study. In subjects with positive ADA or neutralizing antibody (NAb) status, exposure to SHP647 was not affected. Serum concentrations of SHP647 in this study were consistent with what was predicted based on the Feeder Study A7281006. There were 2 deaths in the study: 1 subject died of multiple organ dysfunction syndromes in the treatment period and 1 subject died of metastatic neoplasm in the follow-up period. Neither death was reported as related to treatment with the study drug by the investigators. The most frequently reported serious adverse event (SAE) was CD in either the treatment period or the follow-up period. The system organ class with the most subjects experiencing TEAEs was GI disorders. Although Study A7281007 was not placebo controlled, the exploratory efficacy results (based on the modified Harvey Bradshaw Index) indicated that the effect of SHP647 on disease activity was maintained over the duration of treatment.

The SHP647 dose selection (25 and 75 mg) for this study is based on data from these 3 previous studies, which evaluated the activity of SHP647 in adult patients with moderately to severely active CD based on CDAI scores between 220 and 450. The results of a post hoc analysis of remission rate by baseline elevated serum concentration of hsCRP suggested that the greatest treatment effect was at a dose of 22.5 mg. Similarly, post hoc analysis of remission rates by endoscopic severity assessed using the Simple Endoscopic Score for CD (SES-CD) suggested best efficacy at a dose of 75 mg. Therefore, both dosage regimens 25 and 75 mg Q4W have been selected for the Phase 3 testing. The Phase 1 study (A7281008, TOSCA) and Phase 2 studies (A7281006, OPERA; and A7281007, OPERA II [long-term safety study]) that investigated the safety, tolerance, pharmacokinetics, and pharmacodynamic properties of SHP647 support further clinical development of SHP647 using SC administration in subjects with moderate to severe CD.

15 Dec 2017

### **2.2.3 Exploratory Objectives**

The exploratory objectives are as follows:

A horizontal bar chart showing the percentage of respondents who have heard of various terms related to the study of the brain. The y-axis lists the terms, and the x-axis shows the percentage from 0% to 100%. The bars are black.

Term	Percentage
Neuroscience	95%
Neurologist	92%
Neuron	88%
Neurotransmitter	85%
Brain scan	82%
Neurodegenerative disease	78%
Neuroplasticity	75%
Neuroimaging	72%
Neurodevelopmental disorder	68%
Neurodegeneration	65%

### 3. STUDY DESIGN

#### 3.1 Study Design and Flow Chart

This is a Phase 3, randomized, double-blind, placebo-controlled, parallel-group study designed to evaluate the efficacy and safety of SHP647 in inducing clinical remission and endoscopic response in subjects with moderate to severe CD.

At study initiation, a total of 1032 subjects (387 subjects in the 25 mg SHP647 treatment group, 387 subjects in the 75 mg SHP647 treatment group, and 258 subjects in the placebo group) are planned for enrollment into the study ([Figure 1](#)). Subjects must be at least 16 years of age and no more than 80 years of age at the time of signing the informed consent/assent form.

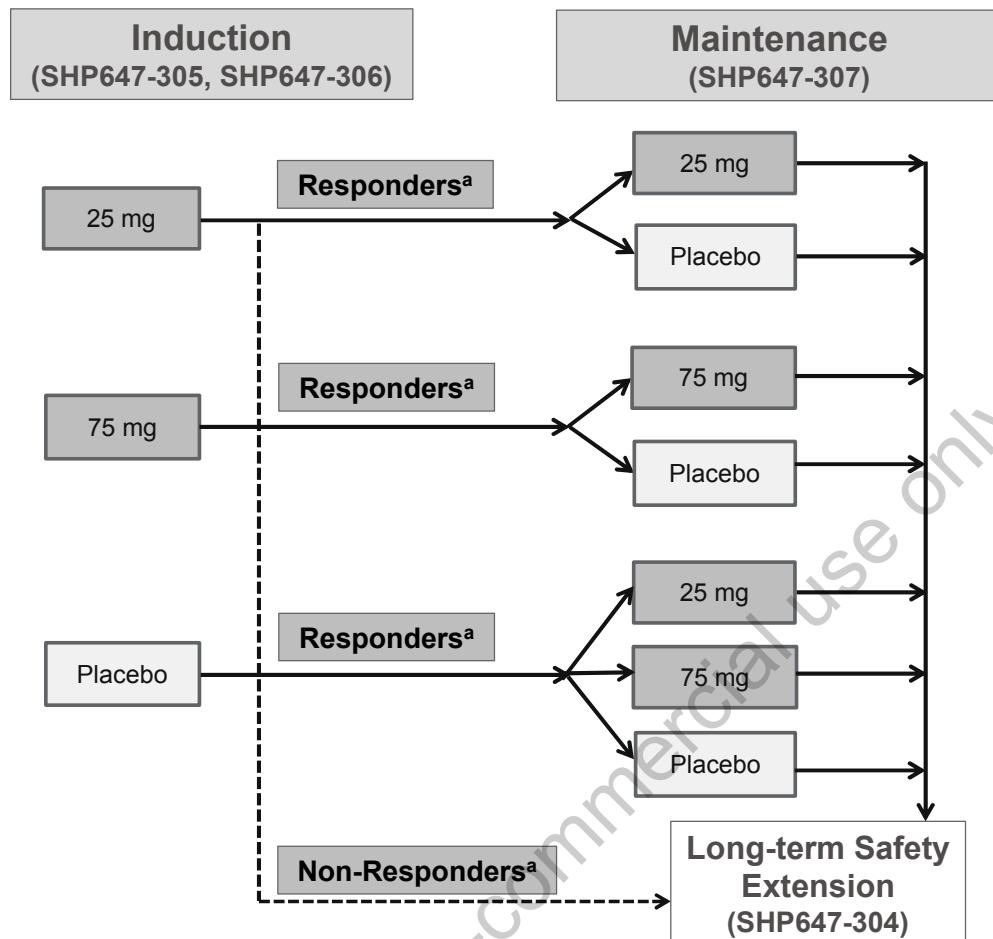
The study consists of a screening period up to 6 weeks and a 16-week treatment period. After the screening period, eligible subjects will be randomized to receive 1 of 3 treatments (25 mg SHP647, 75 mg SHP647, or placebo) in a 3:3:2 ratio. Randomization will be stratified based upon the subject's status of prior anti-TNF treatment (naïve or experienced), glucocorticoid use at baseline (on glucocorticoids at baseline versus not on glucocorticoids at baseline), and SES-CD at baseline (SES-CD  $\geq$ 17 or SES-CD <17). Subjects will receive SC injections of SHP647 or placebo, using a prefilled syringe (PFS), on Week 0/Day 1 (Visit 2), Week 4 (Visit 4), Week 8 (Visit 5), and Week 12 (Visit 6). Subjects will undergo efficacy, [REDACTED], [REDACTED], safety, and health outcome assessments at these visits and at the time points specified in [Table 1](#).

At the end of the 16-week treatment period, subjects will be offered the opportunity to participate in either a double-blind maintenance study (SHP647-307; for subjects who fulfill the entry criteria) or a long-term safety extension (LTS) study (SHP647-304; for subjects who do not fulfill the entry criteria for Study SHP647-307) as shown in [Figure 1](#). Subjects who withdraw early from the 16-week treatment period or who do not wish to enter the maintenance study (SHP647-307) or LTS study (SHP647-304) will continue into a 16-week safety follow-up period. Only those subjects who complete the full course of investigational product treatment in the induction studies (SHP647-305 or SHP647-306) will be eligible to continue in the maintenance study or LTS study.

A planned interim analysis for the coprimary endpoints will take place after approximately the first 50% of all randomized subjects in both the SHP647-305 and SHP647-306 studies have either completed the studies or have prematurely withdrawn from the studies. The sample size will be reassessed as part of this interim analysis. See Section [9.5](#) for further details of the planned interim analysis.

The overall study design is shown in [Figure 2](#).

**Figure 1: Overview of SHP647 Phase 3 Studies in Crohn's Disease**



BSFS=Bristol Stool Form Scale; CD=Crohn's disease; CDAI=CDAI=Crohn's Disease Activity Index; NRS=numerical rating scale; PRO=patient-reported outcomes; SES-CD=Simple Endoscopic Score for Crohn's Disease.

<sup>a</sup> Responders are subjects who either:

(a) Meet endoscopic response criteria of a reduction in SES-CD from baseline by  $\geq 25\%$  at Week 16

OR

(b) Meet at least 1 of the following 4 criteria at Week 16 in addition to no worsening of endoscopic score as measured by SES-CD relative to induction study baseline (SHP647-305 or SHP647-306):

1. Subject is in clinical remission as determined by meeting the criteria for clinical remission using the 2-item PRO, ie, 2-item PRO subscore of average worst daily abdominal pain  $\leq 3$  (based on 11-point NRS) over the 7 most recent days and average daily stool frequency  $\leq 2$  of type 6/7 (very soft stools/liquid stools) as shown in the BSFS over the 7 most recent days.\*
2. Subject has a decrease of at least 70 points in CDAI score (CDAI-70) from baseline.
3. Subject has a decrease of  $\geq 30\%$  and at least 2 points from baseline in the average daily worst abdominal pain over the 7 most recent days\*, with the average daily stool frequency of type 6/7 (very soft stools/liquid stools) either: (i) not worsening from baseline and/or (ii) meeting the criteria for clinical remission, ie, 2-item PRO subscore of average daily stool frequency  $\leq 2$  of type 6/7 (very soft stools/liquid stools) as shown in the BSFS over the 7 most recent days.\*
4. Subject has a decrease of  $\geq 30\%$  from baseline in the average daily stool frequency of type 6/7 (very soft stools/liquid stools) as shown in the BSFS over the 7 most recent days\*, with the average daily worst abdominal pain either: (a) not worsening from baseline and/or (b) meeting the criteria for clinical remission, ie, 2-item PRO subscore of average worst daily abdominal pain  $\leq 3$  (based on 11-point NRS) over the 7 most recent days\*.

\*Note: The 7 most recent days may or may not be contiguous during the 10 days of data collection before colonoscopy preparation, depending on days to be excluded because of missing data. If fewer than 7 days are available, the criterion will be calculated on all available most recent 6 or 5 days. If fewer than 5 days are available, the criterion will be treated as missing.

### Clinical Response

The goal of measuring clinical remission is to have a sensitive clinical measure to assess the complete absence of symptoms or the stabilization of noninflammatory symptoms. However, as response and remission are considered to be on a continuum of improvement or response to treatment, clinical remission is generally the chosen measurement over clinical response. Therefore, clinical remission is used as coprimary endpoint and clinical response is used as a key secondary endpoint in this study.

Clinical response is defined in Section 9.8.2.1. For both clinical response criteria, an additional requirement is that the symptom not being used to assess clinical response (ie, abdominal pain severity or very soft stool/liquid stool frequency) must remain unchanged/not worsen from the baseline score, or meet the criteria for clinical remission for that item of the 2-item PRO (either a 2-item PRO subscore of average daily stool frequency  $\leq 2$  of type 6/7 [very soft stools/liquid stools] as shown in the BSFS or average worst daily abdominal pain  $\leq 3$  [based on 11-point NRS] over the 7 most recent days).

The rationale for needing to meet at a minimum clinical response definition for either abdominal pain severity or very soft stool/liquid stool frequency (and not necessarily both) is based on the supposition that a lack of improvement in 1 of these symptoms is not necessarily an indicator of eventual lack of response (as assessed by the stricter clinical remission criterion). Based on Phase 2 data, it has been observed that the magnitude of placebo response rate can be higher for abdominal pain than for stool frequency. Therefore, the additional criterion of at least a 2-point decrease in abdominal pain severity from baseline is required for assessing clinical response for abdominal pain. Overall, the definition of clinical response used for this study has been chosen to allow for the maximal pool of subjects to be assessed for the effect of treatment and, if appropriate, the continuation of therapy in the maintenance study (SHP647-307).

### Composite Score Endpoint of Both Clinical Remission by 2-item PRO and Endoscopic Response at Week 16

In theory, as the degree of inflammation decreases due to the effect of treatment, both clinical signs and symptoms of CD as well as endoscopic appearance can improve. However, in any given subject, the rates of clinical improvement and endoscopic improvement may not be the same. There are reasons for this discrepancy when evaluating clinical and endoscopic improvement in the same time period, including clinical symptoms not being well correlated to the mucosal inflammation. Due to the transmural feature of the disease, symptoms can correspond to the inflammation in some of the other gut layers as well. Previous clinical studies and clinical observations indicate that the improvement of clinical signs and symptoms and the improvement in endoscopic appearance may not go hand in hand. Significant clinical improvement can precede significant endoscopic improvement. The healing process of the gut mucosa may take a long time and may depend on the baseline severity of the endoscopic appearance, which may not be in line with the actual baseline severity of the symptoms. Therefore, when evaluating clinical remission together with endoscopic endpoints, improvement in endoscopic scores could be more relevant than evaluating mucosal healing in the induction phase.

## 5. PRIOR AND CONCOMITANT TREATMENT

### 5.1 Prior Treatment

Prior treatment includes all treatment (including but not limited to herbal remedies and vitamins) received within 30 days (or PK equivalent of 5 half-lives, whichever is longer) of the first dose of investigational product. Use of biologics for indications other than CD during the 90 days before screening must also be recorded.

Subjects must have had an inadequate response to, or lost response to, or had intolerance to at least 1 conventional treatment such as sulfasalazine or 5-ASA, glucocorticoids, immunosuppressants (AZA, 6-MP, or MTX), or anti-TNF.

**Inadequate response:** The subject experiences continued disease activity despite treatment with an adequate therapeutic dose and treatment course (dictated by the product label and international CD therapeutic guidelines)

**Loss of response:** The subject experiences relapse after an initial clinical response or remission

**Intolerance:** The subject has a history of having experienced an unacceptable or dose limiting toxicity associated with the use of the agent.

All prior and concomitant CD-specific treatments will be recorded. The subject's entire history of biologic CD-specific treatments will be recorded.

### 5.2 Concomitant Treatment

Concomitant treatment refers to all treatment taken between the dates of the first dose of investigational product and the end of the safety follow-up period of this study, inclusive.

#### 5.2.1 Permitted Treatment

Subjects must remain on stable doses of permitted CD treatments until completion of the Week 16 visit, unless decreases are required because of AEs. Stable doses of the following treatments for CD are permitted as concomitant medication:

- Oral sulfasalazine or 5-ASA, providing that the dose is stable for at least 2 weeks before baseline (Visit 2)
- Immunosuppressants (AZA, 6-MP, or MTX), providing that the dose is stable for at least 8 weeks before baseline (Visit 2)
- Oral glucocorticoids (prednisone or equivalent [Appendix 3] up to a maximum of 20 mg/day or budesonide up to a maximum of 9 mg/day), providing that the dose is stable for at least 2 weeks before baseline (Visit 2). After baseline (Visit 2), a stable dose of 20 mg/day prednisone or equivalent oral systemic corticosteroid dose is allowed. Steroids may be decreased due to AEs.

Note: Rectal 5-ASA and parenteral or rectal glucocorticoids are prohibited from within 14 days before screening colonoscopy.

If a subject has had the following procedures performed as a part of standard medical care within 12 weeks before screening (Visit 1), these procedures do not need to be repeated as a part of screening:

- Chest x-ray
- Documented negative PPD test or IGRA for TB.

### **7.1.1.2 Baseline Visit (Visit 2, Week 0)**

The baseline visit (Visit 2) will take place on Day 1 (Week 0). The assessments and procedures specified in [Table 1](#) will be performed.

After eligibility has been reconfirmed and all baseline procedures and assessments have been completed, each subject will be randomized to receive 1 of the 3 treatments as described in Section [6.2.2](#) and the first dose of investigational product will be administered.

Results of the baseline laboratory tests are not required for investigational product administration but must be reviewed as soon as possible thereafter.

### **7.1.2 Treatment Period**

#### **7.1.2.1 Visits 3, 4, 5, and 6 (Weeks 2, 4, 8, and 12)**

Visits 3, 4, 5, and 6 are scheduled to take place on Day  $14 \pm 3$  days (Week 2), Day  $28 \pm 3$  days (Week 4), Day  $56 \pm 3$  days (Week 8), and Day  $84 \pm 3$  days (Week 12), respectively. The assessments and procedures specified in [Table 1](#) will be performed.

#### **7.1.2.2 Final On-treatment Visit: Visit 7, Parts 1, 2, and 3 (Week 16/Early Termination)**

The Week 16/ET visit (Visit 7) consists of 3 parts.

Part 1 of Visit 7 should be scheduled 1 to 3 day(s) before Part 2; this will allow for blood samples to be taken before starting the colonoscopy preparation and before the colonoscopy procedure at Part 2 of the visit. The Week 16/ET assessments and procedures that will take place during Part 1 are specified in [Table 1](#).

Part 2 of Visit 7 should be scheduled preferably within 5 to 7 days before Part 3; this will allow sufficient time to obtain the data from the centrally read colonoscopy. The Week 16/ET assessments and procedures that will take place during Part 2 are specified in [Table 1](#).

Part 3 of Visit 7 will take place on Day  $112 \pm 3$  days. The Week 16/ET assessments and procedures that will take place during Part 3 are specified in [Table 1](#).

At Part 3 of Visit 7, after review of CD assessments, health outcome assessments, and safety assessments, it will be determined whether the subject should enroll in the maintenance (SHP647-307) or LTS (SHP647-304) studies or enter the follow-up period of this study. Entry into the maintenance or LTS studies is dependent upon whether the subject fulfills the efficacy entry criteria of the maintenance study (SHP647-307), including achieving endoscopic and/or clinical response, and whether the subject agrees to participate.

The Week 16 assessments and procedures will also form the ET assessments for any subjects who are withdrawn early or discontinued from the study.

### **7.1.3 Follow-up Period**

Subjects who are withdrawn early from the study, or who do not enter either the maintenance or LTS studies, should enter the 16-week safety follow-up period for safety monitoring.

During the safety follow-up period, the Week 24 visit (Visit 8) will take place on Day  $168 \pm 7$  days, or 8 weeks  $\pm 7$  days after the subject's last visit in the treatment period for subjects who are withdrawn early from the study. This visit will routinely be conducted by telephone; however, as an exception, the visit can be performed as a study site visit if preferred.

At the end of the safety follow-up period, there will be a visit at the site on Day  $224 \pm 7$  days, or 16 weeks  $\pm 7$  days after the subject's last visit in the treatment period for subjects who are withdrawn early from the study; this visit will form the Week 32 visit (Visit 9). The assessments and procedures specified in [Table 1](#) will be performed, including querying for SAEs, AEs, and concomitant medications and treatments. All AEs and SAEs that are not resolved at the time of this visit will be followed to closure ([Section 8.1](#)).

Subjects who are proceeding to the maintenance or LTS studies will not enter the safety follow-up period.

### **7.1.4 Additional Care of Subjects after the Study**

No aftercare is planned for this study.

## **7.2 Study Evaluations and Procedures**

Every effort should be made to ensure that the protocol-required tests and procedures are completed as described. However, it is anticipated that from time to time there may be circumstances, outside the control of the investigator, which may make it unfeasible to perform the tests and procedures. In these cases, the investigator will take all steps necessary to ensure the safety and wellbeing of the subject.

When timing of procedures and assessments coincide, the following order should be followed:

- Health outcome and patient-reported questionnaires
- Vital signs and ECG
- Laboratory sample collection
- Investigational product administration
- Colonoscopy is performed at a separate visit ([Section 7.2.2.4](#)).

Note: Blood and tissue samples may be stored for up to, but for no longer than, 25 years.

## 7.2.1 Demographic and Other Baseline Characteristics

Demographic characteristics will be recorded at screening (Visit 1).

## 7.2.2 Efficacy

### 7.2.2.1 Patient-reported Outcome – Crohn’s Disease Daily E-diary

Patient-reported CD clinical signs and symptom data will be collected daily using a PRO-CD daily e-diary (electronic handheld device) starting from the first screening visit; however, collection of the daily e-diary data must begin at least 14 days before colonoscopy preparation. Subjects will enter data on CD signs and symptoms items using the e-diary, which will be provided to subjects at the start of the study. Compliance will be assessed by site staff at each visit. The site staff will instruct the subject on the appropriate use of the e-diary when compliance is below 80% (eg, <23 out of 28 e-diary entries). If 70% compliance cannot be achieved after repeated instructions during the screening period, noncompliant subjects will be automatically noneligible as they will not fulfill inclusion criterion 1 (Section 4.1).

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

- Abdominal pain severity (numeric rating scale [NRS])
- Very soft stool/liquid stool frequency (as shown by BSFS type 6/7)
- Total stool frequency
- Rectal bleeding frequency
- Rectal urgency frequency
- Nausea severity
- Vomiting frequency
- Incontinence frequency
- Abdominal pain used in CDAI
- General wellbeing.

The first 2 items (abdominal pain severity and very soft stool/liquid stool frequency) will be used to calculate the 2-item PRO. The 2-item PRO will be calculated using the following criteria:

- Screening: the 2-item PRO will be calculated based on the 7 most recent days during the 10 days of data collection before the colonoscopy preparation. If 7 out of the 10 most recent days are not available, then the 2-item PRO cannot be calculated for the subject at screening.
- Visits 3, 4, 5, and 6: the 2-item PRO will be calculated based on the 7 most recent days during the 10 days of data collection before the visit. If fewer than 7 days are available, the 2-item PRO will be calculated on all available most recent 6 or 5 days. If fewer than 5 days are available, the 2-item PRO will be treated as missing.

- Visit 7 (Part 3): the 2-item PRO will be calculated based on the 7 most recent days during the 10 days of data collection before the colonoscopy preparation. If fewer than 7 days are available, the 2-item PRO will be calculated on all available most recent 6 or 5 days. If fewer than 5 days are available, the 2-item PRO will be treated as missing.

For all 2-item PRO calculations, the 7 most recent days may or may not be contiguous during the 10 days of data collection depending on days to be excluded because of missing data.

The PRO-CD daily e-diary is presented in [Appendix 2](#).

### 7.2.2.2 Simple Endoscopic Score for Crohn's Disease

The SES-CD will be performed at the time points specified in [Table 1](#). The SES-CD score at baseline (Visit 2) and at Week 16/ET will be calculated using subscores of each of the segments investigated and centrally read from the colonoscopies performed at screening (Visit 1, Part 2) and Week 16 (Visit 7, Part 2), respectively.

The SES-CD is a simple scoring system based on 4 endoscopic variables (presence and size of ulcers, proportion of surface covered by ulcers, proportion of affected surface, and presence and severity of stenosis [narrowing]) measured in the same 5 ileocolonic segments as the CD index of severity. Overall, values on the SES-CD range from 0 to 56, with higher values indicating more severe disease. The 4 endoscopic variables are scored from 0 to 3 in each bowel segment (ileum, right/transverse/left colon, and rectum):

- Presence and size of ulcers (none = score 0; diameter 0.1–0.5 cm = score 1; 0.5–2 cm = score 2; diameter >2 cm = score 3)
- Extent of ulcerated surface (none = 0; <10% = 1; 10%–30% = 2; >30% = 3); extent of affected surface (none = 0; <50% = 1; 50%–75% = 2; >75% = 3)
- Presence and type of narrowings (none = 0; single, can be passed = 1; multiple, can be passed = 2; cannot be passed = 3).

A complete colonoscopy is required (including visualization of the terminal ileum).

The maximum stenosis score in a segment distal to another evaluable segment cannot exceed 2, so that the stenosis scores cannot exceed a total of 11 ([Reinisch et al., 2017](#)).

Evidence of active inflammation and ulceration is required at screening (Visit 1), in the form of a centrally read score of at least 1 in one or more ileocolonic segments in the Presence of Ulcers component of the SES-CD, as well as a total score of >6.

Study videos will be scored separately by 2 central readers who are blinded to the treatment. If the central readers' scores are not in agreement, there will be a third adjudication read to select the correct read from the first 2 scores. Results of the central reading of the videos will be communicated to sites within 5 business days.

For the evaluation of efficacy, in cases where 1 or 2 segments cannot be fully evaluated by central endoscopic readers, ileocolonic segments that are evaluable during screening (Visit 1) and Week 16/ET (matching segments approach) will be utilized.

The SES-CD is presented in [Appendix 2](#).

Screening laboratory tests, if considered by the investigator to be transient and inconsistent with the subject's clinical condition, may be repeated once during the screening period for confirmation. The following clinical laboratory assessments will be performed at the time points specified in [Table 1](#).

### **Serum chemistry**

- alkaline phosphatase
- aspartate aminotransferase
- alanine aminotransferase
- 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

### **Virology**

- hepatitis B surface antigen (HBsAg)
- hepatitis B core antibody (HBcAb)
- hepatitis B surface antibody (also referred as anti-HBsAg) if HBsAg is negative and HBcAb is positive
- hepatitis C virus antibody (HCVAb)
- HCV ribonucleic acid polymerase chain reaction if HCVAAb is positive

### **Urinalysis**

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

It is important to note that PRO assessments are collected and evaluated in a different manner than observed or volunteered AEs. Given these differences, no attempt will be made to reconcile any apparent discrepancies between observed or volunteered AEs and PRO data collected from subjects. Adverse event incidence rates will not be calculated from these solicited data but rather from the information recorded by the investigator.

### Inflammatory Bowel Disease Questionnaire

The IBDQ is a psychometrically validated PRO instrument for measuring the disease-specific HRQL in subjects with inflammatory bowel disease, including CD. The IBDQ consists of 32 items, which are grouped into 4 domains: bowel, systemic, emotional, and social function (Irvine et al., 1994). The 4 domains are scored as follows:

- Bowel function: 10 to 70
- Systemic function: 5 to 35
- Emotional function: 12 to 84
- Social function: 5 to 35.

The total IBDQ score ranges from 32 to 224. For the total score and each domain, a higher score indicates better HRQL. A score of at least 170 corresponds to clinical remission and an increase of at least 16 points is considered to indicate a clinically meaningful improvement.

The IBDQ is presented in [Appendix 2](#).

[REDACTED]

The [REDACTED] is presented in [Appendix 2](#).

[REDACTED]

The [REDACTED] and [REDACTED] are presented in [Appendix 2](#).

[REDACTED]

**Table 4: Volume of Blood to Be Drawn From Each Subject**

Assessment	Sample Volume (mL)	Number of Samples	Total Volume (mL)
[REDACTED]	4	3	12
[REDACTED]	3	6	18
ADA and NAb sampling	3	7	21
<b>Total (mL)</b>	<b>144.5</b>		

Ab=antibody; ADA=antidrug antibodies;  $\beta$ -hCG=beta-human chorionic gonadotropin; [REDACTED]; FSH=follicle-stimulating hormone; HBsAg=hepatitis B surface antigen; HBcAb=hepatitis B core antibody; HCV=hepatitis C virus; JCV=John Cunningham virus; [REDACTED]; NAb=neutralizing antibody; PPD=purified protein derivative; TB=tuberculosis.

<sup>a</sup>  $\beta$ -hCG testing for female subjects only.

<sup>c</sup> If a catheter is used, the first mL is to be discarded; then take 4 mL into appropriate tube for [REDACTED]. A total of 5 mL of blood drawn has been used in determination of sample volume.

The amount of blood to be drawn for each assessment is an estimate. The amount of blood to be drawn may vary according to the instructions provided by the manufacturer or laboratory for an individual assessment; however, the total volume drawn over the course of the study should be approximately 144.5 mL. When more than 1 blood assessment is to be done at the time point/period, if they require the same type of tube, the assessments may be combined.

### 8.1.2 Relationship Categorization

A physician/investigator must make the assessment of relationship to investigational product for each AE. The investigator should decide whether, in his or her medical judgment, there is a reasonable possibility that the event may have been caused by the investigational product. If there is no valid reason for suggesting a relationship, then the AE should be classified as “not related.” Otherwise, if there is any valid reason, even if undetermined or untested, for suspecting a possible cause-and-effect relationship between the investigational product and the occurrence of the AE, then the AE should be considered “related.” The causality assessment must be documented in the source document.

The following additional guidance may be helpful:

Term	Relationship Definition
Related	The temporal relationship between the event and the administration of the investigational product is compelling and/or follows a known or suspected response pattern to that product, and the event cannot be explained by the subject's medical condition, other therapies, or accident.
Not related	The event can be readily explained by other factors such as the subject's underlying medical condition, concomitant therapy, or accident and no plausible temporal or biologic relationship exists between the investigational product and the event.

### 8.1.3 Outcome Categorization

The outcome of AEs must be recorded during the course of the study in the eCRF. Outcomes are as follows:

- Fatal
- Not recovered/Not resolved
- Recovered/Resolved
- Recovered/Resolved with sequelae
- Recovering/Resolving
- Unknown.

### 8.1.4 Symptoms of the Disease Under Study

Symptoms of the disease under study should not be classed as AEs as long as they are within the normal day-to-day fluctuation or expected progression of the disease and are part of the efficacy data to be collected in the study; however, significant worsening of the symptoms should be recorded as an AE if the investigator considers such symptoms of disease progression related to the investigational product.

### **8.1.5 Clinical Laboratory and Other Safety Evaluations**

A change in the value of a clinical laboratory, vital sign, or ECG assessment can represent an AE if the change is clinically relevant or if, during treatment with the investigational product, a shift of a parameter is observed from a normal value to an abnormal value, or a further worsening of an already abnormal value. When evaluating such changes, the extent of deviation from the reference range, the duration until return to the reference range, either while continuing treatment or after the end of treatment with the investigational product, and the range of variation of the respective parameter within its reference range, must be taken into consideration.

If, at the end of the treatment phase, there are abnormal clinical laboratory, vital sign, or ECG values which were not present at the pretreatment value observed closest to the start of study treatment, further investigations should be performed until the values return to within the reference range or until a plausible explanation (eg, concomitant disease) is found for the abnormal values.

The investigator should decide, based on the above criteria and the clinical condition of a subject, whether a change in a clinical laboratory, vital sign, or ECG parameter is clinically significant and therefore represents an AE.

### **8.1.6 Pregnancy**

All pregnancies are to be reported from the time informed consent is signed until the defined follow-up period stated in Section [7.1.3](#).

Any report of pregnancy for any female study participant or the partner of a male study participant must be reported within 24 hours to Shire Global Drug Safety using the Shire Investigational and Marketed Products Pregnancy Report Form. A copy of the Shire Investigational and Marketed Products Pregnancy Report Form (and any applicable follow-up reports) must also be sent to the CRO/Shire medical monitor using the details specified in the [emergency contact information](#) section of the protocol. The pregnant female study participant must be withdrawn from the study.

Every effort should be made to gather information regarding the pregnancy outcome and condition of the infant. It is the responsibility of the investigator to obtain this information within 30 calendar days after the initial notification and approximately 30 calendar days postpartum.

Pregnancy complications such as spontaneous abortion/miscarriage or congenital abnormality are considered SAEs and must be reported using the Shire Clinical Study Serious Adverse Event and Nonserious AEs Required by the Protocol Form. Note: An elective abortion is not considered an SAE.

In addition to the above, if the investigator determines that the pregnancy meets serious criteria, it must be reported as an SAE using the Shire Clinical Study Serious Adverse Event and Nonserious AEs Required by the Protocol Form as well as the Shire Investigational and Marketed Products Pregnancy Report Form. The test date of the first positive serum/urine  $\beta$ -hCG test or ultrasound result will determine the pregnancy onset date.

### 8.1.7 Abuse, Misuse, Overdose, and Medication Error

Abuse, misuse, overdose, or medication error (as defined below) must be reported to the sponsor according to the SAE reporting procedure whether or not they result in an AE/SAE as described in Section 8.2. Note: The 24-hour reporting requirement for SAEs does not apply to reports of abuse, misuse, overdose, or medication errors unless these result in an SAE.

The categories below are not mutually exclusive; the event can meet more than 1 category.

- **Abuse** – Persistent or sporadic intentional intake of investigational product when used for a nonmedical purpose (eg, to alter one's state of consciousness or get high) in a manner that may be detrimental to the individual and/or society
- **Misuse** – Intentional use of investigational product other than as directed or indicated at any dose (Note: This includes a situation where the investigational product is not used as directed at the dose prescribed by the protocol)
- **Overdose** – Intentional or unintentional administration of investigational product at a dose interval that is less than 2 weeks between doses
- **Medication Error** – An error made in prescribing, dispensing, administration, and/or use of an investigational product. For studies, medication errors are reportable to the sponsor only as defined below.

Cases of subjects missing doses of the investigational product are not considered reportable as medication errors.

Medication errors should be collected/reported for all products under investigation.

The administration and/or use of the unassigned treatment is/are always reportable as a medication error.

The administration and/or use of an expired investigational product should be considered as a reportable medication error.

There is no specific antidote for overdose with SHP647. Treatment should be symptomatic and supportive.

### 8.1.8 Unexpected Adverse Event

An unexpected AE is an AE whose nature, severity, specificity, or outcome is not consistent with the term, representation, or description used in the reference safety information (RSI).

“Unexpected” also refers to the AEs that are mentioned in the IB as occurring with a class of drugs or as anticipated from the pharmacological properties of the product, but are not specifically mentioned as occurring with the particular product under investigation.

The expectedness of AEs will be determined by the sponsor using the IB as the RSI. This determination will include considerations such as the number of AEs previously observed, but not on the basis of what might be anticipated from the pharmacological properties of a product.

### **8.1.9 Suspected Unexpected Serious Adverse Reaction**

A suspected unexpected serious adverse reaction (SUSAR) is defined as any suspected adverse reaction to study treatment (ie, including active comparators) that is both serious and unexpected.

The event(s) must meet all of the following:

- Suspected adverse reaction
- Serious
- Unexpected
- Assessed as related to study treatment.

## **8.2 Serious Adverse Event Procedures**

### **8.2.1 Reference Safety Information**

The reference for safety information for this study is Section [6.8](#) of the SHP647 IB, which the sponsor has provided under separate cover to all investigators.

### **8.2.2 Reporting Procedures**

All initial and follow-up SAE reports must be reported by the investigator to Shire Global Drug Safety and the CRO/Shire medical monitor within 24 hours of the first awareness of the event. Note: The 24-hour reporting requirement for SAEs does not apply to reports of abuse, misuse, overdose, or medication errors (Section [8.1.7](#)) unless they result in an SAE.

The investigator must complete, sign, and date the Shire Clinical Study Serious Adverse Event and Nonserious AEs Required by the Protocol Form and verify the accuracy of the information recorded on the form with the corresponding source documents (Note: Source documents are not to be sent unless requested) and fax or email the form to Shire Global Drug Safety. A copy of the Shire Clinical Study Serious Adverse Event and Nonserious AEs Required by the Protocol Form (and any applicable follow-up reports) must also be sent to the CRO/Shire medical monitor using the details specified in the [emergency contact information](#) section of the protocol.

### **8.2.3 Serious Adverse Event Definition**

An SAE is any untoward medical occurrence (whether considered to be related to investigational product or not) that at any dose:

- Results in death
- Is life-threatening. Note: The term “life-threatening” in the definition of “serious” refers to an event in which the subject was at risk of death at the time of the event; it does not refer to an event which hypothetically might have caused death if it was more severe.
- Requires inpatient hospitalization or prolongation of existing hospitalization.  
Note: Hospitalizations, which are the result of elective or previously scheduled surgery for pre-existing conditions, which have not worsened after initiation of treatment, should not be classified as SAEs.

For example, an admission for a previously scheduled ventral hernia repair would not be classified as an SAE; however, complication(s) resulting from a hospitalization for an elective or previously scheduled surgery that meet(s) serious criteria must be reported as SAE(s).

- Results in persistent or significant disability/incapacity
- Is a congenital abnormality/birth defect
- Is an important medical event. Note: Important medical events that may not result in death, be life-threatening, or require hospitalization may be considered an SAE when, based upon appropriate medical judgment, they may jeopardize the subject and may require medical or surgical intervention to prevent one of the outcomes listed in this definition. Examples of such medical events include allergic bronchospasm requiring intensive treatment in an ED or at home; blood dyscrasias or convulsions that do not result in inpatient hospitalization; or the development of drug dependency or drug abuse.

#### **8.2.4 Serious Adverse Event Collection Time Frame**

All SAEs (regardless of relationship to study) are collected from the time the subject signs the informed consent until the defined follow-up period stated in Section 7.1.3 and must be reported to Shire Global Drug Safety and the CRO/Shire medical monitor within 24 hours of the first awareness of the event.

In addition, any SAE(s) considered “related” to the investigational product and discovered by the investigator at any interval after the study has completed must be reported to Shire Global Drug Safety within 24 hours of the first awareness of the event.

#### **8.2.5 Serious Adverse Event Onset and Resolution Dates**

The onset date of the SAE is defined as the date the event meets serious criteria. The resolution date is the date the event no longer meets serious criteria, the date the symptoms resolve, or the date the event is considered chronic. In the case of hospitalizations, the hospital admission and discharge dates are considered the onset and resolution dates, respectively.

In addition, any signs or symptoms experienced by the subject after signing the informed consent form, or leading up to the onset date of the SAE, or following the resolution date of the SAE, must be recorded as an AE, if appropriate.

#### **8.2.6 Fatal Outcome**

Any SAE that results in the subject’s death (ie, the SAE was noted as the primary cause of death) must have fatal checked as an outcome with the date of death recorded as the resolution date. For all other events ongoing at the time of death that did not contribute to the subject’s death, the outcome should be considered not resolved, without a resolution date recorded.

For any SAE that results in the subject’s death or any ongoing events at the time of death, unless another investigational product action was previously taken (eg, drug interrupted, reduced, withdrawn), the action taken with the investigational product should be recorded as “dose not changed” or “not applicable” (if the subject never received investigational product).

The investigational product action of “withdrawn” should not be selected solely as a result of the subject’s death.

**8.2.7 Regulatory Agency, Institutional Review Board, Ethics Committee, and Site Reporting**

The sponsor or the CRO is responsible for notifying the relevant regulatory authorities, US central Institutional Review Boards (IRBs), and European Union (EU) central Ethics Committees (ECs) of related, unexpected SAEs (ie, SUSARs).

In addition, the CRO is responsible for notifying active sites of all related, unexpected SAEs (ie, SUSARs) occurring during all interventional studies across the SHP647 program.

The investigator is responsible for notifying the local IRB, local EC, or the relevant local regulatory authority of all SAEs that occur at his or her site as required.

Further details regarding the DMC can be found in the DMC charter, which will be available before the administration of investigational product to any subject. Analyses of the data for DMC review will be conducted according to the DMC charter and DMC SAP. Because no formal hypothesis testing for safety assessments is planned, multiplicity concerns regarding repeated analyses are not applicable.

## 9.6 Sample Size Calculation and Power Considerations

Graphical methods are used to control the global family-wise type I error rate (FWER) at the .05 level (2-sided) for the comparisons of the 2 SHP647 treatment groups with the placebo group. Alpha is initially split equally at the .025 level (2-sided) for each of the pairwise treatment comparisons for the coprimary endpoints. Therefore, the power analysis and sample size estimation was calculated based on the chi-square test of proportions using nQuery Advisor® Version 7.0 (Statistical Solutions Ltd, Cork, Ireland) for an individual SHP647 dose compared to placebo.

Power calculations are made based on assuming a .025 (2-sided) significance level for each pairwise treatment comparison, 1720 subjects will be screened to randomize 1032 subjects in a 3:3:2 allocation ratio: 387 subjects in the 25 mg SHP647 treatment group, 387 subjects in the 75 mg SHP647 treatment group, and 258 subjects in the placebo group. These numbers of subjects would yield an approximately 93% power to detect individual pairwise treatment difference in the first coprimary efficacy endpoint, clinical remission by 2-item PRO at Week 16, of 10% (17.5% SHP647 versus 7.5% placebo). Expected clinical remission rates by 2-item PRO at Week 16 are based on observed rates from a post hoc analysis of the A7281006 study and placebo remission rates from literature ([Sandborn et al., 2017](#)). No adjustment for missing data is required in these sample size calculations as subjects with missing data for clinical remission by 2-item PRO at Week 16 are imputed as failures and the above rates account for these subjects.

With the 1032 subjects in allocation noted above, this number of subjects would yield an approximately 94% power to detect individual pairwise treatment difference in the other coprimary efficacy endpoint, endoscopic response at Week 16, of 12.5% (27.5% SHP647 versus 15% placebo). Expected endoscopic response rates at Week 16 are based on observed rates from a post hoc analysis of the A7281006 study and also endoscopic response rates from literature ([Sandborn et al., 2017](#)). No adjustment for missing data is required in these sample size calculations as subjects with missing data for endoscopic response at Week 16 are imputed as failures and the above rates account for these subjects.

The overall power for the coprimary endpoints will be approximately 87% assuming no correlation between the tests on the endpoints and approximately 90% assuming a correlation of 0.4.

With the sample size of 1032 subjects, [Table 5](#) provides the power for detecting a treatment difference between a SHP647 treatment group and the placebo group for the key secondary endpoints.

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

Sensitivity analyses to explore the impact of missing data on the coprimary efficacy endpoints and key secondary endpoints will be conducted. These analyses may compare imputations of the missing values which favor placebo (eg, worst case) and/or imputations which favor active treatment (eg, best case). In addition, imputation methods based on informative missingness and other missing data mechanisms may be performed. Additional sensitivity analyses will also be conducted using the per-protocol set and the completer set. Additional analyses may be developed in the SAP. All sensitivity analyses will be described in the SAP.

Prespecified subgroup analyses are planned for the coprimary endpoints including, but not limited to gender, prior anti-TNF treatment, glucocorticoid use at baseline, SES-CD at baseline, region, age group, randomization stratum, and other important subgroups. A full list of important subgroups will be described in the SAP. Within subgroups, efficacy endpoints will be compared for each active treatment group (25 mg SHP647 and 75 mg SHP647) with the placebo group using a chi-square test. The estimate of the treatment difference, along with the corresponding Newcombe (hybrid-score) 95% CI and chi-square test p-value, will be presented.

#### Statistical Testing and Protection of the Type I Error

The global FWER for the statistical tests of the coprimary and key secondary endpoints will be strongly controlled at .05 (2-sided). To control the FWER, graphical methods discussed in Bretz et al. (2009b) will be utilized to propagate  $\alpha$  from the coprimary endpoints to the key secondary endpoints and between the 2 SHP647 treatment group and placebo comparisons. Alpha is initially split equally at the .025 level (2-sided) for each of the pairwise treatment comparisons for the coprimary endpoints (P) and alpha is propagated in a hierarchical manner to each of the 6 key secondary endpoints (K1–K6) within a pairwise treatment comparison. In order to pass alpha between the coprimary endpoints and the first key secondary endpoint, both coprimary endpoints must attain statistical significance. A graphical visualization of the  $\alpha$  propagation is presented in [Figure 4](#).

- Clinical remission with endoscopic response, ie, both clinical remission by 2-item PRO and endoscopic response, as measured by a decrease in SES-CD of at least 25% at Week 16 (composite endpoint)
- Complete endoscopic healing at Week 16 defined as SES-CD=0-2.

Similar to the coprimary endpoints, the 6 key secondary endpoints will all be tested by the following hypothesis:

$$H_0: \delta = 0$$

$$H_1: \delta \neq 0$$

The key secondary endpoints will be analyzed using the same approach as described for the coprimary endpoints. Subjects with missing key secondary endpoint data at the Week 16 visit will be considered failures and counted as nonresponders.

### 9.8.2.2 Other Secondary Efficacy Endpoints

The other secondary endpoints are as follows:

- Clinical response at the Week 16 visit as measured by at least a 100-point reduction in the CDAI from baseline (CDAI-100 response)
- Clinical response at the Week 16 visit as measured by at least a 70-point reduction in the CDAI from baseline (CDAI-70 response)
- Clinical remission over time, as measured by the 2-item PRO
- Change from baseline in total stool frequency, rectal bleeding frequency, rectal urgency frequency, nausea severity, vomiting frequency, and rectal incontinence frequency scores; and total sign/symptom score based on subject daily e-diary entries
- Endoscopic healing at Week 16 as measured by SES-CD  $\leq 4$  and at least 2-point reduction versus baseline and no subscore  $> 1$  in any individual variable
- Change from baseline in IBDQ domain and total (absolute) scores over time
- Change from baseline in SF-36, version 2, acute (physical and mental component summary scores and individual domain scores) over time
- Incidence of all cause hospitalizations and total inpatient days
- Incidence of CD-related surgeries and other surgical procedures during the entire study period.

Other secondary endpoints will be summarized by descriptive statistics and presented by treatment group. Where appropriate, other secondary efficacy endpoints will be analyzed with the following analysis methods:

- Binary endpoints will be compared between each active treatment group and the placebo group using a CMH chi-square test stratified by status of prior anti-TNF treatment, glucocorticoid use at baseline, and SES-CD at baseline.

The estimate of the common treatment difference along with the corresponding stratified Newcombe 95% CI using the method of Yan and Su (2010) and the p-value computed from CHM test will be provided. Subjects with missing binary endpoint data at the Week 16 visit will be considered failures and counted as nonresponders.

- Continuous endpoints that are only measured at baseline and the Week 16 visit will be analyzed using an analysis of covariance model with fixed effects for treatment group (categorical), status of prior anti-TNF treatment (categorical), glucocorticoid use at baseline (categorical), and SES-CD at baseline (categorical), and the baseline value as a continuous covariate. From this model, estimates of the least squares means, treatment differences, standard errors, p-values, and 95% CIs for least squares mean treatment differences will be provided.
- Continuous endpoints that are measured repeatedly over time will be analyzed using a linear repeated measures mixed model with restricted maximum likelihood estimation. The model will include fixed effects for treatment group (categorical), visit (categorical), treatment group by visit interaction, status of prior anti-TNF treatment (categorical), glucocorticoid use at baseline (categorical), and SES-CD at baseline (categorical); baseline value as a continuous covariate; and repeated measures across visit for subject. From this model, estimates of least squares means, treatment differences, standard errors, p-values, and 95% CIs for least squares mean treatment differences for each visit will be provided.

### 9.8.3 Exploratory Efficacy Endpoints

The exploratory endpoints are as follows:



Exploratory efficacy endpoints will be summarized with descriptive statistics and presented by treatment group using the same approach described as for the other secondary endpoints. See Section 9.8.2.2 for an overview of the planned analyses. Full details for the analysis of exploratory efficacy endpoints will be included in the SAP.

## 9.9 Safety Analyses

All safety analyses will be performed using the safety set. Subjects will be analyzed according to the treatment they actually received.

Adverse events will be coded using the Medical Dictionary for Regulatory Activities.

Treatment-emergent AEs are defined as AEs with start dates at the time of or following the first exposure to investigational product. The number of events and percentage of TEAEs will be calculated by system organ class, by preferred term, and by treatment group.

Treatment-emergent AEs will be further summarized by severity and relationship to investigational product. Adverse events leading to withdrawal, serious AEs, and deaths will be similarly summarized or listed.

Clinical laboratory tests, vital signs, and ECG findings will be summarized by treatment group and visit. Potentially clinically important findings will also be summarized or listed.

Antidrug antibody data will be summarized by treatment group and visit.

Further details of safety analyses will be described in the SAP.

## 9.10 Other Analyses

### Simple Endoscopic Score for Crohn's Disease (SES-CD)

#### SITE WORKSHEET: Simple Endoscopic Score for Crohn's Disease (SES-CD)

Site #	Investigator	Subject ID	Visit Date (dd mmm yyyy)

#### Definitions of Simple Endoscopic Score for Crohn's Disease

Score Variable	0	1	2	3
Size of ulcers	None	Aphthous ulcers (Ø 0.1 to 0.5 cm)	Large ulcers (Ø 0.5 to 2 cm)	Very large ulcers (Ø >2 cm)
Ulcerated surface	None	<10%	10–30%	>30%
Affected surface	Unaffected segment	<50%	50–75%	>75%
Presence of narrowing	None	Single, can be passed	Multiple, can be passed	Cannot be passed

	Ileum	Right colon	Transverse colon	Left colon	Rectum	Total
Presence and size of ulcers (0–3)						
Extent of ulcerated surface (0–3)						
Extent of affected surface (0–3)						
Presence and type of narrowing(s) (0–3)						
					SES-CD =	

4. You indicated you had X bowel movements in the past 24 hours. Of these, how many had blood, either in the stool, the toilet bowel, or on the toilet paper?

Enter number of bowel movements with blood: \_\_\_\_\_

5. You indicated you had X bowel movements in the past 24 hours. How many of these involved urgency (having to suddenly rush to the toilet to make it on time)?

Enter number of bowel movements with urgency: \_\_\_\_\_

6. Please rate your worst feeling of nausea (feeling sick to your stomach or like you might throw up) over the past 24 hours.

None

Mild

Moderate

Severe

7. How many vomiting episodes did you have in the past 24 hours? An episode includes one or multiple heaves (including dry heaves) in quick succession followed by a break in vomiting.

Enter number of vomiting episodes: \_\_\_\_\_

8. How many bowel incontinence episodes (losing control of your bowels before reaching the toilet) did you have in the past 24 hours?

Enter number of bowel incontinence episodes: \_\_\_\_\_

9. Please rate your abdominal pain over the past 24 hours.

None

Mild

Moderate

Severe

3. How often during the last 2 weeks have you felt frustrated, impatient, or restless? Please choose an option from:
1. ALL OF THE TIME
  2. MOST OF THE TIME
  3. A GOOD BIT OF THE TIME
  4. SOME OF THE TIME
  5. A LITTLE OF THE TIME
  6. HARDLY ANY OF THE TIME
  7. NONE OF THE TIME
4. How often during the last 2 weeks have you been unable to attend school or do your work because of your bowel problem? Please choose an option from:
1. ALL OF THE TIME
  2. MOST OF THE TIME
  3. A GOOD BIT OF THE TIME
  4. SOME OF THE TIME
  5. A LITTLE OF THE TIME
  6. HARDLY ANY OF THE TIME
  7. NONE OF THE TIME
5. How much of the time during the last 2 weeks have your bowel movements been loose? Please choose an option from:
1. ALL OF THE TIME
  2. MOST OF THE TIME
  3. A GOOD BIT OF THE TIME
  4. SOME OF THE TIME
  5. A LITTLE OF THE TIME
  6. HARDLY ANY OF THE TIME
  7. NONE OF THE TIME



For non-commercial use only

[REDACTED]

[REDACTED]  
[REDACTED]

[REDACTED]

[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]

[REDACTED]

[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]

[REDACTED]

[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]

[REDACTED]

[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]

For non-commercial use only

**Table 1: Schedule of Assessments**

Study Procedure	Screening <sup>a</sup>		Baseline	Treatment						Follow-up			
	Weeks -6 to -1		Week 0/ Day 1	Week 2	Week 4	Week 8	Week 12	Week 16/ET <sup>b</sup>			Week 24 <sup>c</sup>	Week 32 <sup>c</sup>	
Visit Number	1 (Part 1) <sup>a</sup>	1 (Part 2) <sup>a</sup>	2	3	4	5	6	7 (Part 1) <sup>b</sup>	7 (Part 2) <sup>b</sup>	7 (Part 3) <sup>b</sup>	8	9	
Study Day	<b>-42 to 0</b>		1	<b>14 ±3</b>	<b>28 ±3</b>	<b>56 ±3</b>	<b>84 ±3</b>				<b>112 ±3</b>	<b>168 ±7</b>	<b>224 ±7</b>
TB test (PPD or QuantiFERON-TB Gold Plus) <sup>q</sup>	X												
JCV antibody banked sample <sup>r</sup>			X										
			X					X	X				
			X					X	X				
			X					X	X				
			X	X	X	X	X	X					
ADA and NAb sampling			X	X	X	X	X	X				X	
<b>Endoscopic Procedure</b>													
Colonoscopy (including biopsy) <sup>s</sup>		X								X			
<b>CD Assessments</b>													
CDAI <sup>t</sup>		X			X	X	X				X		
PRO-CD daily e-diary data instruction	X												
PRO-CD daily e-diary data <sup>u</sup>	X	X	X	X	X	X	X	X	X	X			
SES-CD <sup>v</sup>			X								X		
<b>Health Assessment<sup>w</sup></b>													
IBDQ			X		X	X				X			
			X		X	X				X			
Hospitalizations, inpatient days, [REDACTED] (HRUA)					X	X	X			X		X	
					X	X	X				X		
			X		X	X	X				X		
			X								X		
SF-36, version 2, acute			X		X	X	X			X			
			X							X			

### 9.8.1 Coprimary Efficacy Endpoints

The coprimary efficacy endpoints are:

- Clinical remission at the Week 16 visit as defined by the following: 2-item PRO subscores of average worst daily abdominal pain  $\leq 3$  (based on 11-point NRS) over the 7 most recent days and average daily stool frequency  $\leq 2$  of type 6/7 (very soft stools/liquid stools) as shown in the BSFS over the 7 most recent days. The 7 most recent days may or may not be contiguous during the 10 days of data collection before colonoscopy preparation, depending on days to be excluded because of missing data. If fewer than 7 days are available, the endpoint will be calculated on all available most recent 6 or 5 days. If fewer than 5 days are available, the endpoint will be treated as missing.
- Endoscopic response at Week 16 as measured by a decrease in SES-CD of at least 25% from baseline.

The coprimary efficacy endpoints, clinical remission at the Week 16 visit and endoscopic response at the Week 16 visit, will each be compared for each active treatment group (25 mg or 75 mg SHP647) to the placebo group using a Cochran-Mantel-Haenszel (CMH) chi-square test stratified by status of prior anti-TNF treatment, glucocorticoid use, and SES-CD at baseline for each of the stages of the study (stage 1 includes subjects whose primary efficacy data are used in the interim analysis and stage 2 includes all other subjects. Note: classification of stage 1 and stage 2 is based on the time of randomization rather than the time of study completion or termination). Subjects with missing data at the Week 16 visit will be considered failures and counted as nonresponders.

Weighted inverse normal p-value combination methods are used to combine the p-values from stage 1 and stage 2 through the following formula:

$$C(p_1, p_2) = 1 - \Phi[w_1 \Phi^{-1}(1-p_1) + w_2 \Phi^{-1}(1-p_2)]$$

Where  $p_1, p_2$  are the p-values computed from the CMH chi-square test for each stage,  $w_i^2 = n_i/(n_1 + n_2)$ ,  $n_1$  and  $n_2$  are the preplanned stage-wise sample sizes that are fixed at the time of the interim analysis based on an original total sample size, and  $\Phi$  denotes the cumulative distribution function of the standard normal distribution ([Bretz et al., 2009a](#)). Given that there is no possibility of stopping early for efficacy, that any potential stopping for futility of either or both doses of SHP647 is nonbinding, and that weights are prespecified, the test statistic  $C(p_1, p_2)$  can be compared against the nominal alpha level to assess statistical significance ([Chang and Chow, 2008](#)).

The coprimary endpoints will each be tested by the following hypothesis:

$$H_0: \delta = 0$$

$$H_1: \delta \neq 0$$

Where  $\delta$  is the common treatment difference across strata. The common treatment difference is a weighted average of the stratum-specific treatment differences.

This definition is further supported by a study conducted in a similar condition (irritable bowel syndrome) that examined the minimal clinically important difference on the 11-point NRS for abdominal pain ([Spiegel et al., 2009](#)) as well as post hoc analyses of the Phase 2 data from the SHP647 program (Study A7281006, OPERA).

#### Rationale for Very Soft Stool/Liquid Stool Frequency

Diarrhea is the most common sign in the presentation of CD, affecting approximately 85% of patients with a diagnosis of CD. In the CDAI, the number of liquid or soft stools (each day for 7 days) is used with a multiplier of 2. The coprimary endpoints for clinical remission in studies SHP647-305 and SHP647-306 requires the use of a definition without any such multiplying factor and will use the BSFS for defining the very soft or liquid stools according to types 6 and 7, respectively. A retrospective study of PROs in CD based on data from randomized controlled studies using rifaximin and methotrexate showed that a mean daily stool frequency score of  $\leq 1.5$  had an area under the receiving operating characteristic curve of 0.79 ([Khanna et al., 2015](#)) and provided a potential cutoff for defining remission as measured by CDAI. In a recent study to select the attributes determining overall disease severity and to rank the importance of and to score these individual attributes for both CD and UC based on specialist opinion, a sample of at least 10 loose stools per week was considered as an attribute contributing to overall disease severity in CD ([Siegel et al., 2016](#)). Based on post hoc analyses of the Phase 2 data in the SHP647 program (Study A7281006, OPERA) and by choosing the population of subjects satisfying the moderate to severe CD inclusion criteria, various cutoffs were explored and a stool frequency  $\leq 2.0$  was found to be optimal in terms of treatment separation while still allowing for a reasonable threshold for remission. Based on these and other recent data that support this cutoff, an average daily stool frequency  $\leq 2$  of type 6/7 (very soft stools/liquid stools) has been chosen as the stool frequency criterion for clinical remission.

#### Endoscopic Response

Endoscopic response is defined in 2 ways:

- 25% reduction in SES-CD score (“endoscopic response”)
- 50% reduction in SES-CD score (“enhanced endoscopic response”).

“Endoscopic response” will be used as a coprimary endpoint and “enhanced endoscopic response” will be used as a key secondary endpoint in this study as these magnitudes of changes are likely to be clinically relevant. In the recent trial with upadacitinib, the magnitude of difference between the different doses of the active drug and placebo at Week 16 was reported to be 9-36% for endoscopic response and 10-20% for enhanced endoscopic response ([Sandborn et al., 2017](#)). In the post hoc analysis of the Phase 2 OPERA study with SHP647 the magnitude of difference between the active treatment arms and placebo showed a similar pattern (higher for the endoscopic response than for enhanced endoscopic response) ([D’Haens et al., 2018](#)).

Mucosal healing or “endoscopic healing” is considered to be a pivotal long-term target in the treatment of CD; however, partial healing or endoscopic response may also provide benefits. Endoscopic response can be an important indicator that the mucosal inflammation has decreased as an effect of the investigational product.

It is important to note that assessments based on neurological evaluations are collected and evaluated in a different manner than observed or volunteered AEs. Given these differences, no attempt will be made to reconcile any apparent discrepancies between observed or volunteered AEs and data from neurological assessment collected from subjects. Investigators may determine if any finding on neurological testing constitutes an AE. Adverse event incidence rates will not be calculated from these neurological evaluation data but rather from the AE information recorded by the investigator.

#### **7.2.3.4 Adverse Event Collection**

At each study visit, subjects will be questioned in a general way to ascertain if AEs have occurred since the previous visit (eg, “Have you had any health problems since your last visit?”). Adverse events are collected from the time informed consent and/or assent is signed until the end of the study or the defined follow-up period stated in Section 7.1.3 (Section 8, Adverse and Serious Adverse Events Assessment.)

#### **7.2.3.5 Vital Signs**

Vital signs will be measured at the time points specified in Table 1. Additional collection times or changes to collection times will be permitted, as necessary to ensure appropriate collection of safety data. Vital signs include blood pressure, pulse, respiratory rate, and temperature.

Single measurements of sitting blood pressure will be recorded at each time point. Blood pressure should be determined by cuff with the subject’s arm supported at the level of the heart and recorded to the nearest mm Hg using the same method, the same arm (preferably the dominant arm), and the same position throughout the study.

Respiratory rate will be measured with the subject in a comfortable position. The observer should hold the extremity of the subject as a distraction for the subject (ie, pretending he/she is taking the subject’s radial pulse) and count the respiration for 1 minute.

Body temperature should be taken using a thermometer and reported in degrees Celsius or Fahrenheit.

Any deviations from baseline (Visit 2) vital signs that are deemed clinically significant in the opinion of the investigator are to be recorded as an AE, unless documented in the subject’s medical history as a pre-existing medical condition.

#### **7.2.3.6 Clinical Laboratory Evaluations**

All clinical laboratory assays will be performed according to the central laboratory’s normal procedures. Reference ranges are to be supplied by the central laboratory and will be used to assess the clinical laboratory data for clinical significance and out-of-range pathological changes. The investigator should assess out-of-range clinical laboratory values for clinical significance, indicating if the value(s) is/are not clinically significant or clinically significant. Abnormal clinical laboratory values, which are unexpected or not explained by the subject’s clinical condition, may, at the discretion of the investigator or sponsor, be repeated as soon as possible until confirmed, explained, or resolved.

Subjects with inflammatory bowel disease (IBD) (including CD), with or without stoma, who failed or were intolerant to both anti-TNF and immunosuppressant therapy and who had moderate to severe active disease underwent a lumbar puncture (LP), completed induction therapy with 3 doses of 225 mg ontamalimab 4 weeks apart, and then underwent a second LP 2 ( $\pm 1$ ) weeks later. The primary endpoint was the percent change from baseline (pretreatment) in absolute lymphocyte count in cerebrospinal fluid (CSF) in subjects with IBD after receiving 3 doses of 225 mg ontamalimab. The mean percentage change from baseline in absolute lymphocytes in CSF was 61.76% with a median change of 35.2% (range: -70.2% to 267.8%). The post-treatment LP/pretreatment LP geometric mean ratio for CSF lymphocytes was 1.33 with the lower bound of the 80% confidence interval (CI)=1.13, which was greater than 0.5, supporting rejection of the null hypothesis (ie, that the percent decrease in total lymphocytes counts after treatment would be  $\geq 50\%$  (equivalent to the geometric mean ratio in total lymphocyte counts being  $\leq 0.5$ ). This result supports the hypothesis that ontamalimab does not impair trafficking of lymphocytes into the CNS and thus should not impair CNS immune surveillance.

Study A7281007 (OPERA II) was a Phase 2 open-label extension study to provide additional long-term safety data on subjects with moderate to severe CD who completed Study A7281006 or Study A7281008 and wished to continue to receive ontamalimab. Ontamalimab 75 mg (with potential dose escalation to 225 mg) SC given Q4W for 72 weeks was generally well tolerated in subjects with CD over the treatment period evaluated in this study. In subjects with positive ADA or neutralizing antibody (NAb) status, exposure to ontamalimab was not affected. Serum concentrations of ontamalimab in this study were consistent with what was predicted based on the Feeder Study A7281006. There were 2 deaths in the study: 1 subject died of multiple organ dysfunction syndromes in the treatment period and 1 subject died of metastatic neoplasm in the follow-up period. Neither death was reported as related to treatment with the investigational product by the investigators. The most frequently reported SAE was CD in either the treatment period or the follow-up period. The SOC with the most subjects experiencing TEAEs was GI disorders. Although Study A7281007 was not placebo controlled, the exploratory efficacy results (based on the modified Harvey Bradshaw Index) indicated that the effect of ontamalimab on disease activity was maintained over the duration of treatment.

The ontamalimab dose selection (25 mg and 75 mg) for this study is based on data from these 3 previous studies, which evaluated the activity of ontamalimab in adult patients with moderately to severely active CD based on CDAI scores between 220 and 450. The results of a post hoc analysis of remission rate by baseline elevated serum concentration of hsCRP suggested that the greatest treatment effect was at a dose of 22.5 mg. Similarly, post hoc analysis of remission rates by endoscopic severity assessed using the SES-CD suggested best efficacy at a dose of 75 mg. Therefore, both dosage regimens 25 mg and 75 mg Q4W have been selected for the Phase 3 testing. The Phase 1 study (A7281008, TOSCA) and Phase 2 studies (A7281006, OPERA; and A7281007, OPERA II [long-term safety study]) that investigated the safety, tolerance, PK, and pharmacodynamic (PD) properties of ontamalimab support further clinical development of ontamalimab using SC administration in subjects with moderate to severe CD.

22 Nov 2019

### **2.2.3 Exploratory Objectives**

The exploratory objectives are as follows:

Term	Percentage (%)
Climate change	100
Global warming	95
Green energy	90
Sustainable development	85
Environmental protection	80
Ecology	75

### 3. STUDY DESIGN

#### 3.1 Study Design and Flow Chart

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

A total of 1032 subjects (387 subjects in the 25 mg ontamalimab treatment group, 387 subjects in the 75 mg ontamalimab treatment group, and 258 subjects in the placebo group) are planned for enrollment into the study ([Figure 1](#)). Subjects must be at least 16 years of age and no more than 80 years of age at the time of signing the informed consent/assent form.

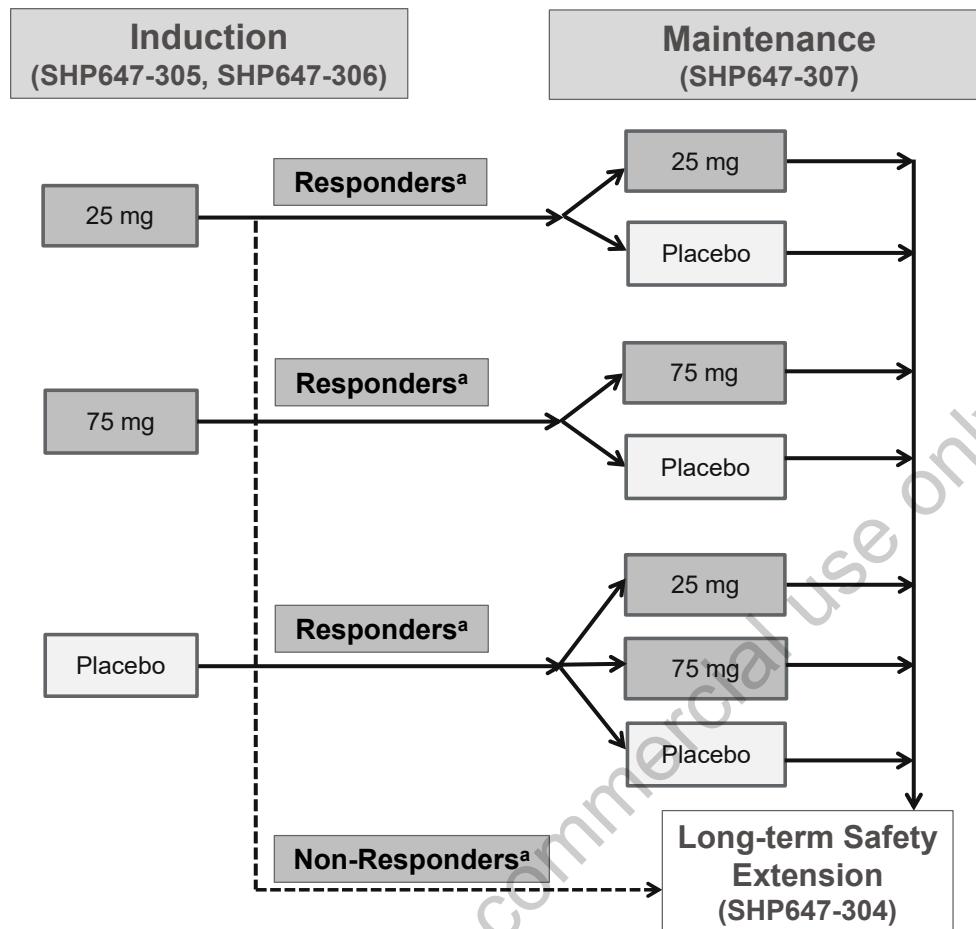
The study consists of a screening period up to 6 weeks and a 16-week treatment period. After the screening period, eligible subjects will be randomized to receive 1 of 3 treatments (25 mg ontamalimab, 75 mg ontamalimab, or placebo) in a 3:3:2 ratio. Randomization will be stratified based upon the subject's status of prior anti-TNF treatment (naïve or experienced), glucocorticoid use at baseline (on glucocorticoids at baseline versus not on glucocorticoids at baseline), and SES-CD at baseline (SES-CD  $\geq$ 17 or SES-CD <17). Subjects will receive SC injections of ontamalimab or placebo, using a prefilled syringe (PFS), on Week 0/Day 1 (Visit 2), Week 4 (Visit 4), Week 8 (Visit 5), and Week 12 (Visit 6). Subjects will undergo efficacy, [REDACTED], [REDACTED], safety, and health outcome assessments at these visits as detailed in [Table 1](#).

At the end of the 16-week treatment period, subjects will be offered the opportunity to participate in either a double-blind maintenance study (SHP647-307; for subjects who fulfill the entry criteria) or a long-term safety extension (LTS) study (SHP647-304; for subjects who do not fulfill the entry criteria for Study SHP647-307) as shown in [Figure 1](#). Subjects who withdraw early from the 16-week treatment period or who do not wish to enter the maintenance study (SHP647-307) or LTS study (SHP647-304) will continue into a 16-week safety follow-up period. Only those subjects who complete the full course of investigational product treatment in the induction studies (SHP647-305 or SHP647-306) will be eligible to continue in the maintenance study or LTS study.

A planned interim analysis for the coprimary endpoints will take place after approximately the first 50% of all randomized subjects in both the SHP647-305 and SHP647-306 studies have either completed the studies or have prematurely withdrawn from the studies. The sample size will be reassessed as part of this interim analysis. See Section [9.5](#) for further details of the planned interim analysis.

The overall study design is shown in [Figure 2](#).

**Figure 1 Overview of Ontamalimab Phase 3 Studies in Crohn's Disease**



BSFS=Bristol Stool Form Scale; CDAI= Crohn's Disease Activity Index; NRS=numerical rating scale; PRO=patient-reported outcome; SES-CD=Simple Endoscopic Score for Crohn's Disease

<sup>a</sup> Responders are subjects who either:

(a) Meet endoscopic response criteria of a reduction in SES-CD from baseline by  $\geq 25\%$  at Week 16  
OR

(b) Meet at least 1 of the following 4 criteria at Week 16 in addition to no worsening of endoscopic score as measured by SES-CD relative to induction study baseline (SHP647-305 or SHP647-306):

1. Subject is in clinical remission as determined by meeting the criteria for clinical remission using the 2-item PRO, ie, 2-item PRO subscore of average worst daily abdominal pain  $\leq 3$  (based on 11-point NRS) over the 7 most recent days and average daily stool frequency  $\leq 2$  of type 6/7 (very soft stools/liquid stools) as shown in the BSFS over the 7 most recent days.\*
2. Subject has a decrease of at least 100 points in CDAI score (CDAI-70) from baseline.
3. Subject has a decrease of  $\geq 30\%$  and at least 2 points from baseline in the average daily worst abdominal pain over the 7 most recent days\*, with the average daily stool frequency of type 6/7 (very soft stools/liquid stools) either: (i) not worsening from baseline and/or (ii) meeting the criteria for clinical remission, ie, 2-item PRO subscore of average daily stool frequency  $\leq 2$  of type 6/7 (very soft stools/liquid stools) as shown in the BSFS over the 7 most recent days.\*
4. Subject has a decrease of  $\geq 30\%$  from baseline in the average daily stool frequency of type 6/7 (very soft stools/liquid stools) as shown in the BSFS over the 7 most recent days\*, with the average daily worst abdominal pain either: (a) not worsening from baseline and/or (b) meeting the criteria for clinical remission, ie, 2-item PRO subscore of average worst daily abdominal pain  $\leq 3$  (based on 11-point NRS) over the 7 most recent days\*.

\*Note: The 7 most recent days may or may not be contiguous during the 10 days of data collection before colonoscopy preparation, depending on days to be excluded because of missing data. If fewer than 7 days are available, the criterion will be calculated on all available most recent 6 or 5 days. If fewer than 5 days are available, the criterion will be treated as missing.

condition (irritable bowel syndrome) that examined the minimal clinically important difference on the 11-point NRS for abdominal pain ([Spiegel et al., 2009](#)) as well as post hoc analyses of the Phase 2 data from the ontamalimab program (Study A7281006, OPERA).

### Rationale for Very Soft Stool/Liquid Stool Frequency

Diarrhea is the most common sign in the presentation of CD, affecting approximately 85% of patients with a diagnosis of CD. In the CDAI, the number of liquid or soft stools (each day for 7 days) is used with a multiplier of 2. The coprimary endpoints for clinical remission in studies SHP647-305 and SHP647-306 requires the use of a definition without any such multiplying factor and will use the BSFS for defining the very soft or liquid stools according to types 6 and 7, respectively. A retrospective study of PROs in CD based on data from randomized controlled studies using rifaximin and MTX showed that a mean daily stool frequency score of  $\leq 1.5$  had an area under the receiving operating characteristic curve of 0.79 ([Khanna et al., 2015](#)) and provided a potential cutoff for defining remission as measured by CDAI. In a recent study to select the attributes determining overall disease severity and to rank the importance of and to score these individual attributes for both CD and UC based on specialist opinion, a sample of at least 10 loose stools per week was considered as an attribute contributing to overall disease severity in CD ([Siegel et al., 2016](#)). Based on post hoc analyses of the Phase 2 data in the ontamalimab program (Study A7281006, OPERA) and by choosing the population of subjects satisfying the moderate to severe CD inclusion criteria, various cutoffs were explored and a stool frequency  $\leq 2.0$  was found to be optimal in terms of treatment separation while still allowing for a reasonable threshold for remission. Based on these and other recent data that support this cutoff, an average daily stool frequency  $\leq 2$  of type 6/7 (very soft stools/liquid stools) has been chosen as the stool frequency criterion for clinical remission.

### Endoscopic Response

Endoscopic response is defined in 2 ways:

- 25% reduction in SES-CD score (“endoscopic response”)
- 50% reduction in SES-CD score (“enhanced endoscopic response”).

“Endoscopic response” will be used as a coprimary endpoint and “enhanced endoscopic response” will be used as a key secondary endpoint in this study as these magnitudes of changes are likely to be clinically relevant. In the recent trial with upadacitinib, the magnitude of difference between the different doses of the active drug and placebo at Week 16 was reported to be 9% to 36% for endoscopic response and 10% to 20% for enhanced endoscopic response ([Sandborn et al., 2017](#)). In the post hoc analysis of the Phase 2 OPERA study with ontamalimab the magnitude of difference between the active treatment arms and placebo showed a similar pattern (higher for the endoscopic response than for enhanced endoscopic response) ([D'Haens et al., 2018](#)).

Mucosal healing or “endoscopic healing” is considered to be a pivotal long-term target in the treatment of CD; however, partial healing or endoscopic response may also provide benefits. Endoscopic response can be an important indicator that the mucosal inflammation has decreased as an effect of the investigational product. Some treatments may result in a partial initial

from literature suggesting that thresholds  $\leq 3$  are likely to be quite stringent (Sandborn et al., 2017), and the refractory nature of the disease in those with moderate to severe CD, it is also important to assess the effects of treatment using a more realistic measure. Hence, for this key secondary endpoint of clinical remission, average daily stool frequency  $\leq 3$  of type 6/7 (very soft stools/liquid stools) has been chosen as the appropriate threshold.

### Clinical Response

The goal of measuring clinical remission is to have a sensitive clinical measure to assess the complete absence of symptoms or the stabilization of noninflammatory symptoms. However, as response and remission are considered to be on a continuum of improvement or response to treatment, clinical remission is generally the chosen measurement over clinical response. Therefore, clinical remission is used as coprimary endpoint and clinical response is used as a key secondary endpoint in this study.

Clinical response is defined in Section 9.8.2.1. For both clinical response criteria, an additional requirement is that the symptom not being used to assess clinical response (ie, abdominal pain severity or very soft stool/liquid stool frequency) must remain unchanged/not worsen from the baseline score, or meet the criteria for clinical remission for that item of the 2-item PRO (either a 2-item PRO subscore of average daily stool frequency  $\leq 2$  of type 6/7 [very soft stools/liquid stools] as shown in the BSFS or average worst daily abdominal pain  $\leq 3$  [based on 11-point NRS] over the 7 most recent days).

The rationale for needing to meet at a minimum clinical response definition for either abdominal pain severity or very soft stool/liquid stool frequency (and not necessarily both) is based on the supposition that a lack of improvement in 1 of these symptoms is not necessarily an indicator of eventual lack of response (as assessed by the stricter clinical remission criterion). Based on Phase 2 data, it has been observed that the magnitude of placebo response rate can be higher for abdominal pain than for stool frequency. Therefore, the additional criterion of at least a 2-point decrease in abdominal pain severity from baseline is required for assessing clinical response for abdominal pain. Overall, the definition of clinical response used for this study has been chosen to allow for the maximal pool of subjects to be assessed for the effect of treatment and, if appropriate, the continuation of therapy in the maintenance study (SHP647-307).

### Composite Score Endpoint of Both Clinical Remission by 2-item PRO and Endoscopic Response at Week 16

In theory, as the degree of inflammation decreases due to the effect of treatment, both clinical signs and symptoms of CD as well as endoscopic appearance can improve. However, in any given subject, the rates of clinical improvement and endoscopic improvement may not be the same. There are reasons for this discrepancy when evaluating clinical and endoscopic improvement in the same time period, including clinical symptoms not being well correlated to the mucosal inflammation. Due to the transmural feature of the disease, symptoms can correspond to the inflammation in some of the other gut layers as well. Previous clinical studies and clinical observations indicate that the improvement of clinical signs and symptoms and the improvement in endoscopic appearance may not go hand in hand. Significant clinical improvement can precede significant endoscopic improvement. The healing process of the gut

## 5. PRIOR AND CONCOMITANT TREATMENT

### 5.1 Prior Treatment

Prior treatment includes all treatment (including but not limited to herbal remedies and vitamins) received within 30 days (or PK equivalent of 5 half-lives, whichever is longer) of the first dose of investigational product. Use of biologics for indications other than CD during the 90 days before screening must also be recorded.

Prior and concomitant CD-specific treatments from the previous 10 years will be recorded. The subject's entire history of biologic CD-specific treatments will be recorded.

Subjects must have had an inadequate response to, or lost response to, or had an intolerance to at least 1 conventional treatment such as sulfasalazine or 5-ASA, glucocorticoids, immunosuppressants (AZA, 6-MP, or MTX), or anti-TNF agents ([Lichtenstein et al., 2018](#)). Please see [Appendix 4](#) for guidance on defining prior treatment failure and intolerance to prior treatment for CD.

### 5.2 Concomitant Treatment

Concomitant treatment refers to all treatment taken between the dates of the first dose of investigational product and the end of the safety follow-up period of this study, inclusive.

#### 5.2.1 Permitted Treatment

Subjects must remain on stable doses of permitted CD treatments until completion of the Week 16 visit, unless decreases are required because of AEs. Stable doses of the following treatments for CD are permitted as concomitant medication:

- Oral sulfasalazine or 5-ASA, providing that the dose is stable for at least 2 weeks before baseline (Visit 2)
- Immunosuppressants (AZA, 6-MP, or MTX), providing that the dose is stable for at least 8 weeks before baseline (Visit 2)
- Oral glucocorticoids (prednisone or equivalent [see [Appendix 3](#)]) up to a maximum of 20 mg/day or budesonide up to a maximum of 9 mg/day), providing that the dose is stable for at least 2 weeks before baseline (Visit 2). After baseline (Visit 2), a stable dose of 20 mg/day prednisone or equivalent oral systemic corticosteroid dose is allowed. Steroids may be decreased due to AEs.

Note: Rectal 5-ASA and parenteral or rectal glucocorticoids are prohibited from within 14 days before screening colonoscopy.

should not proceed with the colonoscopy preparation and/or the colonoscopy. All colonoscopies will be evaluated using the SES-CD (see [Appendix 2](#)).

If a subject has had the following procedures performed as a part of standard medical care within 12 weeks before screening (Visit 1), these procedures do not need to be repeated as a part of screening:

- Chest x-ray
- Documented negative PPD test or IGRA for TB.

#### **7.1.1.2 Baseline Visit (Visit 2; Week 0)**

The baseline visit (Visit 2) will take place on Day 1 (Week 0). The assessments and procedures specified in [Table 1](#) will be performed.

After eligibility has been reconfirmed and all baseline procedures and assessments have been completed, each subject will be randomized to 1 of the 3 treatment groups as described in Section [6.2.2](#) and the first dose of investigational product will be administered.

Results of the baseline laboratory tests are not required for investigational product administration but must be reviewed as soon as possible thereafter.

#### **7.1.2 Treatment Period**

##### **7.1.2.1 Visits 3, 4, 5, and 6 (Weeks 2, 4, 8, and 12)**

Visits 3, 4, 5, and 6 are scheduled to take place on Day  $14 \pm 3$  days (Week 2), Day  $28 \pm 3$  days (Week 4), Day  $56 \pm 3$  days (Week 8), and Day  $84 \pm 3$  days (Week 12), respectively. The assessments and procedures specified in [Table 1](#) will be performed.

##### **7.1.2.2 Final On-treatment Visit: Visit 7, Parts 1, 2, and 3 (Week 16/Early Termination)**

The Week 16/ET visit (Visit 7) consists of 3 parts.

Part 1 of Visit 7 can either be done on the same day as Part 2 or be done up to 3 day(s) before Part 2. If Parts 1 and 2 are done on the same day, blood samples must be taken before starting the colonoscopy preparation. The Week 16/ET assessments and procedures that will take place during Part 1 are specified in [Table 1](#).

Part 2 of Visit 7 must be completed within 10 days (preferably within 5 to 7 days) before Part 3; this will allow sufficient time to obtain the data from the centrally read colonoscopy. The Week 16/ET assessments and procedures that will take place during Part 2 are specified in [Table 1](#).

Part 3 of Visit 7 will take place on Day  $112 \pm 3$  days. The Week 16/ET assessments and procedures that will take place during Part 3 are specified in [Table 1](#).

At Part 3 of Visit 7, after review of CD assessments, health outcome assessments, and safety assessments, it will be determined whether the subject should enroll in the maintenance (SHP647-307) or LTS (SHP647-304) studies or enter the follow-up period of this study. Entry into the maintenance or LTS studies is dependent upon whether the subject fulfills the efficacy entry criteria of the maintenance study (SHP647-307), including achieving endoscopic and/or clinical response, and whether the subject agrees to participate.

The Week 16 assessments and procedures will also form the ET assessments for any subjects who are withdrawn early or discontinued from the study.

### **7.1.3 Follow-up Period**

Subjects who are withdrawn early from the study, or who do not enter either the maintenance or LTS studies, should enter the 16-week safety follow-up period for safety monitoring.

During the safety follow-up period, the Week 24 visit (Visit 8) will take place on Day  $168 \pm 7$  days, or 8 weeks  $\pm 7$  days after the subject's last visit in the treatment period for subjects who are withdrawn early from the study. This visit will routinely be conducted by telephone; however, as an exception, the visit can be performed as a study site visit if preferred.

At the end of the safety follow-up period, there will be a visit at the site on Day  $224 \pm 7$  days, or 16 weeks  $\pm 7$  days after the subject's last visit in the treatment period for subjects who are withdrawn early from the study; this visit will form the Week 32 visit (Visit 9). The assessments and procedures specified in [Table 1](#) will be performed, including querying for SAEs, AEs, and concomitant medications and procedures. All AEs and SAEs that are not resolved at the time of this visit will be followed to closure (see [Section 8.1](#)).

Subjects who are proceeding to the maintenance or LTS studies will not enter the safety follow-up period.

### **7.1.4 Additional Care of Subjects After the Study**

No aftercare is planned for this study.

## **7.2 Study Evaluations and Procedures**

Every effort should be made to ensure that the protocol-required tests and procedures are completed as described. However, it is anticipated that from time to time there may be circumstances, outside the control of the investigator, that may make it unfeasible to perform the tests and procedures. In these cases, the investigator will take all steps necessary to ensure the safety and well-being of the subject.

When timing of procedures and assessments coincide, the following order should be followed:

- Health outcome and patient-reported questionnaires
- Vital signs and ECG
- Laboratory sample collection
- Investigational product administration
- Colonoscopy is performed at a separate visit (see Section 7.2.2.4).

Note: Blood and tissue samples may be stored for up to the duration allowed by local regulations, but for no longer than 25 years.

## 7.2.1 Demographic and Other Baseline Characteristics

Demographic characteristics will be recorded at screening (Visit 1).

## 7.2.2 Efficacy

### 7.2.2.1 Patient-reported Outcome – Crohn’s Disease Daily E-diary

Patient-reported CD clinical signs and symptom data will be collected daily using a PRO-CD daily e-diary (electronic handheld device) starting during the screening period; however, collection of the daily e-diary data must begin at least 10 days before colonoscopy preparation. Subjects will enter data on CD signs and symptoms items using the e-diary, which will be provided to subjects at the start of the study. Compliance will be assessed by site staff at each visit. The site staff will instruct the subject on the appropriate use of the e-diary when compliance is below 80% (eg, <23 out of 28 e-diary entries). If at least 70% compliance cannot be achieved after repeated instructions during the screening period, noncompliant subjects will be automatically noneligible as they will not fulfill inclusion criterion 1 (see Section 4.1).

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

- Abdominal pain severity (numeric rating scale [NRS])
- Very soft stool/liquid stool frequency (as shown by BSFS type 6/7)
- Total stool frequency
- Rectal bleeding frequency
- Rectal urgency frequency
- Nausea severity
- Vomiting frequency
- Incontinence frequency
- Abdominal pain used in CDAI
- General well-being.

The first 2 items (abdominal pain severity and very soft stool/liquid stool frequency) will be used to calculate the 2-item PRO. The 2-item PRO will be calculated using the following criteria:

- Screening: the 2-item PRO will be calculated based on the 7 most recent days during the 10 days of data collection before the colonoscopy preparation. If 7 out of the 10 most recent days are not available, then the 2-item PRO cannot be calculated for the subject at screening. Note that the subject must be confirmed as meeting the PRO subscore requirements at screening **before** a colonoscopy is done.
- Visits 3, 4, 5, and 6: the 2-item PRO will be calculated based on the 7 most recent days during the 10 days of data collection before the visit. If fewer than 7 days are available, the 2-item PRO will be calculated on all available most recent 6 or 5 days. If fewer than 5 days are available, the 2-item PRO will be treated as missing.
- Visit 7 (Part 3): the 2-item PRO will be calculated based on the 7 most recent days during the 10 days of data collection before the colonoscopy preparation. If fewer than 7 days are available, the 2-item PRO will be calculated on all available most recent 6 or 5 days. If fewer than 5 days are available, the 2-item PRO will be treated as missing.

For all 2-item PRO calculations, the 7 most recent days may or may not be contiguous during the 10 days of data collection depending on days to be excluded because of missing data.

The PRO-CD daily e-diary is presented in [Appendix 2](#).

#### 7.2.2.2 Simple Endoscopic Score for Crohn's Disease

The SES-CD will be performed at the time points specified in [Table 1](#). The SES-CD score at baseline (Visit 2) and at Week 16/ET will be calculated using subscores of each of the segments investigated and centrally read from the colonoscopies performed at screening (Visit 1, Part 2) and Week 16 (Visit 7, Part 2), respectively.

The SES-CD is a simple scoring system based on 4 endoscopic variables (presence and size of ulcers, proportion of surface covered by ulcers, proportion of affected surface, and presence and severity of stenosis [narrowing]) measured in the same 5 ileocolonic segments as the CD index of severity ([Daperno et al., 2004](#)). Overall, values on the SES-CD range from 0 to 56, with higher values indicating more severe disease. The 4 endoscopic variables are scored from 0 to 3 in each bowel segment (ileum, right/transverse/left colon, and rectum):

- Presence and size of ulcers (none = score 0; diameter 0.1–0.5 cm = score 1; 0.5–2 cm = score 2; diameter >2 cm = score 3)
- Extent of ulcerated surface (none = 0; <10% = 1; 10%–30% = 2; >30% = 3); extent of affected surface (none = 0; <50% = 1; 50%–75% = 2; >75% = 3)
- Presence and type of narrowings (none = 0; single, can be passed = 1; multiple, can be passed = 2; cannot be passed = 3).

A complete colonoscopy is required (including visualization of the terminal ileum except when it is not possible due to impassable stenosis or previous partial colectomy/ileocolectomy). The maximum stenosis score in a segment distal to another evaluable segment cannot exceed 2, so that the stenosis scores cannot exceed a total of 11 ([Reinisch et al., 2017](#)).

Respiratory rate will be measured with the subject in a comfortable position. The observer should hold the extremity of the subject as a distraction for the subject (ie, pretending he/she is taking the subject's radial pulse) and count the respiration for 1 minute.

Body temperature should be taken using a thermometer and reported in degrees Celsius or Fahrenheit.

Any deviations from baseline (Visit 2) vital signs that are deemed clinically significant in the opinion of the investigator are to be recorded as an AE, unless documented in the subject's medical history as a pre-existing medical condition.

#### 7.2.3.6 Clinical Laboratory Evaluations

All clinical laboratory assays will be performed according to the central laboratory's normal procedures. Reference ranges are to be supplied by the central laboratory and will be used to assess the clinical laboratory data for clinical significance and out-of-range pathological changes. The investigator should assess out-of-range clinical laboratory values for clinical significance, indicating if the value(s) is/are not clinically significant or clinically significant.

Abnormal clinical laboratory values that are unexpected or not explained by the subject's clinical condition may, at the discretion of the investigator or sponsor, be repeated as soon as possible until confirmed, explained, or resolved.

Screening laboratory tests, if considered by the investigator to be transient and inconsistent with the subject's clinical condition, may be repeated once during the screening period for confirmation. The following clinical laboratory assessments will be performed at the time points specified in [Table 1](#).

#### Serum chemistry

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

#### Hematology

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

## Virology

HBsAg	HCVAb
HBcAb	HCV RNA PCR if HCVAb is positive
hepatitis B DNA if HBsAg is negative and HBcAb is positive	HIV

## Urinalysis

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

Virology test results must be confirmed as negative before enrollment in the study; if a virology test result is positive, the subject will be excluded from entering the study. Results of the virology screen will be reviewed and verified by the study monitor but will not be collected in the electronic case report form (eCRF) database.

Stool microbiology will be performed at screening (Visit 1) or at any time a subject experiences an increase in GI symptoms (see Section 7.2.3.12). Diagnosis of *C. difficile* infection should be made using the central laboratory. If, for any reason, the central laboratory is not available, please see [Appendix 5](#) for guidance regarding diagnostic algorithms.

A TB test (PPD or QuantiFERON-TB Gold Plus) will be performed at screening (Visit 1). A documented negative PPD test within 12 weeks before screening (Visit 1) is acceptable. The IGRA official reading and method or test must be located in the source documentation.

A serum sample will be collected and banked for John Cunningham virus antibody testing. It may be analyzed if a subject shows neurological symptoms suggestive of PML.

All laboratory assessments should be performed at central laboratories, with the exception of the following assessments: stool microbiology (local or central laboratory) and TB test (PPD or QuantiFERON-TB Gold Plus) (refer to laboratory manual for details).

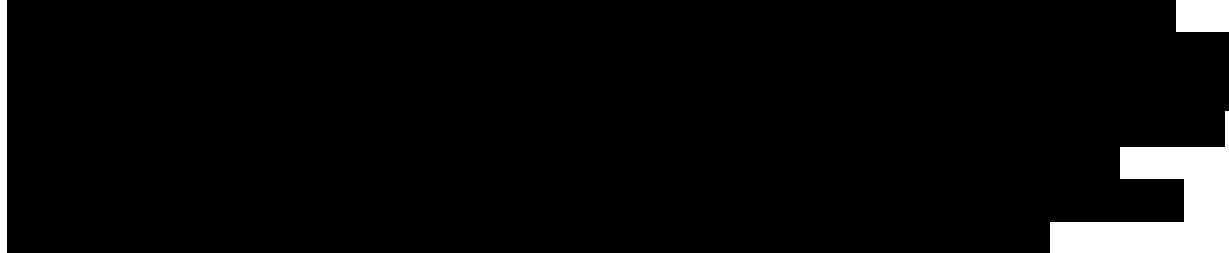
### 7.2.3.7 Pregnancy Test and Follicle-stimulating Hormone Test

A beta-human chorionic gonadotropin ( $\beta$ -hCG) pregnancy test will be performed on all females of childbearing potential at the time points specified in [Table 1](#); if pregnancy is suspected; or on withdrawal of the subject from the study. A serum pregnancy test will be performed at screening (Visit 1); all other pregnancy tests will be urine tests.

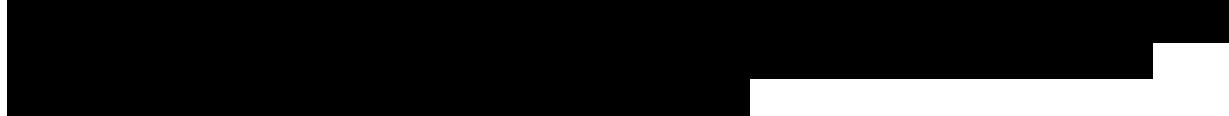
- Emotional function: 12 to 84
- Social function: 5 to 35.

The total IBDQ score ranges from 32 to 224. For the total score and each domain, a higher score indicates better HRQL. A score of at least 170 corresponds to clinical remission and an increase of at least 16 points is considered to indicate a clinically meaningful improvement.

The IBDQ is presented in [Appendix 2](#).



The [REDACTED] is presented in [Appendix 2](#).



The [REDACTED] and [REDACTED] are presented in [Appendix 2](#).



### **Short Form-36 Health Survey (Version 2, Acute Form)**

The SF-36 is a generic quality-of-life instrument that has been widely used to assess HRQL of subjects. Generic instruments are used in general populations to assess a wide range of domains applicable to a variety of health states, conditions, and diseases. The SF-36 consists of 36 items that are aggregated into 8 multi-item scales (physical functioning, role – physical, bodily pain, general health, vitality, social functioning, role – emotional, and mental health), with scores ranging from 0 to 100 ([Ware and Sherbourne, 1992](#)). Higher scores indicate better HRQL.

## 7.2.5 Volume of Blood to Be Drawn From Each Subject

The volume of blood to be drawn from each subject is summarized in [Table 4](#).

**Table 4 Volume of Blood to Be Drawn From Each Subject**

Assessment	Sample Volume (mL)	Number of Samples	Total Volume (mL)
Hematology	2	7	14
Serum chemistry	4	7	28
HBsAg	2	1	2
HBcAb	2	1	2
HCVAb	2	1	2
HBV DNA	6	1	6
HIV	2	1	2
FSH	2	1	2
Serum β-hCG <sup>a</sup>	2	1	2
TB test (QuantiFERON-TB Gold Plus or PPD)	4	1	4
JCV antibody banked sample	3.5	1	3.5
[REDACTED]			
[REDACTED]	2	3	6
[REDACTED]	5	3	15
[REDACTED]	4	3	12
[REDACTED]	5	6	30
ADA and NAb sampling	3	7	21
<b>Total mL</b>			<b>151.5</b>

ADA=antidrug antibody; β-hCG=beta-human chorionic gonadotropin; [REDACTED]; FSH=follicle-stimulating hormone; HBcAb=hepatitis B core antibody; HBsAg=hepatitis B surface antigen; HBV=hepatitis B virus; HCVAb=hepatitis C virus antibody; HIV=human immunodeficiency virus; JCV=John Cunningham virus; [REDACTED]  
[REDACTED]; NAb=neutralizing antibody; PPD=purified protein derivative; TB=tuberculosis

<sup>a</sup> β-hCG testing for female subjects only.

b If a catheter is used, the first mL is to be discarded; then take 4 mL into appropriate tube for [REDACTED]. A total of 5 mL of blood drawn has been used in determination of sample volume.

The amount of blood to be drawn for each assessment is an estimate. The amount of blood to be drawn may vary according to the instructions provided by the manufacturer or laboratory for an individual assessment; however, the total volume drawn over the course of the study should be approximately 151.5 mL. When more than 1 blood assessment is to be done at the time point/period, if they require the same type of tube, the assessments may be combined.

The medical assessment of severity is determined by using the following definitions:

**Mild:** A type of AE that is usually transient and may require only minimal treatment or therapeutic intervention. The event does not generally interfere with usual activities of daily living.

**Moderate:** A type of AE that is usually alleviated with specific therapeutic intervention. The event interferes with usual activities of daily living, causing discomfort but poses no significant or permanent risk of harm to the research subject.

**Severe:** A type of AE that interrupts usual activities of daily living, or significantly affects clinical status, or may require intensive therapeutic intervention.

### 8.1.2 Relationship Categorization

A physician/investigator must make the assessment of relationship to investigational product for each AE. The investigator should decide whether, in his or her medical judgment, there is a reasonable possibility that the event may have been caused by the investigational product. If there is no valid reason for suggesting a relationship, then the AE should be classified as “not related.” Otherwise, if there is any valid reason, even if undetermined or untested, for suspecting a possible cause-and-effect relationship between the investigational product and the occurrence of the AE, then the AE should be considered “related.” The causality assessment must be documented in the source document.

The following additional guidance may be helpful:

Term	Relationship Definition
Related	The temporal relationship between the event and the administration of the investigational product is compelling and/or follows a known or suspected response pattern to that product, and the event cannot be explained by the subject’s medical condition, other therapies, or accident.
Not Related	The event can be readily explained by other factors such as the subject’s underlying medical condition, concomitant therapy, or accident and no plausible temporal or biologic relationship exists between the investigational product and the event.

### 8.1.3 Adverse Events of Special Interest

Adverse events of special interest will be captured and monitored during this study. Investigators will report all AESIs to the sponsor, regardless of causality, using the same timelines as described for SAE reporting (see Section 8.2.2). The following describes the AESIs and the criteria for reporting AESIs.

#### 8.1.3.1 Hypersensitivity

Potential hypersensitivity, serum sickness, vasculitis, and Arthus reactions to ontamalimab will be regarded as AESIs. These events must be reported on Shire “Clinical Study Serious Adverse Event and Nonserious AE as Required by the Protocol Form” and within the time frame mandated for SAEs (see Section 8.2.2).

### 8.1.5 Clinical Laboratory and Other Safety Evaluations

A change in the value of a clinical laboratory, vital sign, or ECG assessment can represent an AE if the change is clinically relevant or if, during treatment with the investigational product, a shift of a parameter is observed from a normal value to an abnormal value, or a further worsening of an already abnormal value. When evaluating such changes, the extent of deviation from the reference range, the duration until return to the reference range, either while continuing treatment or after the end of treatment with the investigational product, and the range of variation of the respective parameter within its reference range, must be taken into consideration.

If, at the end of the treatment phase, there are abnormal clinical laboratory, vital sign, or ECG values which were not present at the pretreatment value observed closest to the start of study treatment, further investigations should be performed until the values return to within the reference range or until a plausible explanation (eg, concomitant disease) is found for the abnormal values.

The investigator should decide, based on the above criteria and the clinical condition of a subject, whether a change in a clinical laboratory, vital sign, or ECG parameter is clinically significant and therefore represents an AE.

### 8.1.6 Pregnancy

All pregnancies are to be reported from the time informed consent is signed until the defined follow-up period stated in Section 7.1.3.

Any report of pregnancy for any female study participant or the partner of a male study participant must be reported within 24 hours to the Shire Global Drug Safety Department using the Shire Investigational and Marketed Products Pregnancy Report Form. A copy of the Shire Investigational and Marketed Products Pregnancy Report Form (and any applicable follow-up reports) must also be sent to the CRO/Shire medical monitor using the details specified in the **emergency contact information** section of the protocol. The pregnant female study participant must be withdrawn from the study.

Every effort should be made to gather information regarding the pregnancy outcome and condition of the infant. It is the responsibility of the investigator to obtain this information within 30 calendar days after the initial notification and approximately 30 calendar days postpartum.

Pregnancy complications such as spontaneous abortion/miscarriage or congenital abnormality are considered SAEs and must be reported using the Shire Clinical Study Serious Adverse Event and Nonserious AEs Required by the Protocol Form. Note: An elective abortion is not considered an SAE.

In addition to the above, if the investigator determines that the pregnancy meets serious criteria, it must be reported as an SAE using the Shire Clinical Study Serious Adverse Event and Nonserious AEs Required by the Protocol Form as well as the Shire Investigational and Marketed Products Pregnancy Report Form. The test date of the first positive serum/urine  $\beta$ -hCG test or ultrasound result will determine the pregnancy onset date.

### 8.1.7 Abuse, Misuse, Overdose, and Medication Error

Abuse, misuse, overdose, or medication error (as defined below) must be reported to the sponsor according to the SAE reporting procedure whether or not they result in an AE/SAE as described in Section 8.2. Note: The 24-hour reporting requirement for SAEs does not apply to reports of abuse, misuse, overdose, or medication errors unless these result in an SAE.

The categories below are not mutually exclusive; the event can meet more than 1 category.

- **Abuse** – Persistent or sporadic intentional intake of investigational product when used for a nonmedical purpose (eg, to alter one's state of consciousness or get high) in a manner that may be detrimental to the individual and/or society
- **Misuse** – Intentional use of investigational product other than as directed or indicated at any dose (Note: This includes a situation where the investigational product is not used as directed at the dose prescribed by the protocol)
- **Overdose** – Intentional or unintentional administration of investigational product at a dose interval that is less than 2 weeks between doses
- **Medication Error** – An error made in prescribing, dispensing, administration, and/or use of an investigational product. For studies, medication errors are reportable to the sponsor only as defined below.

Cases of subjects missing doses of the investigational product are not considered reportable as medication errors.

Medication errors should be collected/reported for all products under investigation.

The administration and/or use of the unassigned treatment is/are always reportable as a medication error.

The administration and/or use of an expired investigational product should be considered as a reportable medication error.

There is no specific antidote for overdose with SHP647. Treatment should be symptomatic and supportive.

### 8.1.8 Unexpected Adverse Event

An unexpected AE is an AE whose nature, severity, specificity, or outcome is not consistent with the term, representation, or description used in the reference safety information (RSI).

“Unexpected” also refers to the AEs that are mentioned in the IB as occurring with a class of drugs or as anticipated from the pharmacological properties of the product, but are not specifically mentioned as occurring with the particular product under investigation.

The expectedness of AEs will be determined by the sponsor using the IB as the RSI. This determination will include considerations such as the number of AEs previously observed, but not on the basis of what might be anticipated from the pharmacological properties of a product.

### **8.1.9 Suspected Unexpected Serious Adverse Reaction**

A suspected unexpected serious adverse reaction (SUSAR) is defined as any suspected adverse reaction to study treatment (ie, including active comparators) that is both serious and unexpected.

The event(s) must meet all of the following:

- Suspected adverse reaction
- Serious
- Unexpected
- Assessed as related to study treatment.

## **8.2 Serious Adverse Event Procedures**

### **8.2.1 Reference Safety Information**

The reference for safety information for this study is Section 6.8 of the SHP647 IB, which the sponsor has provided under separate cover to all investigators.

### **8.2.2 Reporting Procedures**

All initial and follow-up SAE reports must be reported by the investigator to the Shire Global Drug Safety Department and the CRO/Shire medical monitor within 24 hours of the first awareness of the event. Note: The 24-hour reporting requirement for SAEs does not apply to reports of abuse, misuse, overdose, or medication errors (Section 8.1.7) unless they result in an SAE.

The investigator must complete, sign, and date the Shire Clinical Study Serious Adverse Event and Nonserious AEs Required by the Protocol Form and verify the accuracy of the information recorded on the form with the corresponding source documents (Note: Source documents are not to be sent unless requested) and fax or email the form to the Shire Global Drug Safety Department. A copy of the Shire Clinical Study Serious Adverse Event and Nonserious AEs Required by the Protocol Form (and any applicable follow-up reports) must also be sent to the CRO/Shire medical monitor using the details specified in the [emergency contact information](#) section of the protocol.

### **8.2.3 Serious Adverse Event Definition**

An SAE is any untoward medical occurrence (whether considered to be related to investigational product or not) that at any dose:

- Results in death
- Is life-threatening. Note: The term “life-threatening” in the definition of “serious” refers to an event in which the subject was at risk of death at the time of the event; it does not refer to an event which hypothetically might have caused death if it was more severe.

The investigator must complete, sign, and date the Shire “Clinical Study Serious Adverse Event and Nonserious AE as Required by the Protocol Form” and verify the accuracy of the information recorded on the form with the corresponding source documents (Note: Source documents are not to be sent unless requested) and fax or email the form to the Shire Global Drug Safety Department. A copy of the Shire “Clinical Study Serious Adverse Event and Nonserious AE as Required by the Protocol Form” (and any applicable follow-up reports) must also be sent to the CRO/Shire medical monitor using the details specified in the emergency contact information section of the protocol.

### **8.2.3 Serious Adverse Event Definition**

An SAE is any untoward medical occurrence (whether considered to be related to investigational product or not) that at any dose:

- Results in death.
- Is life-threatening. Note: The term “life-threatening” in the definition of “serious” refers to an event in which the subject was at risk of death at the time of the event; it does not refer to an event which hypothetically might have caused death if it was more severe.
- Requires inpatient hospitalization or prolongation of existing hospitalization.  
Note: Hospitalizations, which are the result of elective or previously scheduled surgery for pre-existing conditions, which have not worsened after initiation of treatment, should not be classified as SAEs. For example, an admission for a previously scheduled ventral hernia repair would not be classified as an SAE; however, complication(s) resulting from a hospitalization for an elective or previously scheduled surgery that meet(s) serious criteria must be reported as SAE(s).
- Results in persistent or significant disability/incapacity.
- Is a congenital abnormality/birth defect.
- Is an important medical event. Note: Important medical events that may not result in death, be life-threatening, or require hospitalization may be considered an SAE when, based upon appropriate medical judgment, they may jeopardize the subject and may require medical or surgical intervention to prevent 1 of the outcomes listed in this definition. Examples of such medical events include allergic bronchospasm requiring intensive treatment in an ED or at home; blood dyscrasias or convulsions that do not result in inpatient hospitalization; or the development of drug dependency or drug abuse.

### **8.2.4 Serious Adverse Event Collection Time Frame**

All SAEs (regardless of relationship to study) are collected from the time the subject signs the informed consent until the defined follow-up period stated in Section [7.1.3](#) and must be reported to the Shire Global Drug Safety Department and the CRO/Shire medical monitor within 24 hours of the first awareness of the event.

In addition, any SAE(s) considered “related” to the investigational product and discovered by the investigator at any interval after the study has completed must be reported to the Shire Global Drug Safety Department within 24 hours of the first awareness of the event.

### **8.2.5 Serious Adverse Event Onset and Resolution Dates**

The onset date of the SAE is defined as the date the event meets serious criteria. The resolution date is the date the event no longer meets serious criteria, the date the symptoms resolve, or the date the event is considered chronic. In the case of hospitalizations, the hospital admission and discharge dates are considered the onset and resolution dates, respectively.

In addition, any signs or symptoms experienced by the subject after signing the informed consent form, leading up to the onset date of the SAE, or following the resolution date of the SAE, must be recorded as an AE, if appropriate.

### **8.2.6 Fatal Outcome**

Any SAE that results in the subject's death (ie, the SAE was noted as the primary cause of death) must have fatal checked as an outcome with the date of death recorded as the resolution date. For all other events ongoing at the time of death that did not contribute to the subject's death, the outcome should be considered not resolved, without a resolution date recorded.

For any SAE that results in the subject's death or any ongoing events at the time of death, unless another investigational product action was previously taken (eg, drug interrupted, reduced, withdrawn), the action taken with the investigational product should be recorded as "dose not changed" or "not applicable" (if the subject never received investigational product). The investigational product action of "withdrawn" should not be selected solely as a result of the subject's death.

### **8.2.7 Regulatory Agency, Institutional Review Board, Ethics Committee, and Site Reporting**

The sponsor or the CRO is responsible for notifying the relevant regulatory authorities, US central institutional review boards (IRBs), and European Union (EU) central ethics committees (ECs) of related, unexpected SAEs (ie, SUSARs).

In addition, the CRO is responsible for notifying active sites of all related, unexpected SAEs (ie, SUSARs) occurring during all interventional studies across the ontamalimab program.

The investigator is responsible for notifying the local IRB, local EC, or the relevant local regulatory authority of all SAEs that occur at his or her site as required.

## 8.2.8 Safety Monitoring for Potential Cases of Drug-induced Liver Injury

The following safety monitoring and stopping criteria are provided for elevated hepatic blood tests based on normal and elevated baseline ALT and total bilirubin levels.

Abnormal values in ALT concurrent with abnormal elevations in total bilirubin that meet the criteria outlined below in the absence of other causes of liver injury are considered potential cases of drug-induced liver injury (potential Hy's Law cases) and should always be considered important medical events.

The threshold of laboratory abnormalities for a potential case of drug-induced liver injury depends on the subject's individual baseline values and underlying conditions. Subjects who present with the following laboratory abnormalities per [Table 6](#) should be evaluated further to definitively determine the etiology of the abnormal laboratory values. The measurement(s) should be reconfirmed with another blood draw preferably within 48 to 72 hours of the initial finding of potential concern. Please refer to laboratory manual for further instructions.

Guidance for Dosing Interruption: Investigator-directed delays in dosing due to abnormal laboratory findings or AEs should be discussed with the medical monitor to determine whether the subject should continue with the treatment. Only those subjects who complete the full course of investigational product treatment in the induction studies (SHP647-305 or SHP647-306) will be eligible to continue in the maintenance study or LTS study.

**Table 6 Safety Monitoring Rules for Treatment-emergent Elevated ALT and/or Bilirubin**

Treatment-emergent ALT	Treatment-emergent total bilirubin	Treatment-emergent symptoms	Action
<u>Normal baseline</u>  ALT $\geq 5 \times$ ULN	Normal	None	Repeat ALT, AST, ALP, TBL, in 2-5 days. Follow-up for symptoms.
<u>Elevated baseline<sup>a</sup>:</u>  ALT $\geq 3 \times$ baseline <i>or</i> $\geq 300$ U/L (whichever occurs first)	<u>Patients with Gilbert's syndrome or hemolysis:</u>  No change in baseline TBL		Initiate evaluation for other etiologies of abnormal liver tests.  Testing for hepatitis A, B, and/or C infection may be warranted. Subjects who entered the study with HBcAb with or without HBsAb would need evaluation with HBV DNA to rule out HBV reactivation. <sup>c</sup>
<u>Normal baseline</u>  ALT $\geq 8 \times$ ULN	Normal	None	Interrupt investigational product. <sup>b</sup> Initiate close monitoring and workup for competing etiologies.
<u>Elevated baseline<sup>a</sup>:</u>  ALT $\geq 5 \times$ baseline or $\geq 500$ U/L (whichever occurs first)	<u>Patients with Gilbert's syndrome or hemolysis:</u>  No change in baseline TBL		Investigational product can be restarted only if another etiology is identified and liver enzymes return to baseline.  Testing for hepatitis A, B, and/or C infection may be warranted. Subjects who entered the study with HBcAb with or without HBsAb would need evaluation with HBV DNA to rule out HBV reactivation. <sup>c</sup>

Further details regarding the DMC can be found in the DMC charter, which will be available before the administration of investigational product to any subject. Analyses of the data for DMC review will be conducted according to the DMC charter and DMC SAP. Because no formal hypothesis testing for safety assessments is planned, multiplicity concerns regarding repeated analyses are not applicable.

## 9.6 Sample Size Calculation and Power Considerations

Graphical methods are used to control the global family-wise type I error rate (FWER) at the .05 level (2-sided) for the comparisons of the 2 SHP647 treatment groups with the placebo group. Alpha is initially split equally at the .025 level (2-sided) for each of the pairwise treatment comparisons for the coprimary endpoints. Therefore, the power analysis and sample size estimation was calculated based on the chi-square test of proportions using nQuery Advisor® Version 7.0 (Statistical Solutions Ltd, Cork, Ireland) for an individual SHP647 dose compared to placebo.

Power calculations are made based on assuming a .025 (2-sided) significance level for each pairwise treatment comparison, 1720 subjects will be screened to randomize 1032 subjects in a 3:3:2 allocation ratio: 387 subjects in the 25 mg SHP647 treatment group, 387 subjects in the 75 mg SHP647 treatment group, and 258 subjects in the placebo group. These numbers of subjects would yield an approximately 93% power to detect individual pairwise treatment difference in the first coprimary efficacy endpoint, clinical remission by 2-item PRO at Week 16, of 10% (17.5% SHP647 versus 7.5% placebo). Expected clinical remission rates by 2-item PRO at Week 16 are based on observed rates from a post hoc analysis of the A7281006 study and placebo remission rates from literature ([Sandborn et al., 2017](#)). No adjustment for missing data is required in these sample size calculations as subjects with missing data for clinical remission by 2-item PRO at Week 16 are imputed as failures and the above rates account for these subjects.

With the 1032 subjects in allocation noted above, this number of subjects would yield an approximately 94% power to detect individual pairwise treatment difference in the other coprimary efficacy endpoint, endoscopic response at Week 16, of 12.5% (27.5% SHP647 versus 15% placebo). Expected endoscopic response rates at Week 16 are based on observed rates from a post hoc analysis of the A7281006 study and also endoscopic response rates from literature ([Sandborn et al., 2017](#)). No adjustment for missing data is required in these sample size calculations as subjects with missing data for endoscopic response at Week 16 are imputed as failures and the above rates account for these subjects.

The overall power for the coprimary endpoints will be approximately 87% assuming no correlation between the tests on the endpoints and approximately 90% assuming a correlation of 0.4.

With the sample size of 1032 subjects, [Table 6](#) provides the power for detecting a treatment difference between a SHP647 treatment group and the placebo group for the key secondary endpoints.

The coprimary endpoints will each be tested by the following hypothesis:

$$H_0: \delta = 0$$

$$H_1: \delta \neq 0$$

Where  $\delta$  is the common treatment difference across strata. 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 the p-value computed from the p-value combination method will be presented for each active treatment group to placebo comparison for each endpoint.

Sensitivity/supplementary analyses to explore the impact of missing data on the coprimary efficacy endpoints and key secondary endpoints will be conducted. These analyses may compare imputations of the missing values which favor placebo (eg, worst case) and/or imputations which favor active treatment (eg, best case). In addition, imputation methods based on informative missingness and other missing data mechanisms may be performed. Additional sensitivity/supplementary analyses will also be conducted using the PP set and the completer set. Additional analyses may be developed in the SAP. All sensitivity/supplementary analyses will be described in the SAP.

Prespecified subgroup analyses are planned for the coprimary endpoints including, but not limited to gender, prior anti-TNF treatment, glucocorticoid use at baseline, SES-CD at baseline, region, age group, randomization stratum, and other important subgroups. A full list of important subgroups will be described in the SAP. Within subgroups, efficacy endpoints will be compared for each active treatment group (25 mg ontamalimab and 75 mg ontamalimab) with the placebo group using a chi-square test. The estimate of the treatment difference, along with the corresponding Newcombe (hybrid-score) 95% CI and chi-square test p-value, will be presented.

#### Statistical Testing and Protection of the Type I Error

The global FWER for the statistical tests of the coprimary and key secondary endpoints will be strongly controlled at .05 (2-sided). To control the FWER, graphical methods discussed in Bretz et al. (2009b) will be utilized to propagate  $\alpha$  from the coprimary endpoints to the key secondary endpoints and between the 2 ontamalimab treatment group and placebo comparisons. Alpha is initially split equally at the .025 level (2-sided) for each of the pairwise treatment comparisons for the coprimary endpoints (P) and alpha is propagated in a hierarchical manner to each of the 6 key secondary endpoints (K1–K6) within a pairwise treatment comparison. In order to pass alpha between the coprimary endpoints and the first key secondary endpoint, both coprimary endpoints must attain statistical significance. A graphical visualization of the  $\alpha$  propagation is presented in [Figure 5](#).

- Complete endoscopic healing at Week 16 defined as SES-CD=0-2.

Similar to the coprimary endpoints, the 6 key secondary endpoints will all be tested by the following hypothesis:

$$H_0: \delta = 0$$

$$H_1: \delta \neq 0$$

The key secondary endpoints will be analyzed using the same approach as described for the coprimary endpoints. Subjects with missing key secondary endpoint data at the Week 16 visit will be considered failures and counted as nonresponders.

### 9.8.2.2 Other Secondary Efficacy Endpoints

The other secondary endpoints are as follows:

- Clinical response at the Week 16 visit as measured by at least a 100-point reduction in the CDAI from baseline (CDAI-100 response)
- Clinical response at the Week 16 visit as measured by at least a 70-point reduction in the CDAI from baseline (CDAI-70 response)
- Clinical remission over time, as measured by the 2-item PRO
- Change from baseline in total stool frequency, rectal bleeding frequency, rectal urgency frequency, nausea severity, vomiting frequency, and rectal incontinence frequency scores; and total sign/symptom score based on subject daily e-diary entries
- Endoscopic healing at Week 16 as measured by SES-CD  $\leq 4$  and at least 2-point reduction versus baseline and no subscore  $> 1$  in any individual variable
- Change from baseline in IBDQ domain and total (absolute) scores over time
- Change from baseline in SF-36, version 2, acute (physical and mental component summary scores and individual domain scores) over time
- Incidence of all cause hospitalizations and total inpatient days
- Incidence of CD-related surgeries and other surgical procedures during the entire study period.

Other secondary endpoints will be summarized by descriptive statistics and presented by treatment group. Where appropriate, other secondary efficacy endpoints will be analyzed with the following analysis methods:

- Binary endpoints will be compared between each active treatment group and the placebo group using a CMH chi-square test stratified by status of prior anti-TNF treatment, glucocorticoid use at baseline, and SES-CD at baseline. The estimate of the common treatment difference along with the corresponding stratified Newcombe 95% CI using the method of Yan and Su (2010) and the p-value computed from CHM test will be provided. Subjects with missing binary endpoint data at the Week 16 visit will be considered failures and counted as nonresponders.
- Continuous endpoints that are only measured at baseline and the Week 16 visit will be analyzed using an analysis of covariance model with fixed effects for treatment group (categorical), status of prior anti-TNF treatment (categorical), glucocorticoid use at baseline (categorical), and SES-CD at baseline (categorical), and the baseline value as a continuous covariate. From this model, estimates of the least squares means, treatment differences, standard errors, p-values, and 95% CIs for least squares mean treatment differences will be provided.
- Continuous endpoints that are measured repeatedly over time will be analyzed using a linear repeated measures mixed model with restricted maximum likelihood estimation. The model will include fixed effects for treatment group (categorical), visit (categorical), treatment group by visit interaction, status of prior anti-TNF treatment (categorical), glucocorticoid use at baseline (categorical), and SES-CD at baseline (categorical); baseline value as a continuous covariate; and repeated measures across visit for subject. From this model, estimates of least squares means, treatment differences, standard errors, p-values, and 95% CIs for least squares mean treatment differences for each visit will be provided.

### 9.8.3 Exploratory Efficacy Endpoints

The exploratory endpoints are as follows:



Exploratory efficacy endpoints will be summarized with descriptive statistics and presented by treatment group using the same approach described as for the other secondary endpoints. See Section 9.8.2.2 for an overview of the planned analyses. Full details for the analysis of exploratory efficacy endpoints will be included in the SAP.

## 9.9 Safety Analyses

All safety analyses will be performed using the safety set. Subjects will be analyzed according to the treatment they actually received.

Adverse events will be coded using the Medical Dictionary for Regulatory Activities.

Treatment-emergent AEs are defined as AEs with start dates at the time of or following the first exposure to investigational product. The number of events, incidence, and percentage of TEAEs will be calculated overall, by SOC, by preferred term, and by treatment group.

Treatment-emergent AEs will be further summarized by severity and relationship to investigational product. Adverse events leading to withdrawal, SAEs, and deaths will be similarly summarized or listed. Adverse events of special interest will be summarized by treatment group.

Clinical laboratory tests, vital signs, and ECG findings will be summarized by treatment group and visit. Potentially clinically important findings will also be summarized or listed.

Antidrug antibody data will be summarized by treatment group and visit.

Further details of safety analyses will be described in the SAP.

## 9.10 Other Analyses



### Simple Endoscopic Score for Crohn's Disease (SES-CD)

#### SITE WORKSHEET: Simple Endoscopic Score for Crohn's Disease (SES-CD)

Site #	Investigator	Subject ID	Visit Date (mmm dd yyyy)

#### Definitions of Simple Endoscopic Score for Crohn's Disease

		Simple Endoscopic Score for Crohn's Disease values		
Variable				
Size of Ulcers	None	Apthous ulcers (Ø 0.1 to 0.5 cm)	Large ulcers (Ø 0.5 to 2 cm)	Very large ulcers (Ø >2 cm)
Ulcerated surface	None	<10%	10-30%	>30%
Affected surface	Unaffected segment	<50%	50-75%	>75%
Presence of Narrowing	None	Single, can be passed	Multiple, can be passed	Cannot be passed

Endoscopy Finding	Ileum	Right Colon	Transverse Colon	Left Colon	Rectum
<b>Presence and size of ulcers</b>					
None	○	○	○	○	○
Apthous ulcers (Ø 0.1 to 0.5 cm)	○	○	○	○	○
Large ulcers (Ø 0.5 to 2 cm)	○	○	○	○	○
Very large ulcers (Ø >2 cm)	○	○	○	○	○
<b>Extent of ulcerated surface</b>					
None	○	○	○	○	○
< 10%	○	○	○	○	○
10 – 30%	○	○	○	○	○
> 30%	○	○	○	○	○
<b>Extent of affected surface</b>					
Unaffected segment	○	○	○	○	○
< 50%	○	○	○	○	○
50 – 75%	○	○	○	○	○
>75%	○	○	○	○	○
<b>Presence and type of narrowings</b>					
None	○	○	○	○	○
Single, can be passed	○	○	○	○	○
Multiple, can be passed	○	○	○	○	○
Cannot be passed	○	○	○	○	○

4. You indicated you had X bowel movements in the past 24 hours. Of these, how many had blood, either in the stool, the toilet bowel, or on the toilet paper?

Enter number of bowel movements with blood: \_\_\_\_\_

5. You indicated you had X bowel movements in the past 24 hours. How many of these involved urgency (having to suddenly rush to the toilet to make it on time)?

Enter number of bowel movements with urgency: \_\_\_\_\_

6. Please rate your worst feeling of nausea (feeling sick to your stomach or like you might throw up) over the past 24 hours.

None

Mild

Moderate

Severe

7. How many vomiting episodes did you have in the past 24 hours? An episode includes one or multiple heaves (including dry heaves) in quick succession followed by a break in vomiting.

Enter number of vomiting episodes: \_\_\_\_\_

8. How many bowel incontinence episodes (losing control of your bowels before reaching the toilet) did you have in the past 24 hours?

Enter number of bowel incontinence episodes: \_\_\_\_\_

9. Please rate your abdominal pain over the past 24 hours.

None

Mild

Moderate

Severe

## 9.8 Efficacy Analyses

Unless otherwise specified, all efficacy analyses will be based on the FAS and subjects will be analyzed according to their randomized treatment, regardless of the treatment they actually received.

### 9.8.1 Coprimary Efficacy Endpoints

The coprimary efficacy endpoints are:

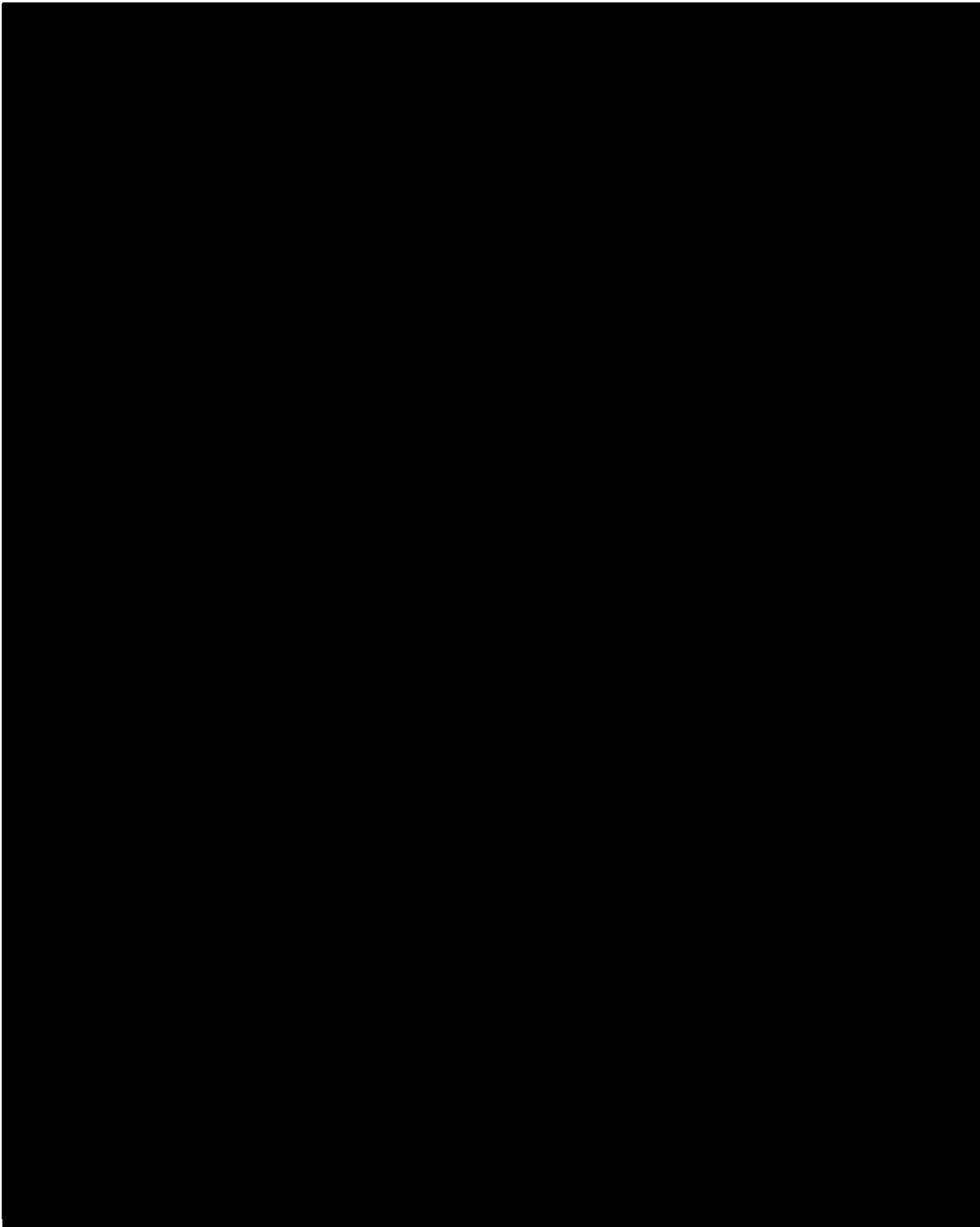
- Clinical remission at the Week 16 visit as defined by the following: 2-item PRO subscores of average worst daily abdominal pain  $\leq 3$  (based on 11-point NRS) over the 7 most recent days and average daily stool frequency  $\leq 2$  of type 6/7 (very soft stools/liquid stools) as shown in the BSFS over the 7 most recent days. The 7 most recent days may or may not be contiguous during the 10 days of data collection before colonoscopy preparation, depending on days to be excluded because of missing data. If fewer than 7 days are available, the endpoint will be calculated on all available most recent 6 or 5 days. If fewer than 5 days are available, the endpoint will be treated as missing.
- Endoscopic response at Week 16 as measured by a decrease in SES-CD of at least 25% from baseline.

The coprimary efficacy endpoints, clinical remission at the Week 16 visit and endoscopic response at the Week 16 visit, will each be compared for each active treatment group (25 mg or 75 mg ontamalimab) to the placebo group using a Cochran-Mantel-Haenszel (CMH) chi-square test stratified by status of prior anti-TNF treatment, glucocorticoid use, and SES-CD at baseline for each of the stages of the study (stage 1 includes subjects whose primary efficacy data are used in the interim analysis and stage 2 includes all other subjects. Note: classification of stage 1 and stage 2 is based on the time of randomization rather than the time of study completion or termination). Subjects with missing data at the Week 16 visit will be considered failures and counted as nonresponders.

Weighted inverse normal p-value combination methods are used to combine the p-values from stage 1 and stage 2 through the following formula:

$$C(p_1, p_2) = 1 - \Phi[w_1 \Phi^{-1}(1-p_1) + w_2 \Phi^{-1}(1-p_2)]$$

Where  $p_1, p_2$  are the p-values computed from the CMH chi-square test for each stage,  $w_i^2 = n_i/(n_1 + n_2)$ ,  $n_1$  and  $n_2$  are the preplanned stage-wise sample sizes that are fixed at the time of the interim analysis based on an original total sample size, and  $\Phi$  denotes the cumulative distribution function of the standard normal distribution ([Bretz et al., 2009a](#)). Given that there is no possibility of stopping early for efficacy, that any potential stopping for futility of either or both doses of ontamalimab is nonbinding, and that weights are prespecified, the test statistic  $C(p_1, p_2)$  can be compared against the nominal alpha level to assess statistical significance ([Chang and Chow, 2008](#)).



[REDACTED]

[REDACTED]  
[REDACTED]

[REDACTED]

[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]

[REDACTED]

[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]

[REDACTED]

[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]

[REDACTED]

[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]

### 7.2.4.3 Health-related Quality-of-life Assessments

Each subject will complete the HRQL assessments at the site during the visits specified in [Table 1](#), using an electronic device. All health outcome and patient-reported questionnaires should be completed before any other assessments. The study site staff should check for completion of all PRO questionnaires.

It is important to note that PRO assessments are collected and evaluated in a different manner than observed or volunteered AEs. Given these differences, no attempt will be made to reconcile any apparent discrepancies between observed or volunteered AEs and PRO data collected from subjects. Adverse event incidence rates will not be calculated from these solicited data but rather from the information recorded by the investigator.

#### Inflammatory Bowel Disease Questionnaire

The IBDQ is a psychometrically validated PRO instrument for measuring the disease-specific HRQL in subjects with IBD, including CD. The IBDQ consists of 32 items, which are grouped into 4 domains: bowel function, systemic symptoms, emotional status, and social function ([Irvine et al., 1994](#)). The 4 domains are scored as follows:

- Bowel symptoms: 10 to 70
- Systemic symptoms: 5 to 35
- Emotional function: 12 to 84
- Social function: 5 to 35.

The total IBDQ score ranges from 32 to 224. For the total score and each domain, a higher score indicates better HRQL. A score of at least 170 corresponds to clinical remission and an increase of at least 16 points is considered to indicate a clinically meaningful improvement.

The IBDQ is presented in [Appendix 2](#).



The [REDACTED] is presented in [Appendix 2](#).

The [REDACTED] and [REDACTED] are presented in [Appendix 2](#).

### Short Form-36 Health Survey (Version 2, Acute Form)

The SF-36 is a generic quality-of-life instrument that has been widely used to assess HRQL of subjects. Generic instruments are used in general populations to assess a wide range of domains applicable to a variety of health states, conditions, and diseases. The SF-36 consists of 36 items that are aggregated into 8 multi-item scales (physical functioning, role – physical, bodily pain, general health, vitality, social functioning, role – emotional, and mental health), with scores ranging from 0 to 100 ([Ware and Sherbourne, 1992](#)). Higher scores indicate better HRQL.

The SF-36 is presented in [Appendix 2](#).

The TSQM is presented in [Appendix 2](#).

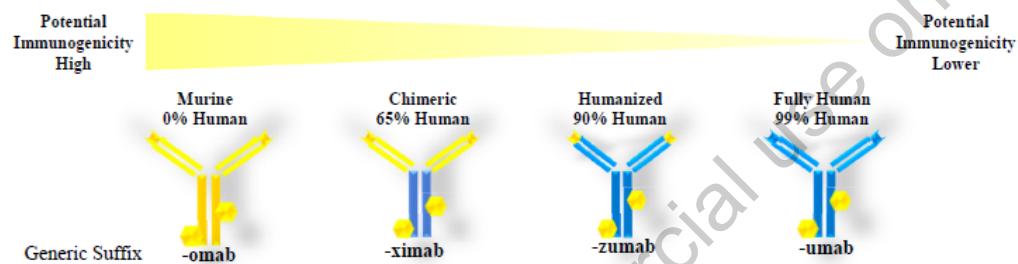
#### 7.2.4.4 Healthcare Resource Utilization Assessments

Hospitalizations, inpatient days, and [REDACTED] will be recorded at the time points specified in [Table 1](#). Information regarding CD-related and other surgeries will be collected from subjects during the treatment period.

It is well known that the administration of foreign proteins can cause immune responses including hypersensitivity reactions such as anaphylaxis and serum sickness. Other immune responses to foreign proteins include the development of ADAs and NAbs.

Monoclonal antibodies have been used in human therapeutics since the 1980s. The first monoclonal antibody approved for human use (ORTHOCLONE OKT3<sup>®</sup>) was a murine protein which caused rapid production of NAbS. Since then, much effort has been expended to reduce the immunogenicity of these useful therapeutic proteins by reducing the extent of “foreignness” from chimeric antibodies such as infliximab, to humanized antibodies such as vedolizumab, and finally to fully human antibodies such as adalimumab and ontamalimab (Isabwe et al., 2018) (see Figure 4).

**Figure 4 Potential Immunogenicity of Therapeutic Monoclonal Antibodies**



Ontamalimab is a fully human antibody of the immunoglobulin G2 subclass. In Phase 1 and Phase 2 clinical trials of ontamalimab, in which over 700 subjects were treated for up to 3 years, there has been no case of anaphylaxis. There have been 2 reported cases of drug hypersensitivity: serum sickness attributed to concomitant administration of penicillin; and a reaction characterized by dyspnea, facial erythema, and chest pain with onset 2 days after administration of the fifth dose of ontamalimab. The latter event mimicked a reaction that the subject had previously experienced after 4 doses of infliximab. In addition, low-titer activity has been observed in ADA assays, including pretreatment samples and placebo-treated subjects, and no subject has had a 2-fold or greater increase in ADA titer. Analysis of PK and clinical parameters has shown no difference between subjects whose ADA assay results are positive as compared with those who are negative.

Nonetheless, the possibility of a hypersensitivity reaction occurring after drug exposure cannot be fully ruled out. The reactions of concern are Type I (anaphylaxis) and Type III (immune complex) reactions. The clinical presentation of anaphylactic reactions is described in Table 5.

reports) must also be sent to the CRO/Shire medical monitor using the details specified in the emergency contact information section of the protocol. The pregnant female study participant must be withdrawn from the study.

Every effort should be made to gather information regarding the pregnancy outcome and condition of the infant. It is the responsibility of the investigator to obtain this information within 30 calendar days after the initial notification and approximately 30 calendar days postpartum. If the pregnancy outcome is a live birth, the vital status and clinical condition of the infant should be obtained and documented at 1 year postpartum.

Pregnancy complications such as spontaneous abortion/miscarriage or congenital abnormality are considered SAEs and must be reported using the Shire “Clinical Study Serious Adverse Event and Nonserious AE as Required by the Protocol Form.” Note: An elective abortion is not considered an SAE.

In addition to the above, if the investigator determines that the pregnancy meets serious criteria, it must be reported as an SAE using the Shire “Clinical Study Serious Adverse Event and Nonserious AE as Required by the Protocol Form” as well as the Shire Investigational and Marketed Products Pregnancy Report Form. The test date of the first positive serum/urine  $\beta$ -hCG test or ultrasound result will determine the pregnancy onset date.

### **8.1.8 Abuse, Misuse, Overdose, and Medication Error**

Abuse, misuse, overdose, or medication error (as defined below) must be reported to the sponsor according to the SAE reporting procedure whether or not they result in an AE/SAE as described in Section 8.2. Note: The 24-hour reporting requirement for SAEs does not apply to reports of abuse, misuse, overdose, or medication errors unless these result in an SAE.

The categories below are not mutually exclusive; the event can meet more than 1 category.

- **Abuse** – Persistent or sporadic intentional intake of investigational product when used for a nonmedical purpose (eg, to alter one’s state of consciousness or get high) in a manner that may be detrimental to the individual and/or society
- **Misuse** – Intentional use of investigational product other than as directed or indicated at any dose (Note: this includes a situation where the investigational product is not used as directed at the dose prescribed by the protocol)
- **Overdose** – Intentional or unintentional administration of investigational product at a dose interval that is less than 2 weeks between doses
- **Medication Error** – An error made in prescribing, dispensing, administration, and/or use of an investigational product. For studies, medication errors are reportable to the sponsor only as defined below.

Cases of subjects missing doses of the investigational product are not considered reportable as medication errors.

Medication errors should be collected/reported for all products under investigation.

Further details regarding the DMC can be found in the DMC charter, which will be available before the administration of investigational product to any subject. Analyses of the data for DMC review will be conducted according to the DMC charter and DMC SAP. Because no formal hypothesis testing for safety assessments is planned, multiplicity concerns regarding repeated analyses are not applicable.

An external hypersensitivity adjudication committee will be established to review data from subjects who experience a suspected Type I or Type III hypersensitivity reaction in order to confirm the nature and etiology of the reaction, to determine whether testing should be performed on stored blood samples, and to finalize recommendations of permanent discontinuation or rechallenge with investigational product. Further details regarding the adjudication committee can be found in the adjudication charter.

## 9.6 Sample Size Calculation and Power Considerations

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

Power calculations are made based on assuming a .025 (2-sided) significance level for each pairwise treatment comparison, 1720 subjects will be screened to randomize 1032 subjects in a 3:3:2 allocation ratio: 387 subjects in the 25 mg ontamalimab treatment group, 387 subjects in the 75 mg ontamalimab treatment group, and 258 subjects in the placebo group. These numbers of subjects would yield an approximately 93% power to detect individual pairwise treatment difference in the first coprimary efficacy endpoint, clinical remission by 2-item PRO at Week 16, of 10% (17.5% ontamalimab versus 7.5% placebo). Expected clinical remission rates by 2-item PRO at Week 16 are based on observed rates from a post hoc analysis of the A7281006 study and placebo remission rates from literature ([Sandborn et al., 2017](#)). No adjustment for missing data is required in these sample size calculations as subjects with missing data for clinical remission by 2-item PRO at Week 16 are imputed as failures and the above rates account for these subjects.

With the 1032 subjects in allocation noted above, this number of subjects would yield an approximately 94% power to detect individual pairwise treatment difference in the other coprimary efficacy endpoint, endoscopic response at Week 16, of 12.5% (27.5% ontamalimab versus 15% placebo). Expected endoscopic response rates at Week 16 are based on observed rates from a post hoc analysis of the A7281006 study and also endoscopic response rates from literature ([Sandborn et al., 2017](#)). No adjustment for missing data is required in these sample size calculations as subjects with missing data for endoscopic response at Week 16 are imputed as failures and the above rates account for these subjects.

**Table 1: Schedule of Assessments**

Study Procedure	Screening <sup>a</sup>		Baseline	Treatment						Follow-up		
	Weeks -6 to -1		Week 0/ Day 1	Week 2	Week 4	Week 8	Week 12	Week 16/ET <sup>b</sup>			Week 24 <sup>c</sup>	Week 32 <sup>c</sup>
Visit Number	1 (Part 1) <sup>a</sup>	1 (Part 2) <sup>a</sup>	2	3	4	5	6	7 (Part 1) <sup>b</sup>	7 (Part 2) <sup>b</sup>	7 (Part 3) <sup>b</sup>	8	9
Study Day	<b>-42 to 0</b>		1	14 ±3	28 ±3	56 ±3	84 ±3			112 ±3	168 ±7	224 ±7
<b>Treatment Procedures</b>												
Randomization <sup>w</sup>			X									
Administration of SHP647 or placebo <sup>w,x</sup>			X		X	X	X					
Adverse events	X	X	X	X	X	X	X	X	X	X	X	X
Prior medications	X											
Concomitant medications and procedures	X	X	X	X	X	X	X	X	X	X	X	X
Dispense stool collection kit for stool sample <sup>y</sup>	X					X	X					

Ab=antibody; ADA=antidrug antibodies; β-hCG=beta-human chorionic gonadotropin; CD=Crohn's disease; CDAI=Crohn's Disease Activity Index; [REDACTED]; ECG=electrocardiogram; [REDACTED]; [REDACTED]; ET=early termination; FSH=follicle-stimulating hormone; GDH=glutamate dehydrogenase; HBcAb=hepatitis B core antibody; HBsAg=hepatitis B surface antigen; HCV=hepatitis C virus; HCV RNA=hepatitis C virus ribonucleic acid; HRUA=Healthcare Resource Utilization Assessment; IBD=inflammatory bowel disease; IBDQ=Inflammatory Bowel Disease Questionnaire; JCV=John Cunningham virus; LTS=long-term safety extension; [REDACTED]; NAb=neutralizing antibody; PCR=polymerase chain reaction; PGA=physician's global assessment; [REDACTED]; PML=progressive multifocal leukoencephalopathy; PPD=purified protein derivative; PRO=patient-reported outcomes; SES-CD=Simple Endoscopic Score for CD; SF-36 v2=Short Form-36 Health Survey, version 2; TB=tuberculosis; [REDACTED].

<sup>a</sup> Screening assessments will take place over more than 1 day (at least 2 visits will be necessary to complete the screening evaluations, including colonoscopy).

<sup>b</sup> Subjects who withdraw early during the treatment period should return for the ET visit and then enter into the safety follow-up period. The Week 16 (Visit 7) and ET visits consist of 3 parts:

Part 1 of Visit 7 should be scheduled 1 to 3 day(s) before Part 2; this will allow for blood samples to be taken before starting the colonoscopy preparation and before the colonoscopy procedure at Part 2 of the visit (Section [7.2.2.4](#))

Part 2 of Visit 7 should be scheduled preferably 5 to 7 days before Part 3; this will allow sufficient time to obtain the data from the centrally read colonoscopy

Part 3 of Visit 7 will take place on Day 112 ±3 days.

<sup>c</sup> Subjects NOT entering the maintenance study (SHP647-307) or LTS (SHP647-304) study at the completion of the Week 16 visit will need to complete the safety follow-up assessments. The Week 24 (Visit 8) 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 32 (Visit 9) visit will take place at the study site.

<sup>d</sup> The outcome of Visit 7, Part 3 is used to assess eligibility to enroll subjects in the maintenance (SHP647-307) or LTS (SHP647-304) studies. Please refer to the respective protocols for further details.

The benefit of endoscopic response was also shown in the SONIC study; the presence of endoscopic response (defined in that study as at least a 50% decrease in endoscopic score at Week 26 of treatment) identified subjects most likely to be in corticosteroid-free clinical remission at Week 50 ([Ferrante et al., 2013](#)). The proportion of patients requiring major abdominal surgery in a single-center cohort study with infliximab was similar with complete healing or with partial healing. ([Schnitzler et al., 2009](#); [Panaccione et al., 2013](#)). Subjects with such a treatment response should be identified by endoscopic assessment in order not to misclassify them as nonresponders and underestimate the response to the treatment.

### 3.1.2 Rationale for Key Secondary Endpoints

#### Clinical Remission Defined by CDAI Score

Conventionally, a CDAI score of <150 has been used to define clinical remission. While there has been widespread use of the CDAI over a long period of time, the items do not contribute equally to the score, and symptom items reported by subjects are not specific for CD and are not sensitive for inflammation seen at colonoscopy. There has been movement away from using the CDAI by regulatory authorities to the use of PROs and objective measures of disease such as endoscopy ([Williet et al., 2014](#)). However, for benchmarking or for comparative effectiveness purposes, CDAI endpoints are expected to be used.

Although this has been the established gold-standard for clinical remission to date, CDAI suffers from requiring complex calculations across 8 individual items including subjective elements.

#### Clinical Remission Defined by Average Worst Daily Abdominal Pain $\leq 1$ (Based on the 4-point Scale) and Average Daily Stool Frequency $\leq 3$ of Type 6/7

The CDAI has been the traditionally used measure to assess clinical response and clinical remission in CD. In the CDAI, the degree of abdominal pain is one of 8 variables and is used with a multiplier of 5 in the overall score. Importantly, it is based on a 4-point scale, with scores ranging from 0 (none) to 3 (severe). With the shift to the new endpoint as evident from the coprimary endpoint of clinical remission in this study, it is still important to allow for a frame of reference to the existing standard for response, based on the CDAI components. A daily average abdominal pain threshold of  $\leq 1$  will help achieve this as 1 on the 4-point scale corresponds to mild abdominal pain. Although direct mapping between the scales has not been established, this will approximate to a score of 3 on the 11-point NRS scale, as this falls within the mild pain range on the NRS based on the literature.

Based on post hoc analyses of the Phase 2 data in the SHP647 program, regulatory requirements, and treatment separation assumptions, a threshold for the average daily stool frequency  $\leq 2$  of type 6/7 (very soft stools/liquid stools) was chosen for the coprimary endpoint of clinical remission. However, given the limited data available for this endpoint, recent evidence from literature suggesting that thresholds  $\leq 3$  are likely to be quite stringent ([Sandborn et al., 2017](#)), and the refractory nature of the disease in those with moderate to severe CD, it is also important to assess the effects of treatment using a more realistic measure. Hence, for this key secondary endpoint of clinical remission, average daily stool frequency  $\leq 3$  of type 6/7 (very soft stools/liquid stools) has been chosen as the appropriate threshold.

PP	per protocol
PPD	purified protein derivative
PRO	patient-reported outcome
Q4W	once every 4 weeks
RSI	reference safety information
SAE	serious adverse event
SAP	statistical analysis plan
SC	subcutaneous
SES-CD	Simple Endoscopic Score for Crohn's disease
SF-36	Short Form-36 Health Survey
SOC	system organ class
SUSAR	suspected unexpected serious adverse reaction
TB	tuberculosis
TEAE	treatment-emergent adverse event
TNF	tumor necrosis factor
[REDACTED]	[REDACTED]
UC	ulcerative colitis
ULN	upper limit of normal
[REDACTED]	[REDACTED]

22 Nov 2019

2. Subjects must be able to voluntarily provide written, signed, and dated (personally or via a legally authorized representative) informed consent and/or assent, as applicable, to participate in the study.
3. Subjects must be between  $\geq 16$  and  $\leq 80$  years of age at the time of the signing of the informed consent/assent form.  
Note: Subjects  $<18$  years of age must weigh  $\geq 40$  kg and must have body mass index  $\geq 16.5$  kg/m<sup>2</sup>.
4. Subjects must have active moderate to severe ileal (terminal ileum), ileocolic, or colonic CD at baseline (Visit 2) as defined by:
  - a. CDAI score between 220 and 450 (inclusive) **AND**
  - b. Meeting the following subscores in the 2-item PRO:
    - i. Abdominal pain subscore  $\geq 5$  (average worst daily pain on the 11-point numerical rating scale [NRS]) **AND** abdominal pain subscore  $\geq 2$  (average daily pain on the 4-point abdominal pain variable of CDAI) over the 7 most recent days out of the 10 days before colonoscopy preparation (may or may not be contiguous) **AND/OR**
    - ii. Average of the daily stool frequency subscore  $\geq 4$  of type 6/7 (very soft stools/liquid stools) as shown in the Bristol Stool Form Scale (BSFS) over the 7 most recent days out of the 10 days before colonoscopy preparation (may or may not be contiguous).

**AND**

- c. Presence of ulcerations that are characteristic to CD, as determined by a colonoscopy performed during screening, and as defined by the SES-CD  $>6$  (SES-CD  $\geq 4$  for isolated ileitis)

Note that the subject must be confirmed as meeting the CDAI score and PRO subscore requirements **before** a colonoscopy is done.

5. Subjects must have a documented diagnosis (endoscopic with histology) of CD for  $\geq 3$  months before screening. Documented diagnosis is defined as:
  - A biopsy report in which the description of the histological findings is consistent with the CD diagnosis **AND**
  - A report documenting disease duration based upon prior colonoscopy.

Note: If a biopsy report is not available in the source document at the time of screening, a biopsy must be performed during the screening colonoscopy and the histology report should be consistent with the CD diagnosis. If the histology description does not support the CD diagnosis at this time point, the subject should not be randomized.

6. Subjects must be willing and able to undergo a colonoscopy during screening after all other inclusion criteria have been met.
7. Subjects must have had an inadequate response to, or lost response to, or had an intolerance to at least 1 conventional treatment such as sulfasalazine or mesalamine (5-aminosalicylic acid [5-ASA]), glucocorticoids, immunosuppressants (azathioprine [AZA], 6-mercaptopurine [6-MP], or methotrexate [MTX]), or anti-TNF (see Appendix 4 of the protocol for guidance). Subjects who have had an inadequate response to sulfasalazine or mesalamine should have also failed at least 1 other conventional treatment such as glucocorticoids.
8. Subjects receiving any treatment(s) for CD described in Section 5.2.1 of the protocol are eligible provided they have been, and are anticipated to be, on a stable dose for the designated period of time.
9. Subjects are males or nonpregnant, nonlactating females who, if sexually active, agree to comply with the contraceptive requirements of the protocol, or females of nonchildbearing potential. Males and females of reproductive potential who are sexually active must agree to use appropriate contraception (ie, highly effective methods for female and medically appropriate methods for male study subjects) (as described in Section 4.4 of the protocol) for the duration of the study.

**Exclusion criteria:**

Subjects are excluded from the study if any of the following exclusion criteria are met.

1. Subjects with indeterminate colitis, microscopic colitis, nonsteroidal anti-inflammatory drug-induced colitis, ischemic colitis, infectious colitis, or clinical/histologic findings suggestive of ulcerative colitis.

**Table 1 Schedule of Assessments**

Study Procedure	Screening <sup>a</sup>		Baseline	Treatment						Follow-up		
	Weeks -6 to -1		Week 0/ Day 1	Week 2	Week 4	Week 8	Week 12	Week 16/ET <sup>b</sup>			Week 24 <sup>c</sup>	Week 32 <sup>c</sup>
Visit Number	1 (Part 1) <sup>a</sup>	1 (Part 2) <sup>a</sup>	2	3	4	5	6	7 (Part 1) <sup>b</sup>	7 (Part 2) <sup>b</sup>	7 (Part 3) <sup>b</sup>	8	9
Study Day	-42 to 0		1	14 ±3	28 ±3	56 ±3	84 ±3			112 ±3	168 ±7	224 ±7
<b>Treatment Procedures</b>												
Randomization <sup>x</sup>			X									
Administration of ontamalimab or placebo <sup>x,y</sup>			X		X	X	X					
Hypersensitivity monitoring <sup>z</sup>			X	X	X	X	X	X	X	X	X	X
Adverse events	X	X	X	X	X	X	X	X	X	X	X	X
Prior medications	X											
Concomitant medications and procedures	X	X	X	X	X	X	X	X	X	X	X	X
Dispense stool collection kit for stool sample <sup>aa</sup>	X						X	X				

ADA=antidrug antibody; β-hCG=beta-human chorionic gonadotropin; CD=Crohn's disease; CDAI=Crohn's Disease Activity Index; [REDACTED]; ECG=electrocardiogram; [REDACTED]; e-diary=electronic diary; [REDACTED]; ET=early termination; FSH=follicle-stimulating hormone; HBcAb=hepatitis B core antibody; HBsAg=hepatitis B surface antigen; HBV=hepatitis B virus; HCVAb=hepatitis C virus antibody; HIV=human immunodeficiency virus; HRUA=Healthcare Resource Utilization Assessment; IBDO=Inflammatory Bowel Disease Questionnaire; IGRA=interferon-γ release assay; ICV=John Cunningham virus; LTS=long-term safety extension; [REDACTED] NAb=neutralizing antibody; [REDACTED]; [REDACTED]; PML=progressive multifocal leukoencephalopathy; PPD=purified protein derivative; PRO=patient-reported outcome; SES-CD=Simple Endoscopic Score for CD; SF-36 v2=Short Form-36 Health Survey, version 2; TB=tuberculosis; [REDACTED]

<sup>a</sup> Screening assessments will take place over more than 1 day (at least 2 visits will be necessary to complete the screening evaluations, including colonoscopy).

<sup>b</sup> Subjects who withdraw early during the treatment period should return for the ET visit and then enter into the safety follow-up period. The Week 16 (Visit 7) and ET visits consist of 3 parts:

- Part 1 of Visit 7 can either be done on the same day as Part 2 or be done up to 3 day(s) before Part 2. If Parts 1 and 2 are done on the same day, blood samples must be taken before starting the colonoscopy preparation
- Part 2 of Visit 7 must be completed within 10 days (preferably within 5 to 7 days) before Part 3; this will allow sufficient time to obtain the data from the centrally read colonoscopy
- Part 3 of Visit 7 will take place on Day 112 ±3 days.

<sup>c</sup> Subjects NOT entering the maintenance study (SHP647-307) or LTS (SHP647-304) study at the completion of the Week 16 visit will need to complete the safety follow-up assessments. The Week 24 (Visit 8) 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 32 (Visit 9) visit will take place at the study site.

## **2.2 Study Objectives**

### **2.2.1 Coprimary Objectives**

The coprimary objectives of this study are to evaluate the efficacy of SHP647 in subjects with moderate to severe CD in:

- Inducing clinical remission based on 2-item patient-reported outcome (PRO) (abdominal pain severity and very soft stool/liquid stool frequency)
- Inducing endoscopic response based on centrally read colonoscopy.

### **2.2.2 Secondary Objectives**

#### **2.2.2.1 Key Secondary Objectives**

The key secondary objectives are as follows:

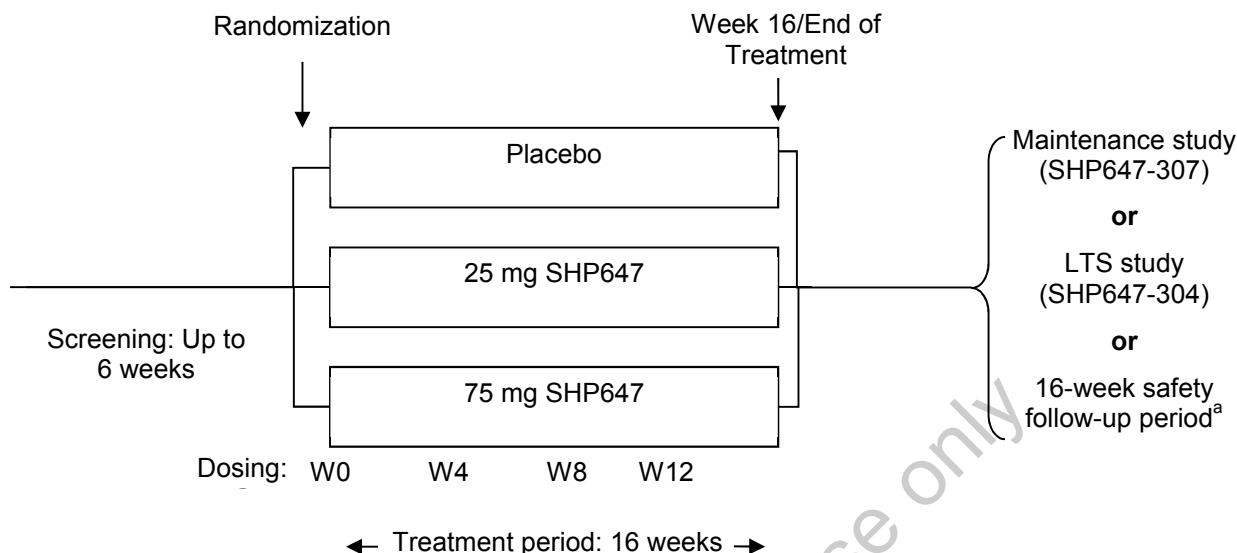
- To evaluate the efficacy of SHP647 in inducing clinical remission as measured by CDAI
- To evaluate the efficacy of SHP647 in inducing enhanced endoscopic response based on centrally read colonoscopy
- To evaluate the efficacy of SHP647 in inducing clinical remission based on abdominal pain severity and very soft stool/liquid stool frequency (alternate thresholds)
- To evaluate the efficacy of SHP647 in inducing clinical response based on patient-reported clinical signs and symptoms (as measured by 2-item PRO)
- To evaluate the efficacy of SHP647 in inducing clinical remission based on patient-reported clinical signs and symptoms (as measured by 2-item PRO) as well as inducing endoscopic response based on centrally read colonoscopy in the same subject
- To evaluate the efficacy of SHP647 in inducing endoscopic healing based on centrally read colonoscopy.

#### **2.2.2.2 Other Secondary Objectives**

The other secondary objectives are as follows:

- To evaluate the safety and tolerability of SHP647
- To evaluate the effect of SHP647 induction treatment on other clinical outcomes (clinical response defined by CDAI, or clinical remission over time, or change from baseline in frequency in CD-related clinical parameters)
- To evaluate the effect of SHP647 induction treatment on other endoscopic outcomes
- To evaluate the effect of SHP647 on health-related quality of life (HRQL) (as measured by the Inflammatory Bowel Disease Questionnaire [IBDQ] and the Short Form-36 Health Survey [SF-36])
- To evaluate the effect of SHP647 on incidence of hospitalizations and total inpatient days
- To evaluate the impact of SHP647 on incidence of CD-related and other surgeries.

**Figure 2: Study Design Flow Chart**



LTS=long-term safety extension; W=week.

<sup>a</sup>Subjects who withdraw early from the 16-week treatment period or who do not wish to enter the maintenance study (SHP647-307) or LTS study (SHP647-304) will continue into a 16-week safety follow-up period.

Note: A planned interim analysis for the coprimary endpoints will take place after approximately the first 50% of all randomized subjects in both the SHP647-305 and SHP647-306 studies have either completed the studies or have prematurely withdrawn from the studies.

### 3.1.1 Rationale for Coprimary Endpoints

In this study, clinical remission, as measured by a decrease below prespecified thresholds in the 2-item PRO (abdominal pain severity and very soft stool/liquid stool frequency [as shown in the Bristol Stool Form Scale, BSFS]), and enhanced endoscopic response, as measured by a decrease in SES-CD, will be the primary instruments to assess the efficacy of SHP647.

#### Rationale for Abdominal Pain Severity

Abdominal pain is one of the most common symptoms of CD, with the cause likely to be multifactorial. In the CDAI, which was the most commonly used primary endpoint in CD studies in the past, the degree of abdominal pain was based on a 4-point scale, with scores ranging from 0 (none) to 3 (severe). However, the new standard is to use the 11-point numerical rating scale (NRS) instead for the degree of abdominal pain. The limitation is that the 4-point and the 11-point scales are not directly comparable. Numerous studies across a variety of conditions have examined cutoff scores for mild, moderate, and severe pain based on the 11-point pain NRS, with the findings across studies generally converging on a cutoff score of 4 (reflecting the maximum score indicating mild pain) and a score of 5 (reflecting the minimum score indicating moderate pain). To ensure that clinical remission criteria for abdominal pain are both clinically meaningful and fall definitively within the mild pain range on the NRS based on the literature, a remission subscore of  $\leq 3$  (a minimum improvement of at least 2 points is required for subjects who enter the study with moderate abdominal pain [subscore of  $\geq 5$ ]) will be used as part of the coprimary endpoints.

For these reasons, the key secondary composite endpoint (which takes into account both clinical and endoscopic response to treatment in the same subject) consists of the evaluation of the clinical remission together with the endoscopic response.

### Complete Endoscopic Healing

Endoscopic healing will be defined in 2 ways:

- Endoscopic healing defined by SES-CD  $\leq 4$  and at least a 2-point reduction versus baseline (Visit 2) and no subscore  $> 1$  in any individual variable
- Complete endoscopic healing defined by SES-CD=0-2.

There is no uniformly accepted definition for endoscopic healing in CD and several different terminologies are used to describe the same endoscopic appearance defined by a certain endoscopic score (eg, endoscopic remission and mucosal healing). Endoscopic healing or mucosal healing is predominantly defined by the absence of mucosal ulcerations in CD during endoscopic assessment of intestinal inflammation ([Atreya and Neurath, 2017](#)). The International Organization for the study of Inflammatory Bowel Disease technical review on endoscopic indices for CD clinical studies defined complete endoscopic healing as SES-CD=0-2 ([Vuitton et al., 2016](#)). Some studies introduced SES-CD  $\leq 4$  as “endoscopic remission”. The more stringent endpoint of “complete endoscopic healing” will be used as a key secondary endpoint in this study. Even in case of complete endoscopic healing, there may still be ongoing histological activity in many cases and it may not always reflect healing of all layers of the tissue, as endoscopy only addresses mucosal rather than transmural healing ([Atreya and Neurath, 2017](#)).

The importance of inducing endoscopic healing is that it may be associated with long-term symptomatic remission; longer relapse-free interval; reduced frequency of hospitalizations, complications, and surgical resections; and the potential for a significant improvement in quality of life ([Peyrin-Biroulet et al., 2011](#)).

### **3.2 Duration and Study Completion Definition**

Each subject's final visit in this study may be at the end of the treatment period (Week 16), if continuing to Study SHP647-307 or SHP647-304, or at the end of the safety follow-up period (Week 32), if not continuing to either of these studies. In either case, the final visit will be in person at the site. A subject's maximum duration of participation is expected to be approximately 38 weeks: a screening period of up to 6 weeks, a treatment period of 16 weeks, and a safety follow-up period of 16 weeks (if applicable). It is expected that the study will be completed in approximately 3 years.

The study completion date is defined as the date the final subject, across all sites, completes their final protocol-defined assessment. Please note that this includes the follow-up visit or contact, whichever is later. The study completion date is used to ascertain timing for study results posting and reporting.

Antidiarrheal opiate drugs such as IMODIUM® (loperamide), LOMOTIL® (diphenoxylate hydrochloride and atropine sulfate), tincture of opium, and codeine will be recorded. Subjects must be using such products in a stable regimen for at least 2 weeks before randomization at baseline (Visit 2). Reported use of any antidiarrheal opiate medicines will assist the investigator response to Question 5 of the CDAI.

Subjects using medicinal marijuana (cannabis) under a physician's prescription, and who obtain the product from a licensed pharmacy or provider, should continue to use it under the same regimen for the duration of the study, unless otherwise instructed by the investigator or treating physician. Such subjects must be using the product, in a stable regimen, for at least 3 months before screening.

Routine nonlive vaccinations are allowed during the study.

Dietary and herbal supplements and probiotics are allowed in the study, provided they are being taken at stable doses at the time of the baseline visit (Visit 2) and for the duration of the study. They should be recorded as concomitant medications.

Use of nicotine-containing preparations should be recorded as concomitant medication.

Antibiotics are permitted, with the exception of antibiotics used to treat the underlying disease or any continuous antibiotic treatment exceeding 2 weeks within 30 days before starting PRO-CD daily e-diary data collection in the screening period or before Week 16 (Visit 7, Part 1).

### 5.2.2 Prohibited Treatment

Table 2 details the minimum required number of days before baseline (Visit 2) for common prior treatments that are excluded medications for this study.

**Table 2: Common Excluded Treatments**

Treatment	Excluded without any timeframe	Minimum Required Number of Days Before Baseline (Visit 2)			
		14 days	30 days	60 days	90 days
SHP647 (PF-00547659) in a previous study	X				
Anti-integrin or antiadhesion molecule treatment (eg, natalizumab, vedolizumab, efalizumab, etrolizumab)	X				
Parenteral and rectal glucocorticoids		X <sup>a</sup>			
Rectal 5-ASA		X <sup>a</sup>			
Investigational products			X <sup>b</sup>		
Live (attenuated) vaccine			X		
Nonbiologics with immunomodulatory properties <sup>c</sup>			X		
Anti-TNF treatment				X	

### 7.2.2.3 Crohn's Disease Activity Index

The CDAI is a composite measure with 8 components; 3 components (abdominal pain severity, very soft stool/liquid stool frequency, and general wellbeing) will be self-reported by the subject and will be recorded as part of the daily e-diary, as described in Section 7.2.2.1 and 5 components will be recorded at the time points specified in [Table 1](#).

The CDAI score at screening (Visit 1, Part 2) will be calculated using the following:

- Components 1 to 3 from subject-reported PRO-CD daily e-diary data collected  $\geq 10$  days before the start of colonoscopy preparation using the same most recent 7 of 10 days as described for the 2-item PRO (Section 7.2.2.1) and
- Components 4 to 8 (weight, medical and physical examination, use of diarrhea treatment, and hematocrit value) collected during screening (Visit 1) Part 1. Note: Hematology samples should be repeated if more than 3 weeks have elapsed before the day of colonoscopy to be able to use the hematocrit central laboratory results for the CDAI score calculation at screening. Hematocrit must not be older than 3 weeks before the day of colonoscopy.

The CDAI scores at Visits 4, 5, and 6 will be calculated using the following:

- Components 1 to 3 from subject-reported PRO-CD daily e-diary data collected  $\geq 10$  days before the visit using the same most recent 7 or 10 days as described for the 2-item PRO (Section 7.2.2.1) and
- Components 4 to 8 (weight, medical and physical examination, use of diarrhea treatment, and hematocrit value) collected at the visit.

The CDAI score at the Week 16/ET visit will be calculated at Visit 7, Part 3 (after all evaluations are complete), using the following:

- Components 1 to 3 from subject-reported PRO-CD daily e-diary data collected  $\geq 10$  days before the start of colonoscopy preparation using the same most recent 7 or 10 days as described for the 2-item PRO (Section 7.2.2.1) and
- Components 4 to 8 (weight, medical and physical examination, use of diarrhea treatment, and hematocrit value) collected at Part 1 and Part 3 of the Week 16/ET visit, where assessed.

Change in CDAI has been used as a primary endpoint in multiple pivotal studies in the CD indication. The algorithm for calculating the CDAI score was first published by William Best and colleagues ([Best et al., 1976](#)).

The CDAI is presented in [Appendix 2](#).

### 7.2.2.4 Colonoscopy and Histology

Colonoscopy will be performed at the time points specified in [Table 1](#).

Virology test results must be confirmed as negative before enrollment in the study; if a virology test result is positive, the subject will be excluded from entering the study. Results of the virology screen will be reviewed and verified by the study monitor but will not be collected in the electronic case report form (eCRF) database.

Stool microbiology will be performed at the screening visit (Visit 1). The detection of *Clostridium difficile* by toxigenic stool culture (stool culture followed by detection of toxin) is considered the gold standard for the diagnosis of the colonization or infection with pathogenic *C. difficile*. Comparable sensitivity may be achieved by direct testing of stool via point-of-use rapid membrane enzyme immunoassay card for both *C. difficile* toxin A and B and glutamate dehydrogenase antigen on a card. Use of the card for point-of-care screening is encouraged where permitted by local regulation; any samples showing a positive result with this method should be sent for toxigenic stool culture. Molecular techniques such as PCR for detection of toxin RNA are also acceptable alternatives. Refer to the laboratory manual for further guidance and instruction for *C. difficile* screening.

A TB test (PPD or QuantiFERON-TB Gold/TB Plus) will be performed at the screening visit (Visit 1). A documented negative PPD test within 12 weeks before baseline (Visit 2) is acceptable provided that an IGRA official reading and method or test is located in the source documentation.

A serum sample will be collected and banked for John Cunningham virus antibody testing at baseline (Visit 2). It may be analyzed if a subject shows neurological symptoms suggestive of PML.

All laboratory assessments should be performed at central laboratories, with the exception of the following assessments: stool microbiology (local or central laboratory) and TB test (PPD or QuantiFERON-TB Gold/TB Plus) (refer to laboratory manual for details).

#### **7.2.3.7      Pregnancy Test and Follicle-stimulating Hormone Test**

A beta-human chorionic gonadotropin ( $\beta$ -hCG) pregnancy test will be performed on all females of childbearing potential at the time points specified in [Table 1](#); if pregnancy is suspected; or on withdrawal of the subject from the study. A serum pregnancy test will be performed at the screening visit (Visit 1); all other pregnancy tests will be urine tests.

Pregnancy tests are not required for females of nonchildbearing potential who have undergone hysterectomy or bilateral oophorectomy, have medically confirmed ovarian failure, or are medically confirmed postmenopausal (cessation of regular menses for at least 12 consecutive months with no alternative pathological or physiological cause; postmenopausal status should be confirmed by FSH testing in females who have had 12 consecutive months of spontaneous amenorrhea and are 51 years of age or older).

#### **7.2.3.8      Electrocardiogram**

A 12-lead ECG will be recorded at the time points specified in [Table 1](#).

### Short Form-36 Health Survey (Version 2, Acute Form)

The SF-36 is a generic quality-of-life instrument that has been widely used to assess HRQL of subjects. Generic instruments are used in general populations to assess a wide range of domains applicable to a variety of health states, conditions, and diseases. The SF-36 consists of 36 items that are aggregated into 8 multi-item scales (physical functioning, role – physical, bodily pain, general health, vitality, social functioning, role – emotional, and mental health), with scores ranging from 0 to 100 ([Ware and Sherbourne, 1992](#)). Higher scores indicate better HRQL.



#### 7.2.4.4 Healthcare Resource Utilization Assessments

Hospitalizations, inpatient days, and [REDACTED] will be recorded at the time points specified in [Table 1](#). Information regarding Crohn's disease-related and other surgeries will be collected from subjects during the treatment period.

#### 7.2.5 Volume of Blood to Be Drawn From Each Subject

The volume of blood to be drawn from each subject is summarized in [Table 4](#).

**Table 4: Volume of Blood to Be Drawn From Each Subject**

Assessment	Sample Volume (mL)	Number of Samples	Total Volume (mL)
Hematology	2	7	14
Serum chemistry	6	7	42
HBsAg	2	1	2
HBcAb	2	1	2
HCV Ab	2	1	2
FSH	2	1	2
Serum β-hCG <sup>a</sup>	2	1	2
TB test (QuantiFERON-TB Gold/TB Plus or PPD)	3	1	3
JCV antibody banked sample	3.5	1	3.5
[REDACTED]			
[REDACTED]	2	3	6
[REDACTED]	5	3	15

## 8. ADVERSE AND SERIOUS ADVERSE EVENTS ASSESSMENT

### 8.1 Definition of Adverse Events, Period of Observation, Recording of Adverse Events

An AE is any untoward medical occurrence in a clinical investigation subject administered a pharmaceutical product and that does not necessarily have a causal relationship with this treatment. An AE can therefore be any unfavorable and unintended sign (including an abnormal laboratory finding), symptom, or disease temporally associated with the use of a medicinal (investigational) product, whether or not related to the medicinal (investigational) product (International Council for Harmonisation [ICH] Guidance E2A [1995]).

All AEs are collected from the time the informed consent and/or assent is signed until the end of the defined follow-up period stated in Section 7.1.3. This includes events occurring during the screening phase of the study, regardless of whether or not investigational product is administered. Where possible, a diagnosis rather than a list of symptoms should be recorded. If a diagnosis has not been made, then each symptom should be listed individually. All AEs should be captured in the subject's source document. In addition to untoward AEs, unexpected benefits outside the investigational product indication should also be captured in the subject's source document.

All AEs must be followed to closure (the subject's health has returned to his/her baseline status or all variables have returned to normal), regardless of whether the subject is still participating in the study. Closure indicates that an outcome is reached, stabilization achieved (the investigator does not expect any further improvement or worsening of the event), or the event is otherwise explained. When appropriate, medical tests and examinations are performed so that resolution of event(s) can be documented.

#### 8.1.1 Severity Categorization

The severity of AEs must be recorded during the course of the event including the start and stop dates for each change in severity. An event that changes in severity should be captured as a new event. Worsening of pretreatment events, after initiation of investigational product, must be recorded as new AEs (for example, if a subject experiences mild intermittent dyspepsia before dosing of investigational product, but the dyspepsia becomes severe and more frequent after first dose of investigational product has been administered, a new AE of severe dyspepsia [with the appropriate date of onset] is recorded in the subject's source document).

The medical assessment of severity is determined by using the following definitions:

**Mild:** A type of AE that is usually transient and may require only minimal treatment or therapeutic intervention. The event does not generally interfere with usual activities of daily living.

**Moderate:** A type of AE that is usually alleviated with specific therapeutic intervention. The event interferes with usual activities of daily living, causing discomfort but poses no significant or permanent risk of harm to the research subject.

**Severe:** A type of AE that interrupts usual activities of daily living, or significantly affects clinical status, or may require intensive therapeutic intervention.

## **9. DATA MANAGEMENT AND STATISTICAL METHODS**

### **9.1 Data Collection**

The investigators' authorized site personnel must enter the information required by the protocol in the eCRF. A study monitor will visit each site in accordance with the monitoring plan and review the eCRF data against the source data for completeness and accuracy. Discrepancies between source data and data entered in the eCRF will be addressed by qualified site personnel. When a data discrepancy warrants correction, the correction will be made by authorized site personnel. Data collection procedures will be discussed with the site at the site initiation visit and/or at the investigator's meeting. Once a subject is randomized, it is expected that site personnel will complete the eCRF entry within approximately 3 business days of the subject's visit.

### **9.2 Clinical Data Management**

Data are to be entered into a clinical database as specified in the CRO's data management plan. Quality control and data validation procedures are applied to ensure the validity and accuracy of the clinical database.

Data are to be reviewed and checked for omissions, errors, and values requiring further clarification using computerized and manual procedures. Data queries requiring clarification are to be communicated to the site for resolution. Only authorized personnel will make corrections to the clinical database, and all corrections are documented in an auditable manner.

### **9.3 Data Handling Considerations**

Data that may potentially unblind the treatment assignment (ie, investigational product serum concentrations, antibodies to investigational product, treatment allocation, and investigational product 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.

### **9.4 Statistical Analysis Process**

The study will be analyzed by the sponsor or its agent.

The statistical analysis plan (SAP) will provide the statistical methods and definitions for the analysis of the efficacy and safety data, as well as describe the approaches to be taken for summarizing other study information such as subject disposition, demographics and baseline characteristics, investigational product exposure, and prior and concomitant medications. The SAP will also include a description of how missing, unused and spurious data will be addressed. The SAP will be finalized before unblinding to preserve the integrity of the statistical analysis and study conclusions.

**Table 5: Power to Detect the Corresponding Treatment Effect for Key Secondary Endpoints**

Key Secondary Endpoint at Week 16	SHP647 Premise	Placebo Premise	Power
Clinical remission by CDAI	26.5%	15%	0.90
Enhanced endoscopic response	25%	13%	0.94
Clinical remission by abdominal pain $\leq 1$ and stool frequency $\leq 3$	24%	14%	0.81
Clinical response by 2-item PRO	52.5%	40%	0.81
Clinical remission by 2-item PRO and endoscopic response	11%	4.5%	0.77
Complete endoscopic healing	6%	2%	0.58

CDAI=Crohn's Disease Activity Index; PRO=patient-reported outcome.

## 9.7 Study Population

The screened set will consist of all subjects who have signed an informed consent document.

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

The safety set will consist of all subjects who have received at least 1 dose of investigational product.

The full-analysis set (FAS) will consist of all subjects in the randomized set who have received at least 1 dose of investigational product.

The per-protocol set will consist of all subjects in the FAS who do not have protocol deviations that may affect the coprimary efficacy endpoints.

The completer set will consist of all subjects in the FAS who have completed the Week 16 assessment for this study.

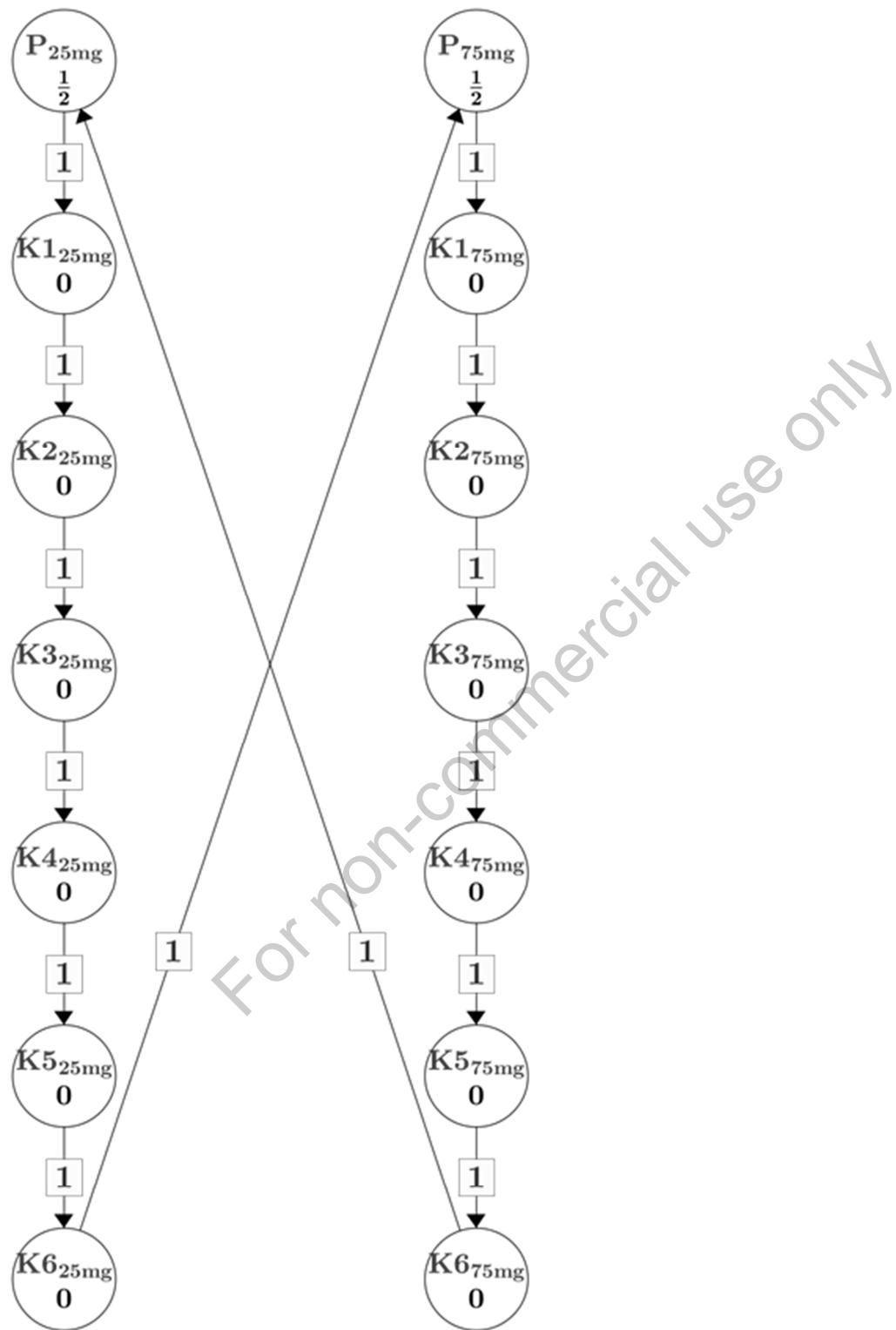
[REDACTED]

[REDACTED]

## 9.8 Efficacy Analyses

Unless otherwise specified, all efficacy analyses will be based on the FAS and subjects will be analyzed according to their randomized treatment, regardless of the treatment they actually received.

Figure 4: Visualization of Multiple Testing Strategy



K=key secondary endpoint; P=coprimay endpoints.

[REDACTED]

[REDACTED]

[REDACTED]

For non-commercial use only

### Crohn's Disease Activity Index (CDAI)

## Crohn's Disease Activity Index (CDAI)

Variable No.	Variable Description	Multiplier	Total
1	No. of liquid or soft stools (each day for 7 days)	X 2	
2	Abdominal pain (0 = none, 1 = mild, 2 = moderate, 3 = severe)	X 5	
3	General well-being (0 = generally well, 1 = slightly under par, 2 = poor, 3 = very poor, 4 = terrible)	X 7	
4	Number of listed complications [arthritis or arthralgia, iritis or uveitis, erythema nodosum or pyoderma gangrenosum or aphthous stomatitis, anal fissure or fistula or abscess, other fistula, fever over 37.8°C (100°F)]	X 20	
5	Use of diphenoxylate or loperamide for diarrhea (0 = no, 1 = yes)	X 30	
6	Abdominal mass (0 = no, 2 = questionable, 5 = definite)	X 10	
7	Hematocrit [Males: 47-Hct (%), Females: 42-Hct (%)]	X 6	
8	Body weight (1-weight/standard weight) X 100 (add or subtract according to sign)	X 1	
<b>CDAI Score</b>			

CDAI=Crohn's Disease Activity Index; ePRO=electronic patient-reported outcome; Hct=hematocrit

Note: Variable 5: This variable covers taking medication for symptomatic relief from diarrhea, eg, bulking agents, opiates etc.

Variable 7: Absolute deviation of hematocrit is the difference in hematocrit from standard. A male subject with a hematocrit of 40% has an absolute deviation of 7. Each percentage deviation has a value of 6 points. If hematocrit subtotal is <0, enter 0.

Variable 8: This variable is based on Metropolitan Life Tables (these are programmed into the ePRO device).

Percent deviation from standard weight is  $(1 - \text{weight}/\text{standard weight}) \times 100$ ; therefore, positive percent deviation represents weight loss, which adds points to the CDAI. Percentage deviation from standard weight = 1 point for each percent deviation. If body weight subtotal is less than -10, enter -10.

CDAI Interpretation:

- 0-149 points: Asymptomatic remission (Note: subjects requiring steroids to remain asymptomatic are not considered to be in remission but are referred to as being "steroid dependent")
- 150-220 points: Mild to moderate active Crohn's disease
- 221-450 points: Moderate to severe active Crohn's disease
- >451 points: Severely active to fulminant disease.

CDAI online estimator: <http://www.ibdjohn.com/cdai/>

Sources: Best et al., 1976; Best et al., 1979.

10. How would you rate your general well-being over the past 24 hours?

Generally well

Slightly below par

Poor

Very poor

Terrible

For non-commercial use only

6. How much energy have you had during the last 2 weeks? Please choose an option from:

1. NO ENERGY AT ALL
2. VERY LITTLE ENERGY
3. A LITTLE ENERGY
4. SOME ENERGY
5. A MODERATE AMOUNT OF ENERGY
6. A LOT OF ENERGY
7. FULL OF ENERGY

7. How often during the last 2 weeks did you feel worried about the possibility of needing to have surgery because of your bowel problem? Please choose an option from:

1. ALL OF THE TIME
2. MOST OF THE TIME
3. A GOOD BIT OF THE TIME
4. SOME OF THE TIME
5. A LITTLE OF THE TIME
6. HARDLY ANY OF THE TIME
7. NONE OF THE TIME

8. How often during the last 2 weeks have you had to delay or cancel a social engagement because of your bowel problem? Please choose an option from:

1. ALL OF THE TIME
2. MOST OF THE TIME
3. A GOOD BIT OF THE TIME
4. SOME OF THE TIME
5. A LITTLE OF THE TIME
6. HARDLY ANY OF THE TIME
7. NONE OF THE TIME

Short Form-36 Health Survey (Version 2), Acute Form

---

## Your Health and Well-Being

---

This survey asks for your views about your health. This information will help keep track of how you feel and how well you are able to do your usual activities.  
*Thank you for completing this survey!*

For each of the following questions, please mark an  in the one box that best describes your answer.

1. In general, would you say your health is:

Excellent	Very good	Good	Fair	Poor
<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5

2. Compared to one week ago, how would you rate your health in general now?

Much better now than one week ago	Somewhat better now than one week ago	About the same as one week ago	Somewhat worse now than one week ago	Much worse now than one week ago
<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5

[REDACTED]

[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]

[REDACTED]

[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]

[REDACTED]

[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]

[REDACTED]

[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]

**Table 1: Schedule of Assessments**

Study Procedure	Screening <sup>a</sup>		Baseline	Treatment					Follow-up			
	Weeks -6 to -1		Week 0/ Day 1	Week 2	Week 4	Week 8	Week 12	Week 16/ET <sup>b</sup>		Week 24 <sup>c</sup>	Week 32 <sup>c</sup>	
Visit Number	1 (Part 1) <sup>a</sup>	1 (Part 2) <sup>a</sup>	2	3	4	5	6	7 (Part 1) <sup>b</sup>	7 (Part 2) <sup>b</sup>	7 (Part 3) <sup>b</sup>	8	9
Study Day	-42 to 0		1	14 ±3	28 ±3	56 ±3	84 ±3			112 ±3	168 ±7	224 ±7
<b>Treatment Procedures</b>												
Randomization <sup>x</sup>			X									
Administration of SHP647 or placebo <sup>x,y</sup>			X		X	X	X					
Adverse events	X	X	X	X	X	X	X	X	X	X	X	
Prior medications	X											
Concomitant medications and procedures	X	X	X	X	X	X	X	X	X	X	X	
Dispense stool collection kit for stool sample <sup>z</sup>	X					X	X					

Ab=antibody; ADA=antidrug antibodies; β-hCG=beta-human chorionic gonadotropin; CD=Crohn's disease; CDAI=Crohn's Disease Activity Index; [REDACTED]; ECG=electrocardiogram; [REDACTED]; [REDACTED]; ET=early termination; FSH=follicle-stimulating hormone; GDH=glutamate dehydrogenase; HBcAb=hepatitis B core antibody; HBsAg=hepatitis B surface antigen; HCV=hepatitis C virus; HCV RNA=hepatitis C virus ribonucleic acid; HRUA=Healthcare Resource Utilization Assessment; IBD=inflammatory bowel disease; IBDQ=Inflammatory Bowel Disease Questionnaire; JCV=John Cunningham virus; LTS=long-term safety extension; [REDACTED]; NAb=neutralizing antibody; PCR=polymerase chain reaction; PGA=physician's global assessment; [REDACTED]; PML=progressive multifocal leukoencephalopathy; PPD=purified protein derivative; PRO=patient-reported outcomes; SES-CD=Simple Endoscopic Score for CD; SF-36 v2=Short Form-36 Health Survey, version 2; TB=tuberculosis; [REDACTED].

<sup>a</sup> Screening assessments will take place over more than 1 day (at least 2 visits will be necessary to complete the screening evaluations, including colonoscopy).

<sup>b</sup> Subjects who withdraw early during the treatment period should return for the ET visit and then enter into the safety follow-up period. The Week 16 (Visit 7) and ET visits consist of 3 parts:

- Part 1 of Visit 7 should be scheduled 1 to 3 day(s) before Part 2; this will allow for blood samples to be taken before starting the colonoscopy preparation and before the colonoscopy procedure at Part 2 of the visit (Section 7.2.2.4)
- Part 2 of Visit 7 should be scheduled preferably 5 to 7 days before Part 3; this will allow sufficient time to obtain the data from the centrally read colonoscopy
- Part 3 of Visit 7 will take place on Day 112 ±3 days.

<sup>c</sup> Subjects NOT entering the maintenance study (SHP647-307) or LTS(SHP647-304) study at the completion of the Week 16 visit will need to complete the safety follow-up assessments. The Week 24 (Visit 8) 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 32 (Visit 9) visit will take place at the study site.

<sup>d</sup> The outcome of Visit 7, Part 3 is used to assess eligibility to enroll subjects in the maintenance (SHP647-307) or LTS(SHP647-304) studies. Please refer to the respective protocols for further details.

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

Sensitivity analyses to explore the impact of missing data on the coprimary efficacy endpoints and key secondary endpoints will be conducted. These analyses may compare imputations of the missing values which favor placebo (eg, worst case) and/or imputations which favor active treatment (eg, best case). In addition, imputation methods based on informative missingness and other missing data mechanisms may be performed. Additional sensitivity analyses will also be conducted using the per-protocol set and the completer set. Additional analyses may be developed in the SAP. All sensitivity analyses will be described in the SAP.

Prespecified subgroup analyses are planned for the coprimary endpoints including, but not limited to gender, prior anti-TNF treatment, glucocorticoid use at baseline, SES-CD at baseline, region, age group, randomization stratum, and other important subgroups. A full list of important subgroups will be described in the SAP. Within subgroups, efficacy endpoints will be compared for each active treatment group (25 mg SHP647 and 75 mg SHP647) with the placebo group using a chi-square test. The estimate of the treatment difference, along with the corresponding Newcombe (hybrid-score) 95% CI and chi-square test p-value, will be presented.

#### Statistical Testing and Protection of the Type I Error

The global FWER for the statistical tests of the coprimary and key secondary endpoints will be strongly controlled at .05 (2-sided). To control the FWER, graphical methods discussed in Bretz et al. (2009b) will be utilized to propagate  $\alpha$  from the coprimary endpoints to the key secondary endpoints and between the 2 SHP647 treatment group and placebo comparisons. Alpha is initially split equally at the .025 level (2-sided) for each of the pairwise treatment comparisons for the coprimary endpoints (P) and alpha is propagated in a hierarchical manner to each of the 6 key secondary endpoints (K1-K6) within a pairwise treatment comparison. In order to pass alpha between the coprimary endpoints and the first key secondary endpoint, both coprimary endpoints must attain statistical significance. A graphical visualization of the  $\alpha$  propagation is presented in Figure 4.

Some treatments may result in a partial initial response, even though at a later stage a complete response may occur. Median duration of remission after 1 year treatment with infliximab was similar in subjects achieving complete absence of mucosal ulcer to subjects who achieved significant but incomplete mucosal healing (D'Haens et al., 2002). The benefit of endoscopic response was also shown in the SONIC study; the presence of endoscopic response (defined in that study as at least a 50% decrease in endoscopic score at Week 26 of treatment) identified subjects most likely to be in corticosteroid-free clinical remission at Week 50 (Ferrante et al., 2013). The proportion of patients requiring major abdominal surgery in a single-center cohort study with infliximab was similar with complete healing or with partial healing. (Schnitzler et al., 2009; Panaccione et al., 2013). Subjects with such a treatment response should be identified by endoscopic assessment in order not to misclassify them as nonresponders and underestimate the response to the treatment.

### 3.1.2 Rationale for Key Secondary Endpoints

#### Clinical Remission Defined by CDAI Score

Conventionally, a CDAI score of <150 has been used to define clinical remission. While there has been widespread use of the CDAI over a long period of time, the items do not contribute equally to the score, and symptom items reported by subjects are not specific for CD and are not sensitive for inflammation seen at colonoscopy. There has been movement away from using the CDAI by regulatory authorities to the use of PROs and objective measures of disease such as endoscopy (Williet et al., 2014). However, for benchmarking or for comparative effectiveness purposes, CDAI endpoints are expected to be used.

Although this has been the established gold-standard for clinical remission to date, CDAI suffers from requiring complex calculations across 8 individual items including subjective elements.

#### Clinical Remission Defined by Average Daily Abdominal Pain $\leq 1$ (Based on the 4-point Scale) and Average Daily Stool Frequency $\leq 3$ of Type 6/7

The CDAI has been the traditionally used measure to assess clinical response and clinical remission in CD. In the CDAI, the degree of abdominal pain is one of 8 variables and is used with a multiplier of 5 in the overall score. Importantly, it is based on a 4-point scale, with scores ranging from 0 (none) to 3 (severe). With the shift to the new endpoint as evident from the coprimary endpoint of clinical remission in this study, it is still important to allow for a frame of reference to the existing standard for response, based on the CDAI components. A daily average abdominal pain threshold of  $\leq 1$  will help achieve this as 1 on the 4-point scale corresponds to mild abdominal pain. Although direct mapping between the scales has not been established, this will approximate to a score of 3 on the 11-point NRS scale, as this falls within the mild pain range on the NRS based on the literature.

Based on post hoc analyses of the Phase 2 data in the SHP647 program, regulatory requirements, and treatment separation assumptions, a threshold for the average daily stool frequency  $\leq 2$  of type 6/7 (very soft stools/liquid stools) was chosen for the coprimary endpoint of clinical remission. However, given the limited data available for this endpoint, recent evidence from literature suggesting that thresholds  $\leq 3$  are likely to be quite stringent (Sandborn et al., 2017), and the refractory nature of the disease in those with moderate to severe CD, it is also important to assess the effects of treatment using a more realistic measure.

Screening laboratory tests, if considered by the investigator to be transient and inconsistent with the subject's clinical condition, may be repeated once during the screening period for confirmation. The following clinical laboratory assessments will be performed at the time points specified in [Table 1](#).

### Serum chemistry

- alkaline phosphatase
- aspartate aminotransferase
- alanine aminotransferase
- 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

### Virology

- hepatitis B surface antigen (HBsAg)
- hepatitis B core antibody (HBcAb)
- hepatitis B DNA reflex testing if HBsAg is negative and HBcAb is positive
- hepatitis C virus antibody (HCVAb)
- HCV ribonucleic acid polymerase chain reaction if HCVAb is positive
- HIV

### Urinalysis

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

Virology test results must be confirmed as negative before enrollment in the study; if a virology test result is positive, the subject will be excluded from entering the study. Results of the virology screen will be reviewed and verified by the study monitor but will not be collected in the electronic case report form (eCRF) database.

## 2.2 Study Objectives

### 2.2.1 Coprimary Objectives

The coprimary objectives of this study are to evaluate the efficacy of ontamalimab in subjects with moderate to severe CD in:

- Inducing clinical remission based on 2-item patient-reported outcome (PRO) (abdominal pain severity and very soft stool/liquid stool frequency)
- Inducing endoscopic response based on centrally read colonoscopy.

### 2.2.2 Secondary Objectives

#### 2.2.2.1 Key Secondary Objectives

The key secondary objectives are as follows:

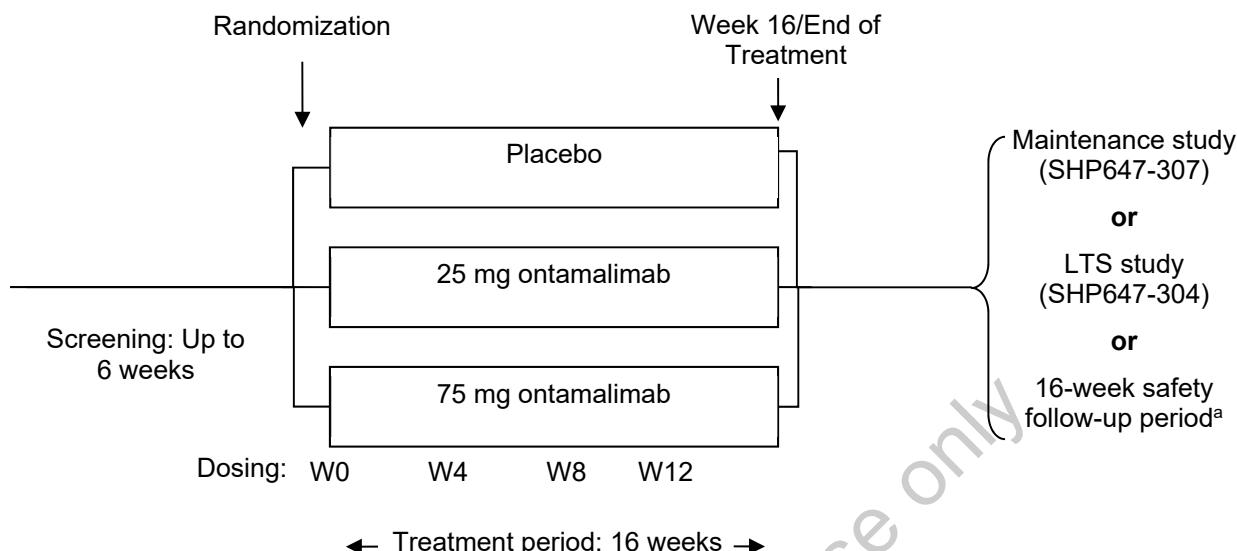
- To evaluate the efficacy of ontamalimab in inducing clinical remission as measured by CDAI
- To evaluate the efficacy of ontamalimab in inducing enhanced endoscopic response based on centrally read colonoscopy
- To evaluate the efficacy of ontamalimab in inducing clinical remission based on abdominal pain severity and very soft stool/liquid stool frequency (alternate thresholds)
- To evaluate the efficacy of ontamalimab in inducing clinical response based on patient-reported clinical signs and symptoms (as measured by 2-item PRO)
- To evaluate the efficacy of ontamalimab in inducing clinical remission based on patient-reported clinical signs and symptoms (as measured by 2-item PRO) as well as inducing endoscopic response based on centrally read colonoscopy in the same subject
- To evaluate the efficacy of ontamalimab in inducing endoscopic healing based on centrally read colonoscopy.

#### 2.2.2.2 Other Secondary Objectives

The other secondary objectives are as follows:

- To evaluate the safety and tolerability of ontamalimab
- To evaluate the effect of ontamalimab induction treatment on other clinical outcomes (clinical response defined by CDAI, or clinical remission over time, or change from baseline in frequency in CD-related clinical parameters)
- To evaluate the effect of ontamalimab induction treatment on other endoscopic outcomes
- To evaluate the effect of ontamalimab on health-related quality of life (HRQL) (as measured by the Inflammatory Bowel Disease Questionnaire [IBDQ] and the Short Form-36 Health Survey [SF-36])
- To evaluate the effect of ontamalimab on incidence of hospitalizations and total inpatient days
- To evaluate the impact of ontamalimab on incidence of CD-related and other surgeries.

**Figure 2 Study Design Flow Chart**



LTS=long-term safety extension; W=week

<sup>a</sup> Subjects who withdraw early from the 16-week treatment period or who do not wish to enter the maintenance study (SHP647-307) or LTS study (SHP647-304) will continue into a 16-week safety follow-up period.

Note: A planned interim analysis for the coprimary endpoints will take place after approximately the first 50% of all randomized subjects in both the SHP647-305 and SHP647-306 studies have either completed the studies or have prematurely withdrawn from the studies.

### 3.1.1 Rationale for Coprimary Endpoints

In this study, clinical remission, as measured by a decrease below prespecified thresholds in the 2-item PRO (abdominal pain severity and very soft stool/liquid stool frequency [as shown in the Bristol Stool Form Scale, BSFS]), and enhanced endoscopic response, as measured by a decrease in SES-CD, will be the primary instruments to assess the efficacy of ontamalimab.

#### Rationale for Abdominal Pain Severity

Abdominal pain is one of the most common symptoms of CD, with the cause likely to be multifactorial. In the CDAI, which was the most commonly used primary endpoint in CD studies in the past, the degree of abdominal pain was based on a 4-point scale, with scores ranging from 0 (none) to 3 (severe). However, the new standard is to use the 11-point numerical rating scale (NRS) instead for the degree of abdominal pain. The limitation is that the 4-point and the 11-point scales are not directly comparable. Numerous studies across a variety of conditions have examined cutoff scores for mild, moderate, and severe pain based on the 11-point pain NRS, with the findings across studies generally converging on a cutoff score of 4 (reflecting the maximum score indicating mild pain) and a score of 5 (reflecting the minimum score indicating moderate pain). To ensure that clinical remission criteria for abdominal pain are both clinically meaningful and fall definitively within the mild pain range on the NRS based on the literature, a remission subscore of  $\leq 3$  (a minimum improvement of at least 2 points is required for subjects who enter the study with moderate abdominal pain [subscore of  $\geq 5$ ]) will be used as part of the coprimary endpoints. This definition is further supported by a study conducted in a similar

response, even though at a later stage a complete response may occur. Median duration of remission after 1 year treatment with infliximab was similar in subjects achieving complete absence of mucosal ulcer to subjects who achieved significant but incomplete mucosal healing (D'Haens et al., 2002).

The benefit of endoscopic response was also shown in the SONIC study; the presence of endoscopic response (defined in that study as at least a 50% decrease in endoscopic score at Week 26 of treatment) identified subjects most likely to be in corticosteroid-free clinical remission at Week 50 (Ferrante et al., 2013). The proportion of patients requiring major abdominal surgery in a single-center cohort study with infliximab was similar with complete healing or with partial healing. (Schnitzler et al., 2009; Panaccione et al., 2013). Subjects with such a treatment response should be identified by endoscopic assessment in order not to misclassify them as nonresponders and underestimate the response to the treatment.

### 3.1.2 Rationale for Key Secondary Endpoints

#### Clinical Remission Defined by CDAI Score

Conventionally, a CDAI score of <150 has been used to define clinical remission. While there has been widespread use of the CDAI over a long period of time, the items do not contribute equally to the score, and symptom items reported by subjects are not specific for CD and are not sensitive for inflammation seen at colonoscopy. There has been movement away from using the CDAI by regulatory authorities to the use of PROs and objective measures of disease such as endoscopy (Williet et al., 2014). However, for benchmarking or for comparative effectiveness purposes, CDAI endpoints are expected to be used.

Although this has been the established gold-standard for clinical remission to date, CDAI suffers from requiring complex calculations across 8 individual items including subjective elements.

#### Clinical Remission Defined by Average Daily Abdominal Pain $\leq 1$ (Based on the 4-point Scale) and Average Daily Stool Frequency $\leq 3$ of Type 6/7

The CDAI has been the traditionally used measure to assess clinical response and clinical remission in CD. In the CDAI, the degree of abdominal pain is one of 8 variables and is used with a multiplier of 5 in the overall score. Importantly, it is based on a 4-point scale, with scores ranging from 0 (none) to 3 (severe). With the shift to the new endpoint as evident from the coprimary endpoint of clinical remission in this study, it is still important to allow for a frame of reference to the existing standard for response, based on the CDAI components. A daily average abdominal pain threshold of  $\leq 1$  will help achieve this as 1 on the 4-point scale corresponds to mild abdominal pain. Although direct mapping between the scales has not been established, this will approximate to a score of 3 on the 11-point NRS scale, as this falls within the mild pain range on the NRS based on the literature.

Based on post hoc analyses of the Phase 2 data in the ontamalimab program, regulatory requirements, and treatment separation assumptions, a threshold for the average daily stool frequency  $\leq 2$  of type 6/7 (very soft stools/liquid stools) was chosen for the coprimary endpoint of clinical remission. However, given the limited data available for this endpoint, recent evidence

mucosa may take a long time and may depend on the baseline severity of the endoscopic appearance, which may not be in line with the actual baseline severity of the symptoms. Therefore, when evaluating clinical remission together with endoscopic endpoints, improvement in endoscopic scores could be more relevant than evaluating mucosal healing in the induction phase. For these reasons, the key secondary composite endpoint (which takes into account both clinical and endoscopic response to treatment in the same subject) consists of the evaluation of the clinical remission together with the endoscopic response.

### Complete Endoscopic Healing

Endoscopic healing will be defined in 2 ways:

- Endoscopic healing defined by SES-CD  $\leq 4$  and at least a 2-point reduction versus baseline (Visit 2) and no subscore  $> 1$  in any individual variable
- Complete endoscopic healing defined by SES-CD=0-2.

There is no uniformly accepted definition for endoscopic healing in CD and several different terminologies are used to describe the same endoscopic appearance defined by a certain endoscopic score (eg, endoscopic remission and mucosal healing). Endoscopic healing or mucosal healing is predominantly defined by the absence of mucosal ulcerations in CD during endoscopic assessment of intestinal inflammation ([Atreya and Neurath, 2017](#)). The International Organization for the study of Inflammatory Bowel Disease technical review on endoscopic indices for CD clinical studies defined complete endoscopic healing as SES-CD=0-2 ([Vuitton et al., 2016](#)). Some studies introduced SES-CD  $\leq 4$  as “endoscopic remission”. The more stringent endpoint of “complete endoscopic healing” will be used as a key secondary endpoint in this study. Even in case of complete endoscopic healing, there may still be ongoing histological activity in many cases and it may not always reflect healing of all layers of the tissue, as endoscopy only addresses mucosal rather than transmural healing ([Atreya and Neurath, 2017](#)).

The importance of inducing endoscopic healing is that it may be associated with long-term symptomatic remission; longer relapse-free interval; reduced frequency of hospitalizations, complications, and surgical resections; and the potential for a significant improvement in quality of life ([Peyrin-Biroulet et al., 2011](#)).

### **3.2 Duration and Study Completion Definition**

Each subject's final visit in this study may be at the end of the treatment period (Week 16), if continuing to Study SHP647-307 or SHP647-304, or at the end of the safety follow-up period (Week 32), if not continuing to either of these studies. In either case, the final visit will be in person at the site. A subject's maximum duration of participation is expected to be approximately 38 weeks: a screening period of up to 6 weeks, a treatment period of 16 weeks, and a safety follow-up period of 16 weeks (if applicable). It is expected that the study will be completed in approximately 3 years.

Antidiarrheal opiate drugs such as IMODIUM® (loperamide), LOMOTIL® (diphenoxylate hydrochloride and atropine sulfate), tincture of opium, and codeine will be recorded as concomitant medications. Subjects must be using such products in a stable regimen for at least 2 weeks before randomization at baseline (Visit 2). Reported use of any antidiarrheal opiate medicines will assist the investigator response to Question 5 of the CDAI. Antidiarrheal opiate drugs must be taken at stable doses for the duration of the study unless dose reduction or discontinuation is required due to clinical improvement of AE. However, escalations of the dose after dose reduction or re-initiation after drug discontinuation are not allowed (see Section 5.2.3).

Subjects using medicinal marijuana (cannabis) under a physician's prescription, and who obtain the product from a licensed pharmacy or provider, should continue to use it under the same regimen for the duration of the study, unless otherwise instructed by the investigator or treating physician. Such subjects must be using the product, in a stable regimen, for at least 3 months before screening.

Routine nonlive vaccinations are allowed during the study.

Dietary and herbal supplements and probiotics are allowed in the study, provided they are being taken at stable doses at the time of the baseline visit (Visit 2) and for the duration of the study. They should be recorded as concomitant medications.

Use of nicotine-containing preparations should be recorded as concomitant medication.

Antibiotics are permitted, with the exception of antibiotics used to treat the underlying disease exceeding 2 weeks within 30 days before baseline (Visit 2) or before Week 16 (Visit 7, Part 1).

## 5.2.2 Prohibited Treatment

Table 2 details the minimum required number of days before baseline (Visit 2) for common prior treatments that are excluded medications for this study.

Evidence of active inflammation and ulceration is required at screening (Visit 1), in the form of a centrally read score of at least 1 in one or more ileocolonic segments in the Presence of Ulcers component of the SES-CD, as well as a total score of >6.

Study videos will be scored separately by 2 central readers who are blinded to the treatment. If the central readers' scores are not in agreement, there will be a third adjudication read to select the correct read from the first 2 scores. Results of the central reading of the videos will be communicated to sites within 5 business days.

For the evaluation of efficacy, in cases where 1 or 2 segments cannot be fully evaluated by central endoscopic readers, ileocolonic segments that are evaluable during screening (Visit 1) and Week 16/ET (matching segments approach) will be utilized.

The SES-CD is presented in [Appendix 2](#).

#### 7.2.2.3 Crohn's Disease Activity Index

The CDAI is a composite measure with 8 components; 3 components (abdominal pain severity, very soft stool/liquid stool frequency, and general well-being) will be self-reported by the subject and will be recorded as part of the daily e-diary, as described in Section [7.2.2.1](#) and 5 components will be recorded at the time points specified in [Table 1](#).

The CDAI score at screening (Visit 1, Part 2) will be calculated using the following:

- Components 1 to 3 from subject-reported PRO-CD daily e-diary data collected  $\geq 10$  days before the start of colonoscopy preparation using the same most recent 7 of 10 days as described for the 2-item PRO (see Section [7.2.2.1](#)) and
- Components 4 to 8 (weight, medical and physical examination, use of diarrhea treatment, and hematocrit value) collected during screening (Visit 1) Part 1. Note: Hematology samples should be repeated if more than 3 weeks have elapsed before the day of colonoscopy to be able to use the hematocrit central laboratory results for the CDAI score calculation at screening. Hematocrit must not be older than 3 weeks before the day of colonoscopy.

Note that the subject must be confirmed as meeting the CDAI score requirements at screening **before** a colonoscopy is done.

The CDAI scores at Visits 4, 5, and 6 will be calculated using the following:

- Components 1 to 3 from subject-reported PRO-CD daily e-diary data collected  $\geq 10$  days before the visit using the same most recent 7 or 10 days as described for the 2-item PRO (see Section [7.2.2.1](#)) and
- Components 4 to 8 (weight, medical and physical examination, use of diarrhea treatment, and hematocrit value) collected at the visit.

Pregnancy tests are not required for females of nonchildbearing potential who have undergone hysterectomy or bilateral oophorectomy, have medically confirmed ovarian failure, or are medically confirmed postmenopausal (cessation of regular menses for at least 12 consecutive months with no alternative pathological or physiological cause; postmenopausal status should be confirmed by FSH testing in females who have had 12 consecutive months of spontaneous amenorrhea and are 51 years of age or older).

#### **7.2.3.8      Electrocardiogram**

A 12-lead ECG will be recorded at the time points specified in [Table 1](#). When timing of measurements coincide, ECGs should be performed before laboratory blood collection and endoscopic procedure.

A central ECG reader will be used in this study. The eligibility of the subject is based on the assessment of the ECG by the investigator. If abnormal results are observed following assessment by the central reader, the investigator, in consultation with the appointed sponsor or contract research organization (CRO) medical monitor, reconfirms subject eligibility to continue.

#### **7.2.3.9      Chest X-ray**

A chest x-ray will be performed during screening (Visit 1). If a subject has had a chest x-ray performed as a part of standard medical care within 12 weeks before screening (Visit 1), it does not need to be repeated as a part of screening. The official reading must be located in the subject's source documentation.

#### **7.2.3.10     Antidrug Antibodies**

Blood samples for measurement of ADAs and NAbs will be collected at the time points specified in [Table 1](#). Blood samples must be collected before administration of investigational product at that visit.

#### **7.2.3.11     Monitoring for Type I and Type III Immune Reactions**

Subjects will be educated on the signs and symptoms of hypersensitivity reactions and how to respond to them. In addition, subjects will be instructed to report hypersensitivity AEs to the investigator at the time of occurrence, and to seek immediate medical care if hypersensitivity develops. At each visit, the subject will be queried for AEs of special interest (AESIs) related to hypersensitivity.

Subjects will be also instructed to report AEs such as serum-sickness, vasculitis, Arthus reaction, and severe injection-related reactions to the investigator, and to seek immediate medical care if these events are severe in intensity.

Subjects who experience a hypersensitivity reaction or severe or serious injection-related reaction (eg, shortness of breath, wheezing, stridor, angioedema, life-threatening change in vital signs) should discontinue investigational product until the adjudication committee assesses the case and finalizes recommendation of permanent discontinuation or rechallenge with investigational product.

21 Aug 2018

#### **7.2.4.4 Healthcare Resource Utilization Assessments**

Hospitalizations, inpatient days, and [REDACTED] will be recorded at the time points specified in **Table 1**. Information regarding Crohn's disease-related and other surgeries will be collected from subjects during the treatment period.

### **7.2.5 Volume of Blood to Be Drawn From Each Subject**

The volume of blood to be drawn from each subject is summarized in Table 4.

**Table 4:** Volume of Blood to Be Drawn From Each Subject

Assessment	Sample Volume (mL)	Number of Samples	Total Volume (mL)
Hematology	2	7	14
Serum chemistry	6	7	42
HBsAg	2	1	2
HBcAb	2	1	2
HCV Ab	2	1	2
HBV DNA	6	1	6
HIV	2	1	2
FSH	2	1	2
Serum β-hCG <sup>a</sup>	2	1	2
TB test (QuantiFERON-TB Gold Plus or PPD)	4	1	4
JCV antibody banked sample	3.5	1	3.5
[REDACTED]			
[REDACTED]	2	3	6
[REDACTED]	5	3	15
[REDACTED]	4	3	12
[REDACTED]	3	6	18
ADA and NAb sampling	3	7	21
<b>Total (mL)</b>			<b>153.5</b>

## 8. ADVERSE AND SERIOUS ADVERSE EVENTS ASSESSMENT

### 8.1 Definition of Adverse Events, Period of Observation, Recording of Adverse Events

An AE is any untoward medical occurrence in a clinical investigation subject administered a pharmaceutical product and that does not necessarily have a causal relationship with this treatment. An AE can therefore be any unfavorable and unintended sign (including an abnormal laboratory finding), symptom, or disease temporally associated with the use of a medicinal (investigational) product, whether or not related to the medicinal (investigational) product (International Council for Harmonisation [ICH] Guidance E2A 1995).

All AEs are collected from the time the informed consent and/or assent is signed until the end of the defined follow-up period stated in Section 7.1.3. This includes events occurring during the screening phase of the study, regardless of whether or not investigational product is administered. Where possible, a diagnosis rather than a list of symptoms should be recorded. If a diagnosis has not been made, then each symptom should be listed individually. All AEs should be captured in the subject's source document. In addition to untoward AEs, unexpected benefits outside the investigational product indication should also be captured in the subject's source document.

All AEs must be followed to closure (the subject's health has returned to his/her baseline status or all variables have returned to normal), regardless of whether the subject is still participating in the study. Closure indicates that an outcome is reached, stabilization achieved (the investigator does not expect any further improvement or worsening of the event), or the event is otherwise explained. When appropriate, medical tests and examinations are performed so that resolution of event(s) can be documented.

#### 8.1.1 Severity Categorization

The severity of AEs must be recorded during the course of the event including the start and stop dates for each change in severity. An event that changes in severity should be captured as a new event. Worsening of pretreatment events, after initiation of investigational product, must be recorded as new AEs (for example, if a subject experiences mild intermittent dyspepsia before dosing of investigational product, but the dyspepsia becomes severe and more frequent after first dose of investigational product has been administered, a new AE of severe dyspepsia [with the appropriate date of onset] is recorded in the subject's source document).

- Requires inpatient hospitalization or prolongation of existing hospitalization.  
Note: Hospitalizations, which are the result of elective or previously scheduled surgery for pre-existing conditions, which have not worsened after initiation of treatment, should not be classified as SAEs. For example, an admission for a previously scheduled ventral hernia repair would not be classified as an SAE; however, complication(s) resulting from a hospitalization for an elective or previously scheduled surgery that meet(s) serious criteria must be reported as SAE(s).
- Results in persistent or significant disability/incapacity
- Is a congenital abnormality/birth defect
- Is an important medical event. Note: Important medical events that may not result in death, be life-threatening, or require hospitalization may be considered an SAE when, based upon appropriate medical judgment, they may jeopardize the subject and may require medical or surgical intervention to prevent one of the outcomes listed in this definition. Examples of such medical events include allergic bronchospasm requiring intensive treatment in an ED or at home; blood dyscrasias or convulsions that do not result in inpatient hospitalization; or the development of drug dependency or drug abuse.

#### **8.2.4 Serious Adverse Event Collection Time Frame**

All SAEs (regardless of relationship to study) are collected from the time the subject signs the informed consent until the defined follow-up period stated in Section [7.1.3](#) and must be reported to the Shire Global Drug Safety Department and the CRO/Shire medical monitor within 24 hours of the first awareness of the event.

In addition, any SAE(s) considered “related” to the investigational product and discovered by the investigator at any interval after the study has completed must be reported to the Shire Global Drug Safety Department within 24 hours of the first awareness of the event.

#### **8.2.5 Serious Adverse Event Onset and Resolution Dates**

The onset date of the SAE is defined as the date the event meets serious criteria. The resolution date is the date the event no longer meets serious criteria, the date the symptoms resolve, or the date the event is considered chronic. In the case of hospitalizations, the hospital admission and discharge dates are considered the onset and resolution dates, respectively.

In addition, any signs or symptoms experienced by the subject after signing the informed consent form, or leading up to the onset date of the SAE, or following the resolution date of the SAE, must be recorded as an AE, if appropriate.

#### **8.2.6 Fatal Outcome**

Any SAE that results in the subject’s death (ie, the SAE was noted as the primary cause of death) must have fatal checked as an outcome with the date of death recorded as the resolution date. For all other events ongoing at the time of death that did not contribute to the subject’s death, the outcome should be considered not resolved, without a resolution date recorded.

**Table 6 Safety Monitoring Rules for Treatment-emergent Elevated ALT and/or Bilirubin**

Treatment-emergent ALT	Treatment-emergent total bilirubin	Treatment-emergent symptoms	Action
<b>Normal baseline</b>  ALT $\geq 3 \times$ ULN	TBL $\geq 2$ mg/dL increased over baseline <i>or</i> <u>Patients with Gilbert's syndrome or hemolysis:</u> Doubling of baseline direct bilirubin	None	Interrupt investigational product. <sup>b</sup> Initiate close monitoring and workup for competing etiologies.  Investigational product can be restarted only if another etiology is identified and liver enzymes return to baseline.  Testing for hepatitis A, B, and/or C infection may be warranted. Subjects who entered the study with HBcAb with or without HBsAb would need evaluation with HBV DNA to rule out HBV reactivation. <sup>c</sup>
<b>Normal baseline</b>  ALT $\geq 5 \times$ ULN	Normal or elevated	Severe fatigue, nausea, vomiting, right upper quadrant pain <i>or</i> Immunologic symptoms Rash Eosinophilia >5%	Interrupt investigational product. <sup>b</sup> Initiate close monitoring and workup for competing etiologies.  Investigational product can be restarted only if another etiology is identified and liver enzymes return to baseline.  Testing for hepatitis A, B, and/or C infection may be warranted. Subjects who entered the study with HBcAb with or without HBsAb would need evaluation with HBV DNA to rule out HBV reactivation. <sup>c</sup>

ALP=alkaline phosphatase; ALT=alanine aminotransferase; AST=aspartate aminotransferase; HBcAb=hepatitis B core antibody; HBsAb=hepatitis B surface antibody; HBV=hepatitis B virus; TBL=total bilirubin; ULN=upper limit of normal

<sup>a</sup> Elevated baseline ALT defined as ALT  $\geq 1.5 \times$  ULN

<sup>b</sup> Confirmatory repeat liver-related blood tests should be performed within 2 to 3 days before the investigational product is interrupted.

<sup>c</sup> If HBV DNA positive antivirals would need to be started as soon as possible.

Source: Adapted from [Chalasani and Regev, 2016](#).

**Table 6: Power to Detect the Corresponding Treatment Effect for Key Secondary Endpoints**

Key Secondary Endpoint at Week 16	SHP647 Premise	Placebo Premise	Power
Clinical remission by CDAI	26.5%	15%	0.90
Enhanced endoscopic response	25%	13%	0.94
Clinical remission by abdominal pain $\leq 1$ and stool frequency $\leq 3$	24%	14%	0.81
Clinical response by 2-item PRO	52.5%	40%	0.81
Clinical remission by 2-item PRO and endoscopic response	11%	4.5%	0.77
Complete endoscopic healing	6%	2%	0.58

CDAI=Crohn's Disease Activity Index; PRO=patient-reported outcome.

## 9.7 Study Population

The screened set will consist of all subjects who have signed an informed consent document.

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

The safety set will consist of all subjects who have received at least 1 dose of investigational product.

The full-analysis set (FAS) will consist of all subjects in the randomized set who have received at least 1 dose of investigational product.

The per-protocol set will consist of all subjects in the FAS who do not have protocol deviations that may affect the coprimary efficacy endpoints.

The completer set will consist of all subjects in the FAS who have completed the Week 16 assessment for this study.

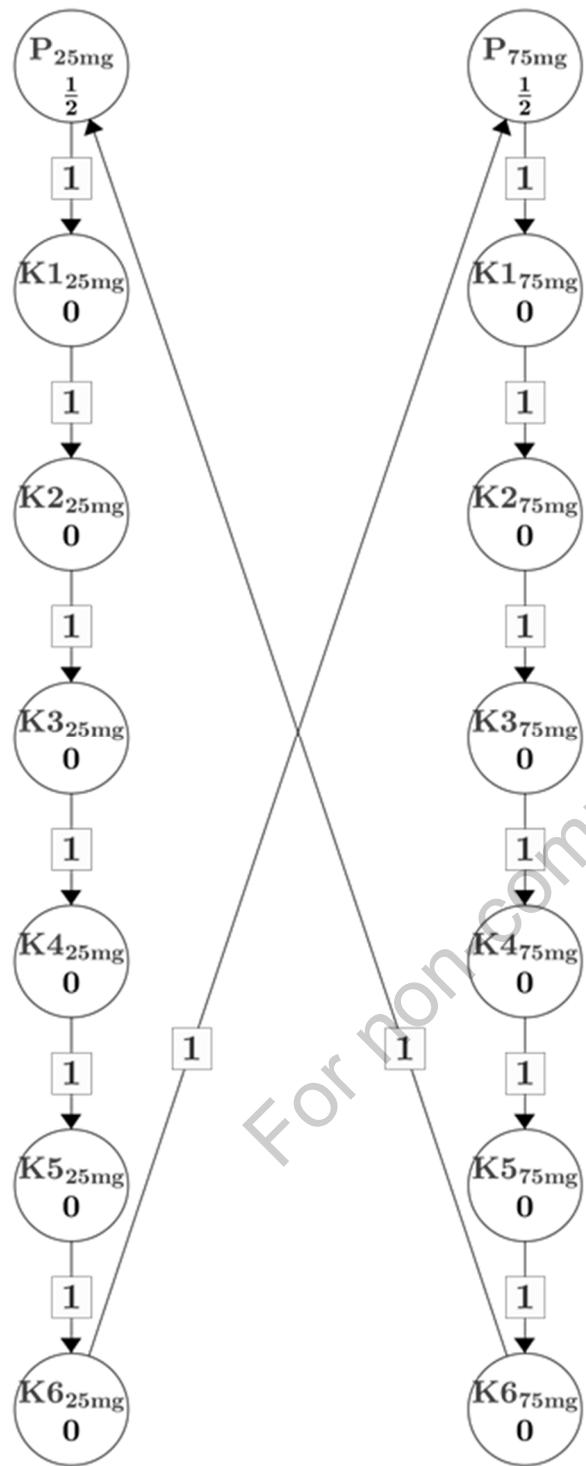
[REDACTED]

[REDACTED]

## 9.8 Efficacy Analyses

Unless otherwise specified, all efficacy analyses will be based on the FAS and subjects will be analyzed according to their randomized treatment, regardless of the treatment they actually received.

Figure 5 Visualization of Multiple Testing Strategy



K=key secondary endpoint; P=coprimary endpoints

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

22 Nov 2019

Term	Percentage
GMOs	~85%
Organic	~80%
Natural	~75%
Artificial	~65%
Organic	~60%
Natural	~55%
Artificial	~50%
Organic	~45%
Natural	~40%
Artificial	~35%
Organic	~30%
Natural	~25%
Artificial	~20%

For non-commercial use

### Crohn's Disease Activity Index (CDAI)

## Crohn's Disease Activity Index (CDAI)

Variable No.	Variable Description	Multiplier	Total
1	No. of liquid or soft stools (each day for 7 days)	X 2	
2	Abdominal pain (0 = none, 1 = mild, 2 = moderate, 3 = severe)	X 5	
3	General well-being (0 = generally well, 1 = slightly under par, 2 = poor, 3 = very poor, 4 = terrible)	X 7	
4	Number of listed complications [arthritis or arthralgia, iritis or uveitis, erythema nodosum or pyoderma gangrenosum or aphthous stomatitis, anal fissure or fistula or abscess, other fistula, fever over 37.8°C (100°F)]	X 20	
5	Use of diphenoxylate or loperamide for diarrhea (0 = no, 1 = yes)	X 30	
6	Abdominal mass (0 = no, 2 = questionable, 5 = definite)	X 10	
7	Hematocrit [Males: 47-Hct (%), Females: 42-Hct (%)]	X 6	
8	Body weight (1-weight/standard weight) X 100 (add or subtract according to sign)	X 1	
<b>CDAI Score</b>			

CDAI=Crohn's Disease Activity Index; ePRO=electronic patient-reported outcome; Hct=hematocrit.

Note: Variable 5: This variable covers taking medication for symptomatic relief from diarrhea, eg, bulking agents, opiates etc.

Variable 7: Absolute deviation of hematocrit is the difference in hematocrit from standard. A male subject with a hematocrit of 40% has an absolute deviation of 7. Each percentage deviation has a value of 6 points. If hematocrit subtotal is <0, enter 0.

Variable 8: This variable is based on Metropolitan Life Tables (these are programmed into the ePRO device).

Percent deviation from standard weight is  $(1 - \text{weight}/\text{standard weight}) \times 100$ ; therefore, positive percent deviation represents weight loss, which adds points to the CDAI. Percentage deviation from standard weight = 1 point for each percent deviation. If body weight subtotal is less than -10, enter -10.

CDAI Interpretation:

- 0-149 points: Asymptomatic remission (Note: subjects requiring steroids to remain asymptomatic are not considered to be in remission but are referred to as being "steroid dependent")
- 150-220 points: Mild to moderate active CD
- 221-450 points: Moderate to severe active CD
- >451 points: Severely active to fulminant disease.

CDAI online estimator: <http://www.ibdjohn.com/cda/>

Sources: [Best et al., 1976](#); [Best et al., 1979](#).

10. How would you rate your general wellbeing over the past 24 hours?

Generally well

Slightly below par

Poor

Very poor

Terrible

For non-commercial use only

[REDACTED]

[REDACTED]

[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]

For non-commercial use only

[REDACTED]

[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]

[REDACTED]

[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]

[REDACTED]

[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]

[REDACTED]

[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]

**Table 5 Clinical Criteria for Diagnosing Anaphylaxis (Type I Hypersensitivity)**

**Anaphylaxis is highly likely when the first criterion below and at least one of the following criteria a and b are fulfilled:**

Acute onset of an illness (minutes to several hours) with involvement of the skin, mucosal tissue, or both (eg, generalized hives, pruritus or flushing, swollen lips-tongue-uvula)

*AND AT LEAST ONE OF THE FOLLOWING:*

- a) Respiratory compromise (eg, dyspnea, wheeze-bronchospasm, stridor, reduced PEF, hypoxemia)
- b) Reduced BP<sup>a</sup> or associated symptoms of end-organ dysfunction (eg, hypotonia [collapse], syncope, incontinence).

BP=blood pressure; PEF=peak expiratory flow

<sup>a</sup> Low systolic BP for children is defined as less than 90 mmHg from 11 to 17 years.

Source: Adapted from [Sampson et al., 2006](#).

Type III hypersensitivity responses, including those mediated by immune complexes and T cells (delayed hypersensitivity responses in the older literature), are relatively rare with respect to therapeutic protein products and a high degree of clinical suspicion is necessary for the diagnosis (Center for Drug Evaluation and Research - Guidance for industry: Immunogenicity assessment for therapeutic protein products, 2014). Type III hypersensitivity reactions involve the formation of biologic/ADA immune complexes in the circulation which, when present in the correct stoichiometric ratio, become deposited in tissues. Once immune complexes are deposited, they can elicit complement activation and inflammation, leading to tissue damage. When immune complexes are deposited in tissues, they tend to localize in small postcapillary venules where there is loss of laminar blood flow, in sites of ultrafiltration where there is high pressure and fenestrated endothelium (eg, choroid plexus, ciliary body, synovium, and glomeruli), in sites of turbulent blood flow (eg, coronary artery branches off aorta, aortic bifurcations, and cardiac valve leaflets), and in renal glomerular endothelium.

Signs and symptoms of immune complex deposition typically have onset 1 to 3 weeks after exposure ([Warrington et al., 2018](#)) usually improving in 7 to 10 days, with full recovery in 2 to 4 weeks and may include fever, rash (including hives), arthralgia, myalgia, vasculitis, Arthus reaction, general ill feeling, itching, and swollen lymph nodes. Some of these findings, such as fever, rash, arthralgia, and myalgia, are consistent with findings associated with IBD and may therefore be very difficult to assign to a particular etiology. When such a reaction is suspected, samples for laboratory assessment will be obtained and stored. Tests will be performed if the diagnosis is confirmed and requested by the adjudication committee.

The administration and/or use of the unassigned treatment is/are always reportable as a medication error.

The administration and/or use of an expired investigational product should be considered as a reportable medication error.

There is no specific antidote for overdose with ontamalimab. Treatment should be symptomatic and supportive.

### **8.1.9      Unexpected Adverse Event**

An unexpected AE is an AE whose nature, severity, specificity, or outcome is not consistent with the term, representation, or description used in the reference safety information (RSI).

“Unexpected” also refers to the AEs that are mentioned in the IB as occurring with a class of drugs or as anticipated from the pharmacological properties of the product but are not specifically mentioned as occurring with the particular product under investigation.

The expectedness of AEs will be determined by the sponsor using the IB as the RSI. This determination will include considerations such as the number of AEs previously observed, but not on the basis of what might be anticipated from the pharmacological properties of a product.

### **8.1.10     Suspected Unexpected Serious Adverse Reaction**

A suspected unexpected serious adverse reaction (SUSAR) is defined as any suspected adverse reaction to study treatment (ie, including active comparators) that is both serious and unexpected.

The event(s) must meet all of the following:

- Suspected adverse reaction
- Serious
- Unexpected
- Assessed as related to study treatment.

## **8.2        Serious Adverse Event Procedures**

### **8.2.1      Reference Safety Information**

The reference for safety information for this study is Section 6.8 of the ontamalimab IB, which the sponsor has provided under separate cover to all investigators.

### **8.2.2      Reporting Procedures**

All initial and follow-up SAE reports must be reported by the investigator to the Shire Global Drug Safety Department and the CRO/Shire medical monitor within 24 hours of the first awareness of the event. Note: The 24-hour reporting requirement for SAEs does not apply to reports of abuse, misuse, overdose, or medication errors (see Section 8.1.8) unless they result in an SAE.

The overall power for the coprimary endpoints will be approximately 87% assuming no correlation between the tests on the endpoints and approximately 90% assuming a correlation of 0.4.

With the sample size of 1032 subjects, **Table 7** provides the power for detecting a treatment difference between a ontamalimab treatment group and the placebo group for the key secondary endpoints.

**Table 7 Power to Detect the Corresponding Treatment Effect for Key Secondary Endpoints**

Key Secondary Endpoint at Week 16	Ontamalimab Premise	Placebo Premise	Power
Clinical remission by CDAI	26.5%	15%	0.90
Enhanced endoscopic response	25%	13%	0.94
Clinical remission by abdominal pain $\leq 1$ and stool frequency $\leq 3$	24%	14%	0.81
Clinical response by 2-item PRO	52.5%	40%	0.81
Clinical remission by 2-item PRO and endoscopic response	11%	4.5%	0.77
Complete endoscopic healing	6%	2%	0.58

CDAI=Crohn's Disease Activity Index; PRO=patient-reported outcome

## 9.7 Study Population

The screened set will consist of all subjects who have signed an informed consent document.

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

The safety set will consist of all subjects who have received at least 1 dose of investigational product.

The full analysis set (FAS) will consist of all subjects in the randomized set who have received at least 1 dose of investigational product.

The per-protocol (PP) set will consist of all subjects in the FAS who do not have predefined protocol deviations that may affect the coprimary efficacy endpoints.

The completer set will consist of all subjects in the FAS who have completed the Week 16 assessment for this study.

