1.0 Title Page

Statistical Analysis Plan for Extension Phase

Study M13-740

A Multicenter, Randomized, Double-Blind, Placebo-Controlled Study of ABT-494 for the Induction of Symptomatic and Endoscopic Remission in Subjects with Moderately to Severely Active Crohn's Disease who have Inadequately Responded to or are Intolerant to Immunomodulators or Anti-TNF Therapy

Date: 14 Aug 2017

Version 1.0

2.0 Table of Contents

List of Tables

List of Figures

3.0 Introduction

This statistical analysis plan (SAP) describes the statistical analysis to be completed by the AbbVie Clinical Statistics Department for study Protocol M13-740 Amendment 5 dated 07 November 2016. It provides details to guide the analyses for baseline, efficacy, and safety variables. This document describes the populations that will be analyzed, the variables that will be analyzed, and the statistical methods that will be utilized.

This is the first version of the SAP for Protocol M13-740 Extension Phase.

This document describes the analyses of data except pharmacokinetic, microbiota metagenomic, pharmacogenetic and serologic data which will be analyzed separately. It takes into account ICH Guidelines E3 and E9.

The statistical analysis for Study M13-740 Induction was performed separately. The SAP for Induction has been signed off prior to Induction database lock in December, 2016.

Unless noted otherwise, all analyses will be performed using SAS version 9.2 or later (SAS Institute Inc., Cary, NC 27513) under the UNIX operating system.

4.0 Study Objectives, Design and Procedures

4.1 Objectives

The objective of this study is to determine the efficacy and safety of multiple doses of ABT-494 versus placebo and to assess the pharmacokinetics of ABT-494 following oral administration in subjects with moderately to severely active Crohn's disease with a history of inadequate response to or intolerance to immunomodulators or anti-TNF therapy.

4.2 Design Diagram

This is a Phase 2, multicenter, randomized, double-blind, placebo-controlled study designed to evaluate the efficacy, safety, and pharmacokinetics of ABT-494 as induction

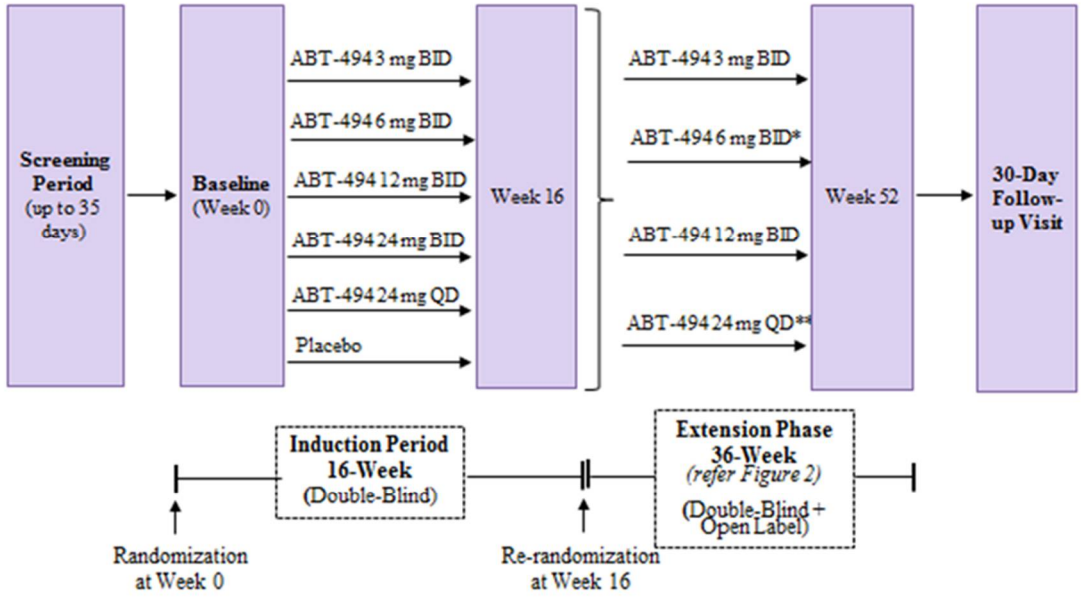
therapy in subjects with moderately to severely active Crohn's disease and evidence of mucosal inflammation defined by a SES-CD ≥ 6 (or ≥ 4 for subjects with disease limited to the ileum) and average daily soft/liquid stool frequency ≥ 2.5 or average daily abdominal pain score of ≥ 2.0; and CDAI ≥ 220 and ≤ 450, with a history of inadequate response or intolerance to immunomodulators and/or anti-TNF therapy. The study will allow enrollment of up to 30% of subjects who were primary non responders to anti-TNF treatment.

The study is designed to enroll approximately 210 subjects to meet scientific and regulatory objectives without enrolling an undue number of subjects in alignment with ethical considerations. Therefore, if the target numbers of subjects have been enrolled, there is a possibility that additional subjects in screening will not be enrolled.

The study duration could be up to 60 weeks, including a Screening Period of up to 35 days, a 16-week double-blind induction period, a 36-week double-blind extension phase, and a 30-day follow-up period. Subjects who met eligibility criteria were randomized in a 1:1:1:1:1:1 ratio to one of the double-blinded induction treatment arms.

The schematics of the overall study design are shown in Figure 1 and Figure 2.

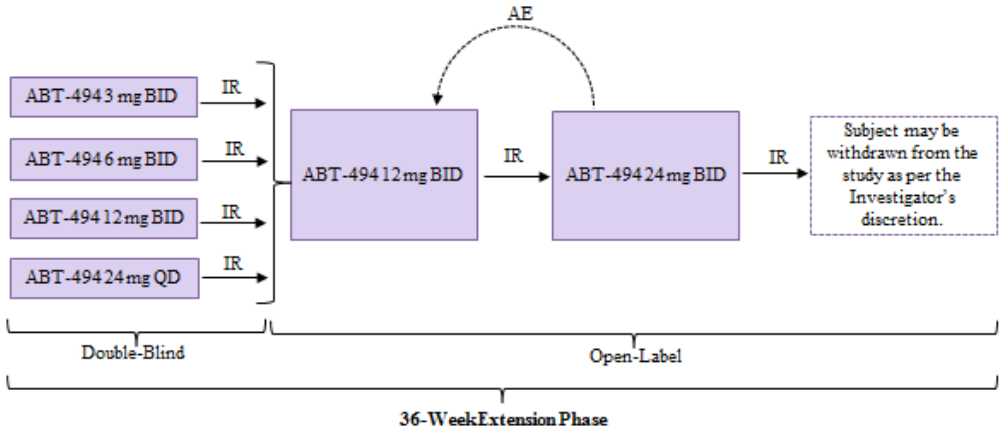
Figure 1. Study Design Schematic



*The 6 mg BID dose arm was added to replace the 24 mg QD dose arm in Protocol Amendment 2

**Subjects who were re-randomized at Week 16 to 24 mg QD prior to Protocol Amendment 2 will continue to receive the double-blind 24 mg QD dose until Week 52/PD.

Figure 2. Study Design Schematic for Open-Label Extension Option for Patients Who Do Not Adequately Respond During the Extension Phase



IR: Inadequate Response

Screening Period

Within 35 days prior to the Baseline Visit, subjects received a full explanation of the study design and study procedures, provided a written informed consent, and underwent the screening procedures as outlined in Table 1.

An e-diary was dispensed and training were provided to the subject at the Screening Visit.

Laboratory tests performed during the Screening period can be repeated in case the abnormalities are considered to be transient by the investigator.

16-Week Induction Period

This period began at the Baseline Visit (Week 0) and end at the Week 16 Visit. At the Baseline Visit, subjects who met all the inclusion criteria and none of the exclusion criteria described in the protocol (Section 5.2.1 and Section 5.2.2) were enrolled into the study and randomized to double-blind induction period. The randomization at Baseline

were stratified by endoscopic disease severity (SES-CD < 15 and ≥ 15), prior anti-TNF use (naïve and experienced), and by participation in the substudy of the gene expression in intestinal biopsies (Yes and No). As part of the randomization at Baseline, subjects were randomly assigned (1:1) to have their follow-up ileocolonoscopy done at Week 12 or Week 16. During this period of the study, subjects visited the study site at Weeks 2, 4, 8, 12 and 16. A ± 3-day window is permitted around scheduled study visits. The last dose of study drug during this period is taken the evening prior to the Week 16 visit.

However, the stratification factor Prior anti-TNF use (naïve and experienced) was added under Amendment 3, and there are less than 10 anti-TNF naïve subjects enrolled under Amendment 3. Therefore, the Prior anti-TNF use will not be included in all analysis as a stratification factor.

Subjects may discontinue study drug treatment at any time during study participation. Subjects that end study participation early will have an Early Termination (ET) Visit and complete the procedures outlined for the ET Visit in Table 1 as soon as possible after the last dose of study drug and preferably prior to the administration of any new therapies.

36-Week Extension Phase

At Week 16, subjects who have completed the induction period will be re-randomized in a ratio of 1:1:1 to one of the three double-blinded doses of ABT-494 3 mg BID, 6 mg BID or 12 mg BID. The re-randomization will be stratified by dose received during the first 16 weeks, and overall response (responder versus non-responder) at Week 16. The subjects who were re-randomized at Week 16 to ABT-494 24 mg QD prior to Protocol Amendment 2 will continue to receive the same double-blind dose until Week 52/PD.

The central reader endoscopic score will be used for calculating the Endoscopic response for the evaluation of the efficacy endpoints. However, for stratification at the time of re-randomization, the endoscopic score at BL from central reader and the endoscopic score at Week 12 or Week 16 from site local reader will be used in order to determine response status.

Clinical response is defined as average daily liquid/soft stool frequency reduction of at least 30% from Baseline and average daily abdominal pain score not worse than Baseline OR average daily abdominal pain score reduction at least 30% from baseline and average daily stool frequency not worse than Baseline.

During this period of the study, subjects will visit the study site at Weeks 20, 28, 36, 44, and 52/Early Termination. A ± 3-day window is permitted around scheduled study visits. The last dose of study drug is taken the evening prior to the Week 52 visit.

Subjects will be expected to remain on double-blinded therapy throughout the entire 36-week extension phase. However, subjects who are considered by the investigator to have not achieved meaningful symptomatic relief and meet the criteria for inadequate response at or after Week 20 will be eligible to receive the open-label therapy with ABT-494 12 mg BID, Figure 2. Subjects who were re-randomized at Week 16 prior to Protocol Amendment 2 and are taking open-label ABT-494 24 mg QD will be transitioned to ABT-494 12 mg BID.

Subjects who still continue to meet criteria for inadequate response following a 4-week course of open-label ABT-494 12 mg BID will be eligible to dose escalate to open-label ABT-494 24 mg BID.

Subjects with persistent inadequate response while on ABT-494 12 mg BID or 24 mg BID may be withdrawn from the study at the investigator's discretion.

Criteria for Inadequate response are as follows:

- Average daily liquid/soft Stool Frequency > 2.2 OR average daily Abdominal Pain score > 1.8 AND
- An increase level of hs-CRP of at least 1 mg/L from Baseline or a hs-CRP ≥ 5 mg/L, either at the previous visit or at the current visit.

Note: hs-CRP will be performed by Central Lab. However, for the purpose of evaluating Inadequate Response criteria, hs-CRP can also be done at local lab.

Assessment of inadequate response should include consideration by the Investigator to rule out symptoms caused by reasons other than Crohn's disease related inflammation.

In the open-label period, if the subject treated with 24 mg BID has an occurrence of an adverse event thought to be possibly related to study medication that in the opinion of the investigator warrants dose reduction, then the dose can be de-escalated to ABT-494 12 mg BID at the investigator's discretion.

Note: Dose escalation and de-escalation are permitted only once during the study.

Subjects may discontinue study drug treatment at any time during study participation. Subjects that end study participation early will have an ET Visit and complete the procedures outlined for the ET Visit in Table 1 as soon as possible after the last dose of study drug and preferably prior to the administration of any new therapies.

Subjects who are not re-randomized at Week 16 or subjects who early terminated from the study will have a follow-up visit approximately 30 days after the last administration of study drug to obtain information on any new or ongoing adverse events (AEs), and to collect vital signs and clinical laboratory tests.

Follow-Up Period

Subjects who have completed the Week 52 visit will have a follow-up visit approximately 30 days after the last administration of study drug to obtain information on any new or ongoing adverse events (AEs), and to collect vital signs and clinical laboratory tests.

Re-Screen

Subjects that initially screen fail for the study may be permitted to re-screen following re-consent. The subject must meet all the inclusion and none of the exclusion criteria at the time of re-screening in order to qualify for the study. There is no minimum period of time a subject must wait to re-screen for the study. If the subject had a complete initial screening evaluation including the assessment of a purified protein derivative (PPD) test (or equivalent), or Interferon-Gamma Release Assay (IGRA; QuantiFERON-TB Gold test

or T-SPOT TB test), chest x-ray, HBV, HCV and ECG, these tests will not be required to be repeated for re-screening provided the conditions noted in the protocol (Section 5.3.1.1) are met and no more than 90 days have passed.

An endoscopy with biopsy will not be required to be repeated for re-screening provided the conditions noted in the protocol (Section 5.3.1.1) are met and no more than 30 days have passed (from the previous screening endoscopy). All other screening procedures will be repeated. As appropriate, sites are encouraged to contact the AbbVie Medical Monitor to confirm if subjects should or should not be re-screened.

Unscheduled Visits

Unscheduled Visits are for purposes when the subject is coming in for a medical visit for evaluation and assessment. During Unscheduled Visits, blood and urine samples will be obtained for the laboratory tests listed in the protocol (Table 2).

Visits for dispensing new study drug in case of temperature excursion, loss or damage are not considered an Unscheduled Visit. In addition, visits to only retest a lab will not be considered an Unscheduled Visit.

Table 1. Study Activities

Activity	Screening[a]	Baseline (Week 0)	Wk 2	Wk 4	Wk 8	Wk 12	Wk 16	Wk 20	Wk 28	Wk 36	Wk 44	Wk 52	Early Termination (ET) Visit	Unscheduled Visit	30-Day Follow-Up Visit
Informed Consent	X														
Inclusion/Exclusion[b]	X	X													
Medical/Surgical History[b]	X	X													
Prior and Concomitant Medications[b]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Physical Exam[c]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Endoscopy[d]	X					X	X					X	X		
Biopsy[e]	X					X	X					X	X		
SES-CD		X				X	X					X	X		
Vital Signs[f]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
12-Lead ECG[g]	X													X	
Chest X-Ray[h]	X													X	
PPD Skin Test or QuantiFERON TB Gold Test[i]	X														
Latent TB Risk Factor Questionnaire	X														
Blood Chemistry and Hematology	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

Table 1. Study Activities (Continued)

Activity	Screening[a]	Baseline (Week 0)	Wk 2	Wk 4	Wk 8	Wk 12	Wk 16	Wk 20	Wk 28	Wk 36	Wk 44	Wk 52	Early Termination (ET) Visit	Unscheduled Visit	30-Day Follow-Up Visit
Urinalysis[j]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Pregnancy Test[k]	X	X	X	X	X	X	X	X	X	X	X	X	X		X
hsCRP	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
HBV and HCV Screening[l]	X														
ABT-494 Concentration[m]			X	X	X	X	X	X	X	X	X	X	X		
Pharmacodynamic biomarkers		X			X		X					X	X		
C. difficile toxin	X														
Stool Sample (fecal calprotectin)[n,o]	X			X			X		X			X	X		
Bristol Stool Chart[p]	X			X			X		X			X	X		
Corticosteroid Taper[q]			X												
Crohn's Disease Activity Index (CDAI)		X	X	X	X	X	X	X	X	X	X	X	X	X	
Inflammatory Bowel Disease Questionnaire (IBDQ)		X			X		X					X	X		

Table 1. Study Activities (Continued)

Activity	Screening[a]	Baseline (Week 0)	Wk 2	Wk 4	Wk 8	Wk 12	Wk 16	Wk 20	Wk 28	Wk 36	Wk 44	Wk 52	Early Termination (ET) Visit	Unscheduled Visit	30-Day Follow-Up Visit
European Quality of Life 5 Dimensions (EQ-5D)		X			X		X					X	X		
Work Productivity and Impairment Questionnaire (WPAI)		X			X		X					X	X		
Abdominal Pain Rating Scale (0 – 10 Scale)		X				X	X								
Monitor Adverse Events[r]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Study Drug Dispensing/ Administration		X	X	X	X	X	X	X	X	X	X				
Dispense Subject Diary	X														
Subject Diary Review		X	X	X	X	X	X	X	X	X	X	X	X	X	

Table 1. Study Activities (Continued)

Activity	Screening[a]	Baseline (Week 0)	Wk 2	Wk 4	Wk 8	Wk 12	Wk 16	Wk 20	Wk 28	Wk 36	Wk 44	Wk 52	Early Termination (ET) Visit	Unscheduled Visit	30-Day Follow-Up Visit
Additional Samples Collection (Optional)															
Pharmacogenetic[s]		X													
Serologic markers[t]		X					X								
mRNA[t]		X					X								
Biopsy for gene expression[u]	X					X	X								

Wk = Week

a. The Screening period will be a minimum of 7 days but no more than 35 days. Baseline Visit date will serve as the reference for all subsequent visits. A ± 3-day window is permitted around scheduled study visits.

b. Update inclusion/exclusion, prior and concomitant therapy, and medical/surgical history information to assure subject eligibility.

c. Physical examination performed at Screening, Baseline, Week 16 and Week 52/Early Termination Visits are full physical examinations which must include an assessment of extra-intestinal manifestations (EIMs) and a count of the number of cutaneous fistulas. Physical examinations and those performed at all other visits are symptom based but must include a count of the number of cutaneous fistulas.

d. Colonoscopy at/during the screening period or within 15 days of the Screening Visit will be used to calculate the SES-CD score at Baseline. Subject will be randomly assigned at Baseline to have endoscopy performed at Week 12 or Week 16.

e. Biopsy may be done when performing the endoscopy. Biopsy to confirm CD diagnosis may be done when performing the endoscopy if appropriate documentation for confirmation of the diagnosis does not exist. Biopsies to rule out dysplasia and colon cancer may be taken at the investigator's discretion.

f. Blood pressure, pulse rate, temperature, respiratory rate and weight should be performed before blood draws are performed. Height will be measured at Screening only (with shoes off and then adding 1 inch or 2.5 cm).

g. Subjects with normal ECG within 90 days of Screening would not require a repeat ECG, if documentation is available. Subjects can have a repeat ECG at any time during the study as warranted based on the opinion of the Investigator.

Table 1. Study Activities (Continued)

h. Chest x-ray includes posterior-anterior (PA) and lateral views. Obtain chest x-ray for subjects with TB risk factors as identified by the TB risk factor questionnaire or for subjects living in areas endemic for TB or for subjects with positive PPD or QuantiFERON-TB Gold. Subjects can have a chest x-ray anytime during the study as warranted based on the opinion of the Investigator.

i. PPD skin test is to be read 48 to 72 hours after placement.

j. Dipstick urinalysis will be completed by the sites at all required visits. A microscopic analysis will be performed by the central laboratory, in the event the dipstick results show protein, ketones or blood greater than negative or glucose greater than normal.

k. Serum pregnancy test will be performed on all women of childbearing potential at Screening and at Week 52/ET. Urine pregnancy test will be performed locally at every visit for all women of childbearing potential. If any urine pregnancy test is positive, a serum pregnancy test will be performed by the central laboratory.

If required by country regulatory authorities, monthly pregnancy tests will be performed throughout the study.

l. Subjects will be tested for the presence of the hepatitis B Virus (HBV) and hepatitis C Virus (HCV) at Screening. A positive result for the hepatitis B surface antigen (HBs Ag) or hepatitis C (HCV RNA detectable in any subject with anti-HCV Ab) will be exclusionary. For subjects who are negative for HBs Ag but are positive for core antibodies (HBc Ab), HBV DNA PCR will be performed and any result that meets or exceeds detection sensitivity will be exclusionary.

m. For all subjects, a PK sample will be collected at Week 2, within 1 to 3 hours after dosing, if possible. At Week 4, Week 8, Week 12, Week 16, Week 20, Week 28, Week 36, Week 44 and Week 52/ET, a blood sample for PK will be collected at any time during the visit, and preferably within 1 to 8 hours after the last dose. Patients can take the dose on visit days at their regular schedule and not necessarily at the clinic. The date and accurate time of the PK sample collection and the last ABT-494 dose will be recorded in the CRF to the nearest minute.

n. Stool sample will be collected at each time point indicated. For the visit that endoscopy will be conducted, stool sample should be collected prior to endoscopy.

o. A stool sample will be collected for fecal calprotectin analysis at each time point indicated. Subjects will be asked to provide a stool sample at the visit, if possible, or subjects will be sent home with instructions and stool sample supplies (supplies will be provided).

p. Stool samples for microbiota analysis should be assessed using the Bristol Stool Chart.

q. Subject is mandated to begin corticosteroid taper at Week 2.

r. Collection of SAEs begins the day the subject signs the informed consent.

s. Only if subject provides written consent to collect the pharmacogenetic sample; if the informed consent form is not signed, no pharmacogenetic sample can be collected. The sample is preferred to be collected at BL.

Table 1. Study Activities (Continued)

t. Only if subject provides written consent to collect the serum biomarkers and mRNA samples; if the informed consent form is not signed, no serum biomarkers and mRNA can be collected.

u. Only if subject provides written consent to collect biopsy samples for gene expression analysis; if the informed consent form is not signed, no biopsy for gene expression can be collected. Biopsies will be obtained at baseline and as the Week 12/16 endoscopy, based on the colonoscopy assignment for Week 12 or 16.

4.3 Sample Size

Approximately 210 subjects were equally allocated to five treatment groups and the placebo group, representing a randomization ratio of 1:1:1:1:1:1. The sample size for this study was based on the expected proportion of subjects who achieve endoscopic remission at Week 12/16 and on the expected proportion of subjects who achieve clinical remission at Week 16. Assuming an endoscopic remission (as well as clinical remission) rate of 12% in the placebo arm and maximum of 35% in at least one of the ABT-494 BID treatment arms (3 mg BID, 6 mg BID, 12 mg BID and 24 mg BID) at Week 12/16, a sample size of 35 subjects per treatment group is sufficient to test for the presence of a dose response signal, to select the best dose response model for the observed data out of a prespecified set of candidate models, and to estimate target doses of interest (e.g., the minimum effective dose, MED) via modeling using MCP-Mod (Multiple comparison procedure and modeling) approach. This approach provides at least 80% average power to detect a dose effect at 5% level of significance (one-sided) with the linear, E_{max}, exponential, logistic, and sigEmax models pre-specified as likely candidates to characterize the dose-response for ABT-494 for the two co-primary endpoints of clinical and endoscopic remission.

5.0 Analysis Populations

5.1 Definition for Analysis Populations

The intent-to-treat responders (ITT-R) and intent-to-treat non-responders (ITT-NR) analysis sets includes all re-randomized subjects in the double-blind extension phase and are responders and non-responders at Week 16, respectively.

For the ITT-R and ITT-NR analysis sets, subjects are assigned to a treatment group based on the re-randomization schedule, regardless of the treatment actually received. The ITT-R and ITT-NR analysis sets will be used as primary sets for all efficacy analysis. Subjects who were randomized to placebo during the Induction Period will be summarized separately.

The intent-to-treat clinical responders (ITT-CR) and intent-to-treat clinical non-responders (ITT-CNR) analysis sets includes all re-randomized subjects in the double-blind extension phase and are clinical responders and clinical non-responders at Week 16, respectively.

For the ITT-CR and ITT-CNR analysis sets, subjects are assigned to a treatment group based on the re-randomization schedule, regardless of the treatment actually received. The ITT-CR and ITT-CNR analysis sets will be used for sensitivity analysis. Subjects who were randomized to placebo during the Induction Period will be summarized separately.

The safety analysis set consists of all subjects who received at least one dose of study medication in the Extension Phase. For the safety analysis sets, subjects are assigned to a treatment group based on the treatment actually received, regardless the treatment randomized. The safety analysis set will be used for all safety analysis of Extension phase.

Any ABT-494 analysis set will consist of all subjects who received at least one dose of ABT-494 during Induction or Extension. Any ABT-494 analysis set will be used for cumulative safety analysis (from Induction baseline to study end).

5.2 Variables Used for Stratification of Randomization

At Baseline, subjects were randomized in a 1:1:1:1:1:1 ratio to receive one of the six treatment groups (double-blind ABT-494 induction doses or matching placebo).

The following are the treatment groups:

Group 1: ABT-494 3 mg BID

Group 2: ABT-494 6 mg BID

Group 3: ABT-494 12 mg BID

Group 4: ABT-494 24 mg BID

Group 5: ABT-494 24 mg QD

Group 6: Placebo

The randomization at Baseline was stratified by endoscopic disease severity (SES-CD < 15 and ≥ 15), prior anti-TNF use (naïve and experienced), and by participation in the substudy of the gene expression in intestinal biopsies (Yes and No). As part of the randomization at Baseline, subjects were randomly assigned (1:1) to have their follow-up ileocolonoscopy done at Week 12 or Week 16.

At Week 16, subjects who have completed the induction period were re-randomized in a ratio of 1:1:1 to one of the three double-blinded doses of ABT-494 3 mg BID, 6 mg BID or 12 mg BID. The re-randomization was stratified by dose received during the first 16 weeks, and overall response (responder versus non-responder) at Week 16. The subjects who were re-randomized at Week 16 to ABT-494 24 mg QD prior to Protocol Amendment 2 continued to receive the same double-blind dose until Week 52/PD.

Overall response consists of both Endoscopic response and Clinical response. The central reader endoscopic score will be used for calculating the Endoscopic response for the evaluation of the efficacy endpoints. However, for stratification re-randomization, the endoscopic score at baseline from central reader and the endoscopic score at Week 12 or Week 16 from site local reader will be used in order to determine response status.

6.0 Analysis Conventions

Definition of Baseline for Induction Period

The Baseline visit date for Induction Period is the date when the first dose of study drug is received and referred to as Day 1 or Week 0. The Baseline value for a variable is defined as the last non-missing value on or before the date of the first dose of study drug for Induction Period.

Definition of Baseline for Extension Phase

The Baseline visit date for Extension Phase is the date when the first dose of study drug for Extension Phase is received and referred to as Day 1 or Week 0 of Extension Phase. The Baseline value for a variable is defined as the last non-missing value on or before the date of the first dose of study drug for Extension Phase.

Definition of Rx Days (Days Relative to the First Dose of Study Drug)

Rx Days are calculated for each time point of interest and it provides a quantitative measure of days between the event and the first dose date. That is, the Rx Day is calculated as the event date minus the date of first dose of study drug plus 1.

The Rx Day will be a negative value when the time point of interest is prior to the date of first dose of study drug, and the Rx Day will be a positive value when the time point of interest is after the first dose date. By this calculation algorithm the first dose day is Rx Day 1, while the day prior to the date of first dose is defined as Rx Day –1 (there is no Rx Day 0). Rx Days are used to map actual study visits to the protocol specified study visits.

Dealing with Multiple Measurements Collected on the Same Day

For efficacy related analyses other than imaging endpoints, if multiple measurements for a particular parameter are collected on the same day for the same subject, the average of those measurements will be used.

For safety related analyses, if multiple measurements are made for a particular laboratory or vital sign parameter on the same day for the same subjects, the average of the values will be used in the analyses of change from Baseline. For summaries and listings for shift from baseline and potentially significant values, all collected values within the pre-specified treatment window will be used.

Definition of Analysis Windows

Since subjects do not always adhere to the study visit schedule, the following rules will be applied to assign actual visits to protocol-specified visits including early termination visits. For each study visit mentioned in the protocol, a nominal or target day will be selected to represent the corresponding visit along with a window around the target day. Windows will be selected in a non-overlapping fashion so that a date collected on the CRF does not correspond to multiple visit windows. Moreover, windows will not discard any Post-Baseline measurement recorded on the CRF. If a subject had two or more actual visits in one visit window, the visit closest to the schedule visit will be used as the study visit for that window. If two visits are equidistant from the target, then the later visit will be used for reporting.

Table 2. Visit Windows for Analysis of CDAI and CRP, Laboratory Parameters and Vital Signs for 36-Week Double-Blind Extension Phase

Scheduled Week	Nominal Day	Time Window (Rx Day Range)
Week 0 (Baseline of Extension Phase)	1[a]	≤ 1
Week 4 (Week 20 of the study)	29	2 – 57
Week 12 (Week 28 of the study)	85	58 – 113
Week 20 (Week 36 of the study)	141	114 – 169
Week 28 (Week 44 of the study)	197	170 – 225
Week 36 (Week 52 of the study)	253	226 – 999

Rx Day B = date of visit – date of first study drug in double-blind extension phase + 1

a. Day of first dose of study drug in double-blind extension phase.

Table 3. Visit Windows for Analysis of SES-CD, EIM, IBDQ, EQ5D, and WPAI for 36-Week Double-Blind Extension Phase

Scheduled Week	Nominal Day	Time Window (Rx Day Range)
Week 0 (baseline of Extension Phase)	1[a]	≤ 1
Week 36 (Week 52 of the study)	253	2 – 999

Rx Day B = date of visit – date of first study drug in double-blind extension phase + 1

a. Day of first dose of study drug in double-blind extension phase.

Table 4. Visit Windows for Analysis of Bristol Stool Chart and Fecal Calprotectin for 36-Week Double-Blind Extension Phase

Scheduled Week	Nominal Day	Time Window (Rx Day Range)
Week 0 (baseline of Extension Phase)	1[a]	≤ 1
Week 12 (Week 28 of the study)	85	2 – 169
Week 36 (Week 52 of the study)	253	170 – 999

Rx Day B = date of visit – date of first study drug in double-blind extension phase + 1

a. Day of first dose of study drug in double-blind extension phase.

CD-Related Corticosteroid Use

Subjects not on CD-related corticosteroids (systemic or rectal corticosteroids) at Induction Baseline which are then initiated during the study or who have dosages of these medications increased to greater than the dose taken at Baseline will be censored for efficacy assessments (i.e., will be considered non-responders for categorical endpoints and will have Baseline values carried forward for non-categorical assessments) from that point through the end of the study. These subjects will continue to be evaluated in the safety population.

The equivalent steroid dose will be determined based on the table below:

Corticosteroid	Equivalent Dose (mg)
Cortisone	25
Hydrocortisone	20
Prednisone	5
Prednisolone	5
Triamcinolone	4.0
Methylprednisolone	4.0
Betamethasone	0.75
Dexamethasone	0.75
Budesonide	1
Beclomethasone	5
Beclometasone	5

Open-Label Therapy

Subjects who are considered by the investigator to have not achieved meaningful symptomatic relief and have met the criteria for inadequate response at or after Week 20 of the study (Week 4 of the Extension Phase) will be eligible to receive open-label therapy with ABT-494 12 mg BID. These subjects will be censored for efficacy assessments (i.e., will be considered non-responders for categorical endpoints and will have the non-missing value prior to open-label therapy carried forward for non-categorical assessments) from that point through the end of the study.

Definition of Missing Data Imputation

The following imputation methods will be used to impute missing values in the efficacy analyses. In addition, an observed case analysis will be performed.

Non-Responder Imputation (NRI)

The NRI approach is used for binary efficacy variables. These variables can take values of 'Achieved' or 'Not Achieved' or may be missing for any reason including discontinuation from study. According to the NRI imputation approach, all missing values will be considered as 'Not Achieved.'

Last Observation Carried Forward (LOCF)

For all variables (categorical variables and continuous variables), the following rules will be used for the LOCF approach:

1. Extension Baseline and Pre-Baseline values will not be used to impute the missing Post-Baseline values.
2. Missing values after Study Day 1 will be imputed using the latest non-missing values after Day 1 and prior to the missing value. If there are no non-missing values after Baseline, then the LOCF value will be missing.

Observed Case (OC)

Observed case analysis will be performed such that missing values will not be imputed. However, all observed values will be replaced by Induction Baseline value once subjects received CD-related corticosteroid therapy, or value prior to open label therapy once subjects moved to open-label therapy.

Mixed-Effect Model Repeated Measure (MMRM)

The MMRM model will be used for continuous efficacy variables with longitudinal data as the primary analysis. The MMRM model includes the baseline values as covariate, treatment, time point and treatment-by-time point interaction as fixed effects; and subjects within treatment as random effect. An unstructured (co)variance structure will be used to model the within-subject error. The comparison at a time point will be the contrast between treatments at that time point. Satterthwaite's approximation will be used to estimate denominator degrees of freedom.

Imputation of Missing Dates

For Baseline, efficacy, and safety parameters, if the day and/or month are missing, the following conventions will be used to impute the missing dates:

- 01 for missing start day

- End of month for missing end day
- January 1st for missing start month
- December 31st for missing end month

In case of partially missing AE start and stop dates, the dates will be imputed by comparing to first dose date of study medication so than the corresponding AEs will be made treatment-emergent whenever possible. If the start date of an AE is partially missing and the month is the same as the start date of a new therapy, the AE will be made treatment emergent to the new therapy.

In case of missing or partially missing study drug dosing dates, the dates will not be imputed. Subjects will be treated as not receiving dose on that date.

Rule for CDAI Calculation

Up to 14 days of diary entries will be evaluated from the ePRO tool for the CDAI calculation for each visit. The diary entries on the days the subjects receive endoscopy preparation medications, the day of endoscopy procedure, and 2 days after endoscopy procedure will be excluded. For each CDAI subscore, the available scores from the most recent diary days (at least 4 days, up to 7 days) prior to actual day of the study visit will be summed, and then multiplied by the corresponding multiplier to get subtotal score. If available diary entries are fewer than 7 days, the subtotal score will be calculated as (summed total available score/number of days) × 7 × corresponding multiplier. The three subtotal scores that are based on ePRO (number of liquid/very soft stools, abdominal pain rating, and general well-being) will then be rounded to one decimal. The final CDAI is rounded to a whole number.

If a subject has less than 4 days of diary data, the total CDAI score will not be calculated and will be considered missing.

Same rule will be applied to average daily stool frequency and average daily abdominal pain.

Bristol Stool Chart

Up to 14 days of diary entries will be evaluated from the ePRO tool for the Bristol Stool Chart calculation. These assessments on the day of endoscopy procedure, the day before endoscopy procedure (due to preparation medications), and 2 days after endoscopy procedure are excluded. The 7 most recent non-missing assessments will be used for the analysis. The Bristol Stool Chart is calculated as the number of days the subjects with Type 6 or Type 7 divided by the total number days with non-missing assessments, rounded to two decimal.

If a subject has less than 4 days of diary data, the Bristol Stool Chart will not be calculated and will be considered missing.

If a subject achieved reduction from Baseline ≥ 50%, this subject will be classified as Responder.

SES-CD Scoring

All colonoscopies shall be performed and recorded in video format. To be eligible for the study, the screening colonoscopy may be performed during the Screening Period (or up to 45 days before the Baseline visit if Amendment 2 is in place at the site). Colonoscopies will also be performed at the Week 12/16 visit and Week 52 visit. Investigators shall provide their SES-CD score assessment on the SES-CD score sheet for the baseline, Week 12/16 visit and Week 52 visit. The investigator SES-CD scores are to be entered in the appropriate eCRF, however neither the Investigator's Baseline, nor Week 12/16, Week 52 SES-CD scores will be used for the study's efficacy analyses unless the adjudicator agree with the site reading.

All videotaped colonoscopies will undergo central review. One primary central reviewer will evaluate the videotaped colonoscopies and provide the SES-CD scores to Parexel. A second central reviewer will adjudicate between the investigator's and the central reviewer's SES-CD scores if there is discrepancy (more than 10% difference) in total SES-CD score. The adjudicator will select the final SES-CD score that he/she most

agrees with from those provided by the investigator and the central reviewer, and this final SES-CD score will be entered into the study database and be used for the study's efficacy analyses. If there is no discrepancy between the investigator and the central reviewer, that central reviewer's score will be entered into the study database and serve as the SES-CD score to be used for the study's efficacy analyses.

If there is a missing SES-CD individual variable in the SES-CD score for the study's efficacy analyses (for all variables except the "Presence of Narrowing" variable), the following imputation rules will be applied:

If the same individual variable is:

- Missing at Week 52 but present at Week 12/16: the missing Week 52 value will be imputed based on the value at Week 12/16.
- Missing at Week 52 but no endoscopy performed at Week 12/16: the missing Week 52 value will be imputed based on the value at Induction Baseline.

If there is a missing SES-CD "Presence of Narrowing" individual variable in the SES-CD score for the study's efficacy analyses, the following imputation rules will be applied:

If the Presence of Narrowing individual variable is:

- Missing at Week 52 but present at Week 12/16: the missing Week 52 value will be imputed based on score at Week 12/16.

  However, if a Week 52 individual variable is scored for intestinal segments proximal to (above) the segment with the missing Presence of Narrowing variable, the maximum imputed value for the missing Presence of Narrowing individual variable will be 2.
- Missing at Week 52 but no endoscopy performed at Week 12/16: the missing Week 52 value will be imputed based on score at Induction Baseline.

  However, if a Week 52 individual variable is scored for intestinal segments proximal to (above) the segment with the missing Presence of Narrowing

variable, the maximum imputed value for the missing Presence of Narrowing individual variable will be 2.

(Gastrointestinal anatomy places the ileum proximal to the right colon, which is proximal to the transverse colon, which is proximal to the sigmoid and left colon, which is proximal to the rectum.)

A sensitivity analysis (Worst Case Imputation) will be performed by imputing all missing Induction Baseline individual variables as '0' and all missing Week 52 variables as '3' except the Week 52 Presence of Narrowing individual variable. At Week 52, the worst values for Presence of Narrowing individual variable will be 2 for Rectum, Sigmoid and Left Colon, Transverse Colon, and Right Colon, and 3 for Ileum.

Isolated Ileal Crohn's Disease at Induction Baseline

Isolated Ileal Crohn's disease at Induction Baseline is defined as all individual variables for Rectum, Sigmoid and Left Colon, Transverse Colon, and Right Colon are zero at Induction Baseline and subscore for ileal ≥ 4. If any of the individual variables is missing or not zero, this subject is classified as Not Isolated Ileal Crohn's disease at Induction Baseline.

Definition of Endoscopic Remission

SES-CD ≤ 4 and at least two point reduction versus Induction baseline and no subscore > 1 in any individual variable.

Definition of Endoscopic Response

SES-CD at least 25% reduction from Induction baseline.

Definition of Enhanced Endoscopic Response

SES-CD reduction from Induction baseline > 50% (or for an Induction Baseline SES-CD of 4, at least a 2 point reduction from Induction Baseline).

Definition of Endoscopic Improvement

SES-CD reduction from Induction baseline > 50% or endoscopic remission.

Definition of Endoscopic Healing

SES-CD ulcerated surface subscore of 0 in subjects with SES-CD ulcerated surface subscore ≥ 1 at Induction Baseline.

Definition of Clinical Remission

Average daily stool frequency ≤ 1.5 and not worse than Induction baseline AND average daily abdominal pain ≤ 1.0 and not worse than Induction baseline.

Definition of Clinical Response

Average daily stool frequency at least 30% reduction from Induction baseline and average daily abdominal pain not worse than Induction baseline OR average daily abdominal pain at least 30% reduction from Induction baseline and average daily stool frequency not worse than Induction baseline.

Definition of Modified Clinical Remission

Average daily stool frequency ≤ 2.8 and not worse than Induction baseline AND average daily abdominal pain ≤ 1.0 and not worse than Induction baseline. This endpoint will only be analyzed among subjects with Induction baseline average daily stool frequency ≥ 4 or average daily abdominal pain ≥ 2.0.

Definition of Enhanced Clinical Response

Average daily stool frequency at least 60% reduction from Induction baseline and average daily abdominal pain not worse than Induction baseline OR average daily abdominal pain at least 35% reduction from Induction baseline and average daily stool frequency not worse than Induction baseline or modified clinical remission.

Definition of Remission

Remission at Week 52 is defined as both Endoscopic remission at Week 52 AND Clinical remission at Week 52.

Definition of Response

Response at Week 52 is defined as both Endoscopic response at Week 52 AND Clinical response at Week 52.

EuroQol-5D-5L (EQ-5D)

EQ-5D is a standardized measure of health status developed by the EuroQol Group in order to provide a simple, generic measure of health for clinical and economic appraisal. The EQ-5D consists of 2 pages. The first page measures 5 dimensions of the health status (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) with 5 levels per dimension (no problems, slight problems, moderate problems, severe problems, and unable corresponding to Level 1 to Level 5 respectively). The second page is an EQ Visual Analogue Scale (EQ VAS). EQ-5D health states, defined by the EQ-5D-5L descriptive system on the first page, may be converted into a single index value. The change from baseline of the index value and EQ VAS will be analyzed and reported. UK scoring algorithm will be used (Table 5).

Table 5. An EQ-5D-5L Value Set for England

	Central Estimate	Value for Health State 23245
Mobility		
No Problem	0	
Slight	0.051	0.051
Moderate	0.063	
Severe	0.212	
Unable	0.275	
Self-Care		
No Problem	0	
Slight	0.057	
Moderate	0.076	0.076
Severe	0.181	
Unable	0.217	
Usual Activities		
No Problem	0	
Slight	0.051	0.051
Moderate	0.067	
Severe	0.174	
Unable	0.190	
Pain/Discomfort		
No	0	
Slight	0.060	
Moderate	0.075	
Severe	0.276	0.276
Unable	0.341	
Anxiety/Depression		
Not	0	
Slight	0.079	
Moderate	0.104	
Severe	0.296	
Unable	0.301	0.301
The value for health state 23245: 1 – 0.9675 × (sum of the subscores)		**1 – 0.9675** × (0.051 + 0.076 + 0.051 + 0.276 + 0.301) = 0.270

The index value is 1 – 0.9675 × sum of 5 components based on central estimates.

The minimum index value is – 0.281 (health state 55555), and the maximum index value is 1 (health state 11111).[1]

If one of the 5 dimensions is missing, the EQ-5D index value will be missing.

7.0 Demographics, Baseline Characteristics, Medical History, and Previous/Concomitant Medications

7.1 Demographic and Baseline Characteristics

For the subjects in ITT and safety analysis set, demographic information and Baseline values will be summarized by descriptive statistics. Categorical data will be summarized by number and percent; and quantitative data will be presented by n, mean, standard deviation, minimum value, median, and maximum value.

In general, continuous variables will be analyzed using analysis of variance (using SAS procedure 'PROC GLM') with treatment group as factor. Categorical variable will be analyzed using chi-square test or Fisher's exact test if ≥ 20% of the cells have expected cell count < 5.

The following demographic and Baseline values will be summarized.

Continuous Variables:

- Age (years)
- Body weight (kg)
- Height (cm)
- Body Mass Index (kg/m^2)
- Blood Pressure (systolic/diastolic) (mmHg)
- Pulse (bpm)
- Temperature (°C)
- CDAI
- Average Daily Stool Frequency

- Average Daily Abdominal Pain
- Total SES-CD score
- Crohn's Disease Duration (years)
- hs-CRP mg/L
- Fecal calprotectin
- IBDQ score
- WPAI and its components
- EQ-5D
- Abdominal Pain Rating Scale
- Bristol Stool Chart Score
- Mean dose of corticosteroids for those on baseline corticosteroid use.

Categorical Variables:

- Sex (male, female)
- Race
- Ethnicity
- Age (< 40, 40 to < 65, 65 to < 75, ≥ 75)
- Baseline corticosteroid use (yes, no)
- Baseline immunosuppressant use (yes, no)
- Number of prior immunosuppressant use
- Number of prior biologic use
- hs-CRP at Baseline (< 10 and ≥ 10 mg/L)
- Baseline fecal calprotectin (≤ 250 µg/g, > 250 µg/g)
- Crohn's disease severity (CDAI ≤ 300, > 300) at Baseline
- Baseline SES-CD (< 15, ≥ 15)
- Disease duration (≤ 3 years, > 3 years)
- Tobacco use (user, ex-user, never used, unknown)
- Alcohol use (drinker, ex-drinker, non-drinker, unknown)
- Region (US, ex-US)

7.2 Medical History

Medical and Surgical History: A complete medical and surgical history (which includes CD-onset date), history of tobacco and alcohol use, and TB history will be obtained from each subject during the Screening period. Medical history will be summarized using body system and condition/diagnosis by treatment group. No statistical tests will be performed.

Chest X-Ray Results: All subjects undergo a standard X-ray of chest (including a posteroanterior [PA] and lateral view) at Screening period. Number and percent of subjects with presence or absence of finding for the previous TB infection, calcified granulomas, Pleural scarring/thickening, and other findings will be presented by treatment group. No statistical tests will be performed.

TB Test Results: Results of PPD skin test, QuantiFERON-TB Gold test at screening visit will be summarized. Induration will be summarized descriptively using n, mean, standard deviation, minimum values, median, and maximum values. The frequency distribution of induration ≥ 5 and < 5 will be provided. QuantiFERON-TB tests will be described as positive or negative. Indetermiante QuantiFERON-TB test results will be repeated. If the second QuantiFERON-TB test is positive or indeterminate, the final assessment will be considered positive. If the second QuantiFERON-TB test is negative, the final assessment will be considered negative. No statistical tests will be performed.

TB Prophylaxis: History of use of TB Prophylaxis or initiation of TB prophylaxis will be summarized.

ECG Results: ECG results at screening will be presented as frequency distribution showing results as Normal, Abnormal (Not clinically significant), Abnormal (Clinically significant) and Unable to evaluate/missing. No statistical tests will be performed.

7.3 Previous Treatment and Concomitant Medications

Based on generic medication names, these categories of medications used by subjects before and during the study will be summarized by number and percent for ITT and safety analysis sets for the treatment groups. No statistical tests will be performed.

The number and percent of subjects using Crohn's disease specific medications (including corticosteroids, aminosalicylates, immunosuppressants [defined as azathioprine, 6-mercaptopurine, or methotrexate], antibiotics) within past 90 days prior to the Baseline, and at the Baseline will be tabulated. In addition, the number and percent of subjects using Crohn's disease specific immunosuppressants, and biologic therapies at any time prior to Baseline will be tabulated.

8.0 Patient Disposition

Subject disposition will be presented for subjects in the ITT and safety analysis sets using the following information by treatment group:

- Number and percent of subjects in various analysis sets by treatment group and by investigator and/or site number
- Number and percent of subjects completing extension phase and discontinuing on or before Week 52 visit.
- Subject disposition including the number and percent of subjects who prematurely discontinued the study) by primary reason and by any reason

Summary of protocol deviations will be provided.

9.0 Study Drug Exposure and Compliance

Study drug exposure and compliance will be summarized using the mean, standard deviation, minimum, median, and maximum. Exposure to study drug will be summarized by separately for double-blind extension phase and open-label extension phase.

Compliance is defined as the number of capsules taken (i.e., the difference between the number of capsules dispensed and the number of capsules returned) divided by the number of capsules a subject is supposed to take each day times the length of time that the subject was in the Treatment Phase of the study (i.e., Final/Discontinuation Visit date during Treatment Phase – Day 1 [Baseline] Visit date + 1). Subjects with missing data for the number of capsules returned will be excluded from the summary.

10.0 Efficacy Analysis

10.1 General Considerations

All statistical tests will be two-sided with the significance level of 0.100. Descriptive statistics will be provided. These include the number of observations, mean, standard deviation, minimum, median, and maximum for continuous variables; and number and percent for discrete variables. The analysis will be performed using SAS® (SAS Institute Inc., Cary, NC, USA).

10.2 Efficacy Analyses

The endpoints (Double-Blind Extension Phase) include:

- Proportion of subject who achieve remission at Week 52.
- Proportion of subjects who achieve endoscopic remission at Week 52.
- Proportion of subjects who achieve both endoscopic remission and modified clinical remission at Week 52.
- Proportion of subjects who achieve clinical remission over time.
- Proportion of subjects who maintain clinical remission over time among subjects in clinical remission at Week 16.
- Proportion of subjects who achieve modified clinical remission over time.
- Proportion of subjects who maintain modified clinical remission over time among subjects in modified clinical remission at Week 16.
- Proportion of subject who achieve response at Week 52.
- Proportion of subjects with SES-CD ≤ 2 at Week 52.

- Proportion of subjects with SES-CD = 0 at Week 52.
- Proportion of subjects who achieve endoscopic response at Week 52.
- Proportion of subjects who achieve enhanced endoscopic response at Week 52.
- Proportion of subjects who achieve endoscopic improvement at Week 52.
- Proportion of subjects who achieve endoscopic healing at Week 52.
- Proportion of subjects who achieve clinical response over time.
- Proportion of subjects who achieve enhanced clinical response over time.
- Proportion of subjects who maintain enhanced clinical response over time among subjects in enhanced clinical response at Week 16.
- Proportion of subjects with an average daily SF $\geq$ 2.5 AND average daily AP $\geq$ 2.0 at Induction Baseline who achieve clinical remission over time.
- Proportion of subjects taking corticosteroids at Induction Baseline who discontinued corticosteroid use and achieved CDAI < 150 over time.
- Proportion of subjects taking corticosteroids at Induction Baseline who discontinued corticosteroid use and achieve remission at Week 52.
- Proportion of subjects taking corticosteroids at Induction Baseline who discontinued corticosteroid use and achieve clinical remission over time.
- Proportion of subjects taking corticosteroids at Induction Baseline who discontinued corticosteroid use and achieve modified clinical remission over time.
- Proportion of subjects taking corticosteroids at Induction Baseline who discontinued corticosteroid use and achieve endoscopic remission at Week 52.
- Proportion of subjects who achieve CDAI < 150 at over time.
- Proportion of subjects with decrease in CDAI $\geq$ 70 points from Induction Baseline over time.
- Change from Induction Baseline in fecal calprotectin level over time.
- Change from Induction Baseline in hs-CRP at over time.
- Change from Induction Baseline in IBDQ at Week 52.
- Change from Induction Baseline in EQ-5D at Week 52.
- Change from Induction baseline in EIMs at Week 52.

- Proportion of subjects with isolated ileal Crohn's disease who achieve remission at Week 52.
- Proportion of subjects with isolated ileal Crohn's disease who achieve modified clinical remission at Week 52.

For categorical efficacy endpoints for double-blind extension phase, the pairwise comparisons for the difference in proportions of subjects between ABT-494 3 mg BID group and ABT-494 higher dose groups will be analyzed using the Chi-square test (or Fisher's exact test if ≥ 20% of the cells have expected cell count < 5.). Additionally, the 95% confidence interval based on normal approximation for the difference in proportions will be provided.

The non-responder imputation will be used for subjects with missing data at the endpoint evaluated. The last observation carried forward (LOCF) method will also be used as the sensitivity analyses.

In general, continuous efficacy variables with repeated measurements will be analyzed using a Mixed Effect Repeated Measure (MMRM) model. The mixed model includes the categorical fixed effects of treatment, week and treatment-by-week interaction and the continuous fixed covariates of Induction baseline measurement. An unstructured variance-covariance matrix will be used. If convergence is not achieved, other variance-covariance matrix will be explored until coverage. The parameter estimations will be based on assumption of data being missing at random and using the method of restrictive maximum likelihood (REML).

Continuous secondary efficacy variables will also be analyzed using an Analysis of Covariance (ANCOVA) model including factors for treatment group, and Induction Baseline values as a covariate. The MMRM analysis is considered primary for inferential purposes.

10.3 Handling of Multiplicity

The analyses are for extension phase endpoints only, and will not be controlled for multiplicity.

10.4 Efficacy Subgroup Analysis

No subgroup analyses will be conducted for double-blind extension phase efficacy endpoints.

11.0 Safety Analysis

11.1 General Considerations

Adverse events (AEs), laboratory data and vital signs are the primary safety parameters in this study. The safety variable will be summarized by treatment according to the treatment a subject actually received.

11.2 Analysis of Adverse Events

11.2.1 Treatment-Emergent Adverse Events

Treatment-emergent AE for double-blind Extension Phase is defined as any adverse event with an onset date on or after the first dose of double-blind maintenance dose and prior to the open-label dose or up to 30 days after the last dose of double-blind maintenance dose if subject discontinued prematurely in double-blind maintenance phase.

For subjects who entered open-label due to inadequate response, treatment-emergent AEs for open-label Extension Phase will also be summarized by open-label dose. Treatment-emergent AE for open-label Extension Phase is defined as any adverse event with an onset date on or after the first dose of open-label dose and up to 30 days after the last dose of open-label dose.

An overview of treatment-emergent AEs, including AEs of special interest such as adverse events leading to death and adverse events leading to early termination, AEs by

Medical Dictionary for Drug Regulatory Activities (MedDRA version 20.0 or later) preferred term and system organ class, AEs by maximum relationship to study drug, and AEs by maximum severity will be summarized by number and percentage.

Changes in laboratory data will be summarized.

In addition, shift tables and listings will be provided for abnormal values, whereby the normal range of the analyzing laboratory will be used. Vital signs will be analyzed similarly.

The number and percent of subjects experiencing treatment-emergent adverse events will be summarized for the following adverse event categories.

- Any treatment-emergent adverse event.
- Any treatment-emergent adverse event that was rated as possibly related to study drug by the investigator (Reasonable Possibility).
- Any treatment-emergent severe adverse event.
- Any treatment-emergent serious adverse event.
- Any treatment-emergent adverse event leading to discontinuation of study drug.
- Any treatment-emergent adverse event leading to death.
- Any treatment-emergent adverse event of special interest.

Treatment-emergent adverse events will be summarized as follows:

- Grouped by System Organ Class and Preferred Term.
- Grouped by System Organ Class, Preferred Term and Severity.
- Grouped by System Organ Class, Preferred Term and Relationship to Study Drug.
- Grouped by System Organ Class and Preferred Term with subject numbers.

In treatment-emergent AE tables, a subject who reports more than one treatment-emergent AE in different system organ classes will be counted only once in the overall total. A subject who reports two or more different preferred terms which are in the same SOC will be counted only once in the SOC total. A subject who reports more than one treatment AE with the same preferred term will be counted only once for that preferred term using the most extreme incident (i.e., most "severe" for the severity tables and most "related" for the relationship tables).

Adverse events will also be summarized by maximum severity. If a subject has an adverse event with unknown severity, then the subject will be counted in the severity category of "unknown," even if the subject has another occurrence of the same adverse event with a severity present. The only exception is if the subject has another occurrence of the same adverse event with the most extreme severity – "Severe." In this case, the subject will be counted under the "Severe" category.

Adverse events will also be summarized by maximum relationship to study, as assessed by the investigator. If a subject has an adverse event with unknown relationship, then the subject will be counted in the relationship category of "unknown," even if the subject has another occurrence of the same adverse event with a relationship present. The only exception is if the subject has another occurrence of the same adverse event with a relationship assessment of "Reasonable Possibility." In this case, the subject will be counted under the "Reasonable Possibility" category, respectively.

Incidence rates per 100 patient years of exposure to study drug will be presented for AE overviews and for AEs by SOC and preferred term where the number of events will be used as the numerator.

$$100 \times \frac{Number\ of\ TEAEs}{Total\ Patient\ Years}$$

where total patient years is defined as the sum of the study drug exposure of all subjects, as defined in an earlier section, normalized by 365.25, and rounded to 1 decimal place.

Overview of treatment-emergent adverse events and treatment-emergent adverse events during the administration of ABT-494 grouped by System Organ Class and Preferred Term will also be summarized for the Any ABT-494 analysis set by quartile of average daily dose.

11.2.2 Adverse Events of Special Interest

The Adverse Events of Special Interests (AESI) categories will be summarized and presented for each treatment group using SMQ/CMQ/LMQ/Preferred Term.

AESI	Type of MedDRA Query	Broad or Narrow Search	SMQ/CMQ/LMQ Search Criteria
Serious Infections	CMQ		"Infection" – Subset for SAEs
Opportunistic Infection	CMQ		"Opportunistic infection"
Herpes zoster	CMQ		"Herpes zoster"
Tuberculosis	CMQ		"Tuberculosis"
Malignancy	SMQ	Narrow	"Malignancies,"
Non-Melanoma Skin Cancer (NMSC)	SMQ	Broad	Skin Malignant tumours (Broad SMQ) removing Melanoma CMQ
Malignancy other than NMSC	SMQ	Narrow	Malignancy and removing NMSC output
Lymphoma	SMQ		"Malignant Lymphomas"
Hepatic Disorder	SMQ	Narrow	"Drug Related Hepatic disorders"
Intestinal Perforations	LMQ		Terms for/related to any part of GI Tract, GI Ulcers, obstruction, hemorrhage
Anemia	CMQ		"Non-Hemolytic and Non-Aplastic anemias"
Adjudicated Cardiovascular Events	Output from CAC		
Neutropenia	CMQ		"Hematological toxicity – Neutropenia"
Lymphopenia	CMQ		"Hematological Toxicity – Lymphopenia (Veliparib Product Specific)"
Renal dysfunction	SMQ	Narrow	"Acute renal failure"
Rhabdomyolysis/myopathy	SMQ	Narrow	"Rhabdomyolysis/Myopathy"
Creatine phosphokinase (CPK) Elevation	Preferred Term		"blood creatine phosphokinase increased"

Additional AEs may be considered for tabulation/summary based on recommendations from Clinical and Safety as deemed appropriate.

11.3 Analysis of Laboratory Data

Changes from Induction Baseline during Extension in continuous laboratory parameters will be summarized by n, mean, standard deviation, and median value value for each treatment group.

Cross (Shift) tables from Induction Baseline to the final value according to the normal range will be provided for each hematology, clinical chemistry parameter and urinalysis parameter except for the microscopic examination.

For selected laboratory parameter with Common Toxicity Criteria (CTC) a listing of all subjects with any laboratory determinations meeting CTC Version 4.03 (or later) of Grade ≥ 3 will be provided. For hemoglobin, an additional list will also be provided based on CTC Version 3.0 (or later) of Grade ≥ 2. For each of these subjects, the whole course of the parameter will be listed. For subjects with laboratory values with CTC ≥ 3, all of the laboratory parameters for those subjects will be listed. For subjects with laboratory values with CTC 4, all of the laboratory parameters for those subjects will be listed.

For subjects who switched op open-label therapy, the above tables will also be presented by the open-label dose.

In addition, Cross (Shift) tables from Induction Baseline to the final value according to the normal range will be provided for each hematology, clinical chemistry parameter and urinalysis parameter except for the microscopic examination for the Any ABT-494 analysis set by quartile of average daily dose.

11.3.1 Analysis of Liver Specific Laboratory Tests

According to FDA's Guidance for Industry "Drug-Induced Liver Injury: Premarketing clinical evaluation" (July 2009), when aminotransferase (AT) abnormalities indicating hepatocellular injury are accompanied by evidence of impaired hepatic function (bilirubin elevation > 2 × ULN), in the absence of evidence of biliary obstruction (i.e., significant

elevation of ALP) or some other explanation of the injury (e.g., viral hepatitis, alcohol hepatitis), the combined finding (i.e., Hy's Law cases) represents a signal of a potential for the drug to cause severe DILI.

For the purpose of assessing for potential Hy's law cases, the frequencies and percentages of subjects with post baseline liver specific function test values that meet the following criteria of potential clinical interest will be summarized by "as treated" treatment group:

- ALT ≥ 3 × ULN
- ALT ≥ 5 × ULN
- ALT ≥ 10 × ULN
- ALT ≥ 20 × ULN
- AST ≥ 3 × ULN
- AST ≥ 5 × ULN
- AST ≥ 10 × ULN
- AST ≥ 20 × ULN
- TBL ≥ 2 × ULN
- Alkaline phosphatase ≥ 1.5 × ULN
- ALT and/or AST ≥ 3 × ULN and concurrent TBL ≥ 1.5 × ULN
- ALT and/or AST ≥ 3 × ULN and concurrent TBL ≥ 2 × ULN

11.4 Analysis of Vital Signs and Weight

The following vital signs are measured at every visit during the study.

- Body Weight (kg)
- Blood Pressure (Systolic/Diastolic) (mmHg)
- Pulse (bpm)
- Temperature (°C)

Changes from Induction Baseline in vital sign values will be summarized by n, mean, standard deviation, and median value for each treatment group.

In additional, incidence of potential clinically significant results will be summarized.

Criteria for potentially clinically significant vital signs results:

Vital Sign	Category	Criteria for Potentially Clinically Significant Vital Sign
Systolic blood pressure	Low	Value ≤ 90 mmHg and decrease ≥ 20 mmHg from Baseline
	High	Value ≥ 160 mmHg and increase ≥ 20 mmHg from Baseline
Diastolic blood pressure	Low	Value ≤ 50 mmHg and decrease ≥ 15 mmHg from Baseline
	High	Value ≥ 105 mmHg and increase ≥ 15 mmHg from Baseline
Pulse	Low	Value ≤ 50 bpm and decrease ≥ 15 bpm from Baseline
	High	Value ≥ 120 bpm and increase ≥ 15 bpm from Baseline

For subjects who switched op open-label therapy, the above tables will also be presented by the open-label dose.

12.0 Summary of Changes

12.1 Summary of Changes Between the Latest Version of Protocol and the Current SAP

The following efficacy variables are not mentioned in the protocol:

Modified Clinical Remission: Average daily stool frequency ≤ 2.8 and not worse than baseline AND average daily abdominal pain ≤ 1.0 and not worse than Induction baseline. This endpoint will only be analyzed among subjects with baseline average daily stool frequency ≥ 4 or average daily abdominal pain ≥ 2.0.

Enhanced Clinical Response: Average daily stool frequency at least 60% reduction from baseline and average daily abdominal pain not worse than baseline OR average daily abdominal pain at least 35% reduction from baseline and average daily stool frequency not worse than baseline.

Enhanced Endoscopic Response: SES-CD reduction from Induction baseline > 50% (or for a Baseline SES-CD of 4, at least a 2 point reduction from Baseline).

Endoscopic Improvement: SES-CD reduction from Induction baseline > 50% or endoscopic remission.

Endoscopic Healing: SES-CD ulcerated surface subscore of 0 in subjects with SES-CD ulcerated surface subscore ≥ 1 at Induction Baseline.

SES-CD ≤ 2

SES-CD = 0

12.2 Summary of Changes Between the Previous Version and the Current Version of the SAP

13.0 Appendix

14.0 References

1. Devlin N, Shah K, Feng Y, et al. Valuing health-related quality of life: an EQ-5D-5L value set for England, January 2016. Available from: https://www.ohe.org/publications/valuing-health-related-quality-life-eq-5d-5l-value-set-england or http://www.euroqol.org/about-eq-5d/valuation-of-eq-5d/eq-5d-5l-value-sets.html.

1.0 Title Page

Statistical Analysis Plan for Induction Period

Study M13-740

A Multicenter, Randomized, Double-Blind, Placebo-Controlled Study of ABT-494 for the Induction of Symptomatic and Endoscopic Remission in Subjects with Moderately to Severely Active Crohn's Disease who have Inadequately Responded to or are Intolerant to Immunomodulators or Anti-TNF Therapy

05 December 2016

2.0 Table of Contents

List of Tables

List of Figures

3.0 Introduction

This statistical analysis plan (SAP) describes the statistical analysis to be completed by the AbbVie Clinical Statistics Department for study Protocol M13-740 Amendment 4 dated 15 April 2016. It provides details to further elaborate statistical methods as outlined in the protocol and describes analysis conventions to guide the statistical programming work.

This is the first version of the SAP for Protocol M13-740.

This analysis plan describes the primary and secondary efficacy analyses as well as the safety analysis for the double-blind Induction Period. The statistical analyses presented in this SAP represent the analyses for protocol Section 8.1.10 "Interim Analyses" and are the final analyses of the Induction Period of this study, and will be referred to as "Final Induction Analyses" thereafter in this SAP.

This document describes the analysis of data except pharmacokinetic, microbiota metagenomic, pharmacogenetic and serologic data which will be analyzed separately. It takes into account ICH Guidelines E3 and E9.

Unless noted otherwise, all analyses will be performed using SAS version 9.2 or later (SAS Institute Inc., Cary, NC 27513) under the UNIX operating system.

4.0 Study Objectives, Design and Procedures

4.1 Objectives

The objective of this study is to determine the efficacy and safety of multiple doses of ABT-494 versus placebo and to assess the pharmacokinetics of ABT-494 following oral administration in subjects with moderately to severely active Crohn's disease with a history of inadequate response to or intolerance to immunomodulators or anti-TNF therapy.

4.2 Design Diagram

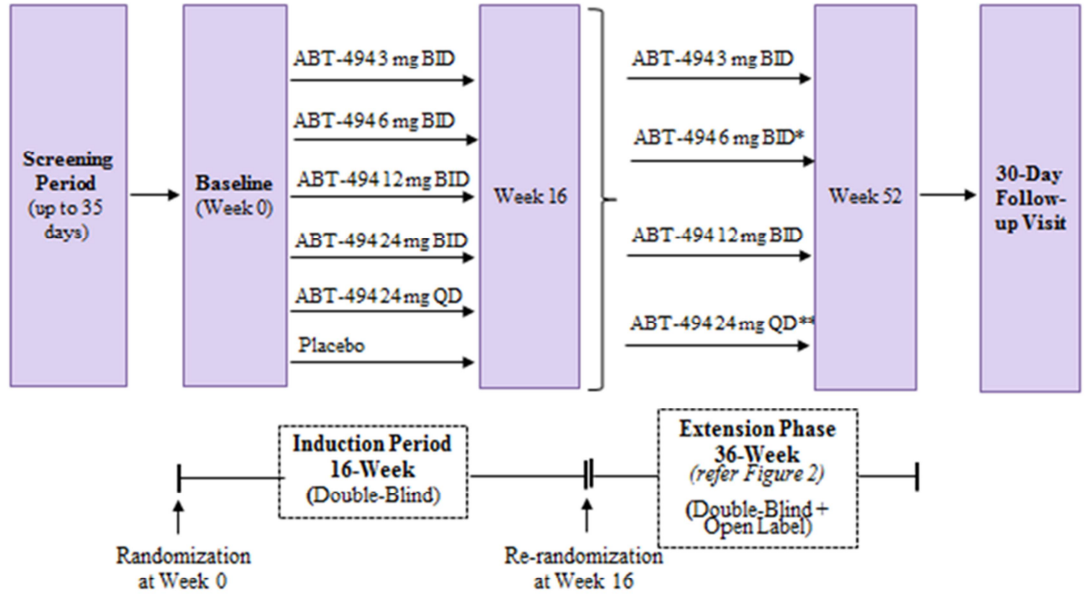
This is a Phase 2, multicenter, randomized, double-blind, placebo-controlled study designed to evaluate the efficacy, safety, and pharmacokinetics of ABT-494 as induction therapy in subjects with moderately to severely active Crohn's disease and evidence of mucosal inflammation defined by a SES-CD ≥ 6 (or ≥ 4 for subjects with disease limited to the ileum) and average daily soft/liquid stool frequency ≥ 2.5 or average daily abdominal pain score of ≥ 2.0; and CDAI ≥ 220 and ≤ 450, with a history of inadequate response or intolerance to immunomodulators and/or anti-TNF therapy. The study will allow enrollment of up to 30% of subjects who were primary non responders to anti-TNF treatment.

The study is designed to enroll 210 subjects to meet scientific and regulatory objectives without enrolling an undue number of subjects in alignment with ethical considerations. Therefore, if the target numbers of subjects have been enrolled, there is a possibility that additional subjects in screening will not be enrolled.

The study duration could be up to 60 weeks, including a Screening Period of up to 35 days, a 16-week double-blind induction period, a 36-week double-blind extension phase, and a 30-day follow-up period. Subjects who meet eligibility criteria will be randomized in a 1:1:1:1:1:1 ratio to one of the double-blinded induction treatment arms.

The schematics of the overall study design are shown in Figure 1 and Figure 2.

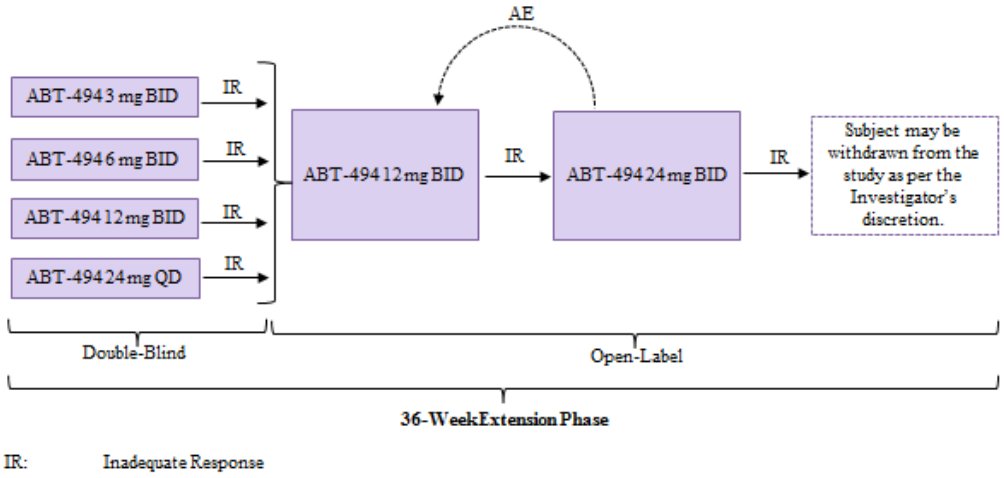
Figure 1. Study Design Schematic



*The 6 mg BID dose arm was added to replace the 24 mg QD dose arm in Protocol Amendment 2

**Subjects who were re-randomized at Week 16 to 24 mg QD prior to Protocol Amendment 2 will continue to receive the double-blind 24 mg QD dose until Week 52/PD.

Figure 2. Study Design Schematic for Open-Label Extension Option for Patients Who Do Not Adequately Respond During the Extension Phase



IR: Inadequate Response

Screening Period

Within 35 days prior to the Baseline Visit, subjects will receive a full explanation of the study design and study procedures, provide a written informed consent, and undergo the screening procedures as outlined in Table 1.

An e-diary will be dispensed and training will be provided to the subject at the Screening Visit.

Laboratory tests performed during the Screening period can be repeated in case the abnormalities are considered to be transient by the investigator.

16-Week Induction Period

This period will begin at the Baseline Visit (Week 0) and will end at the Week 16 Visit. At the Baseline Visit, subjects who meet all the inclusion criteria and none of the exclusion criteria described in the protocol (Section 5.2.1 and Section 5.2.2) will be

enrolled into the study and randomized to double-blind induction period. The randomization at Baseline will be stratified by endoscopic disease severity (SES-CD < 15 and ≥ 15), prior anti-TNF use (naïve and experienced), and by participation in the substudy of the gene expression in intestinal biopsies (Yes and No). As part of the randomization at Baseline, subject will be randomly assigned (1:1) to have their follow-up ileocolonoscopy done at Week 12 or Week 16. During this period of the study, subjects will visit the study site at Weeks 2, 4, 8, 12 and 16. A ± 3-day window is permitted around scheduled study visits. The last dose of study drug during this period is taken the evening prior to the Week 16 visit.

However, the stratification factor Prior anti-TNF use (naïve and experienced) was added under Amendment 3, and there are less than 10 anti-TNF naïve subjects enrolled under Amendment 3. Therefore, the Prior anti-TNF use will not be included in all analysis as a stratification factor.

Subjects may discontinue study drug treatment at any time during study participation. Subjects that end study participation early will have an Early Termination (ET) Visit and complete the procedures outlined for the ET Visit in Table 1 as soon as possible after the last dose of study drug and preferably prior to the administration of any new therapies.

36-Week Extension Phase

At Week 16, subjects who have completed the induction period will be re-randomized in a ratio of 1:1:1 to one of the three double-blinded doses of ABT-494 3 mg BID, 6 mg BID or 12 mg BID. The re-randomization will be stratified by dose received during the first 16 weeks, and overall response (responder versus non-responder) at Week 16. The subjects who were re-randomized at Week 16 to ABT-494 24 mg QD prior to Protocol Amendment 2 will continue to receive the same double-blind dose until Week 52/PD.

The central reader endoscopic score will be used for calculating the Endoscopic response for the evaluation of the efficacy endpoints. However, for stratification at the time of re-randomization, the endoscopic score at BL from central reader and the endoscopic

score at Week 12 or Week 16 from site local reader will be used in order to determine response status.

Clinical response is defined as average daily liquid/soft stool frequency reduction of at least 30% from Baseline and average daily abdominal pain not worse than Baseline OR average daily abdominal pain reduction at least 30% from baseline and average daily stool frequency not worse than Baseline.

During this period of the study, subjects will visit the study site at Weeks 20, 28, 36, 44, and 52/Early Termination. A ± 3-day window is permitted around scheduled study visits. The last dose of study drug is taken the evening prior to the Week 52 visit.

Subjects will be expected to remain on double-blinded therapy throughout the entire 36-week extension phase. However, subjects who are considered by the investigator to have not achieved meaningful symptomatic relief and meet the criteria for inadequate response at or after Week 20 will be eligible to receive the open-label therapy with ABT-494 12 mg BID, Figure 2. Subjects who were re-randomized at Week 16 prior to Protocol Amendment 2 and are taking open-label ABT-494 24 mg QD will be transitioned to ABT-494 12 mg BID.

Subjects who still continue to meet criteria for inadequate response following a 4-week course of open-label ABT-494 12 mg BID will be eligible to dose escalate to open-label ABT-494 24 mg BID.

Subjects with persistent inadequate response while on ABT-494 12 mg BID or 24 mg BID may be withdrawn from the study at the investigator's discretion.

Criteria for Inadequate response are as follows:

- Average daily liquid/soft Stool Frequency > 2.2 OR average daily Abdominal Pain score > 1.8 AND
- An increase level of hs-CRP of at least 1 mg/L from Baseline or a hs-CRP ≥ 5 mg/L, either at the previous visit or at the current visit.

Note: hs-CRP will be performed by Central Lab. However, for the purpose of evaluating Inadequate Response criteria, hs-CRP can also be done at local lab.

Assessment of inadequate response should include consideration by the Investigator to rule out symptoms caused by reasons other than Crohn's disease related inflammation.

In the open-label period, if the subject treated with 24 mg BID has an occurrence of an adverse event thought to be possibly related to study medication that in the opinion of the investigator warrants dose reduction, then the dose can be de-escalated to ABT-494 12 mg BID at the investigator's discretion.

Note: Dose escalation and de-escalation are permitted only once during the study.

Subjects may discontinue study drug treatment at any time during study participation. Subjects that end study participation early will have an ET Visit and complete the procedures outlined for the ET Visit in Table 1 as soon as possible after the last dose of study drug and preferably prior to the administration of any new therapies.

Subjects who are not re-randomized at Week 16 or subjects who early terminated from the study will have a follow-up visit approximately 30 days after the last administration of study drug to obtain information on any new or ongoing adverse events (AEs), and to collect vital signs and clinical laboratory tests.

Follow-Up Period

Subjects who have completed the Week 52 visit will have a follow-up visit approximately 30 days after the last administration of study drug to obtain information on any new or ongoing adverse events (AEs), and to collect vital signs and clinical laboratory tests.

Re-Screen

Subjects that initially screen fail for the study may be permitted to re-screen following re-consent. The subject must meet all the inclusion and none of the exclusion criteria at the time of re-screening in order to qualify for the study. There is no minimum period of

time a subject must wait to re-screen for the study. If the subject had a complete initial screening evaluation including the assessment of a purified protein derivative (PPD) test (or equivalent), or Interferon-Gamma Release Assay (IGRA; QuantiFERON-TB Gold test or T-SPOT TB test), chest x-ray, HBV, HCV and ECG, these tests will not be required to be repeated for re-screening provided the conditions noted in the protocol (Section 5.3.1.1) are met and no more than 90 days have passed.

An endoscopy with biopsy will not be required to be repeated for re-screening provided the conditions noted in the protocol (Section 5.3.1.1) are met and no more than 30 days have passed (from the previous screening endoscopy). All other screening procedures will be repeated. As appropriate, sites are encouraged to contact the AbbVie Medical Monitor to confirm if subjects should or should not be re-screened.

Unscheduled Visits

Unscheduled Visits are for purposes when the subject is coming in for a medical visit for evaluation and assessment. During Unscheduled Visits, blood and urine samples will be obtained for the laboratory tests listed in the protocol (Table 2).

Visits for dispensing new study drug in case of temperature excursion, loss or damage are not considered an Unscheduled Visit. In addition, visits to only retest a lab will not be considered an Unscheduled Visit.

Table 1. Study Activities

Activity	Screening[a]	Baseline (Week 0)	Wk 2	Wk 4	Wk 8	Wk 12	Wk 16	Wk 20	Wk 28	Wk 36	Wk 44	Wk 52	Early Termination (ET) Visit	Unscheduled Visit	30-Day Follow-Up Visit
Informed Consent	X														
Inclusion/Exclusion[b]	X	X													
Medical/Surgical History[b]	X	X													
Prior and Concomitant Medications[b]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Physical Exam[c]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Endoscopy[d]	X					X	X					X	X		
Biopsy[e]	X					X	X					X	X		
SES-CD		X				X	X					X	X		
Vital Signs[f]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
12-Lead ECG[g]	X													X[v]	
Chest X-Ray[h]	X													X[v]	
PPD Skin Test or QuantiFERON TB Gold Test[i]	X														
Latent TB Risk Factor Questionnaire	X														
Blood Chemistry and Hematology	X[w]	X	X	X	X	X	X	X	X	X	X	X	X	X	X

Table 1. Study Activities (Continued)

Activity	Screening[a]	Baseline (Week 0)	Wk 2	Wk 4	Wk 8	Wk 12	Wk 16	Wk 20	Wk 28	Wk 36	Wk 44	Wk 52	Early Termination (ET) Visit	Unscheduled Visit	30-Day Follow-Up Visit
Urinalysis[j]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Pregnancy Test[k]	X	X	X	X	X	X	X	X	X	X	X	X	X		X
hsCRP	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
HBV and HCV Screening[l]	X														
ABT-494 Concentration[m]			X	X	X	X	X	X	X	X	X	X	X		
Pharmacodynamic biomarkers		X			X		X					X	X		
C. difficile toxin	X														
Stool Sample (fecal calprotectin)[n,o]	X			X			X		X			X	X		
Bristol Stool Chart	X			X			X		X			X	X		
Corticosteroid Taper[q]			X												
Crohn's Disease Activity Index (CDAI)		X	X	X	X	X	X	X	X	X	X	X	X	X	
Inflammatory Bowel Disease Questionnaire (IBDQ)		X			X		X					X	X		

Table 1. Study Activities (Continued)

Activity	Screening[a]	Baseline (Week 0)	Wk 2	Wk 4	Wk 8	Wk 12	Wk 16	Wk 20	Wk 28	Wk 36	Wk 44	Wk 52	Early Termination (ET) Visit	Unscheduled Visit	30-Day Follow-Up Visit
European Quality of Life 5 Dimensions (EQ-5D)		X			X		X					X	X		
Work Productivity and Impairment Questionnaire (WPAI)		X			X		X					X	X		
Abdominal Pain Rating Scale (0 – 10 Scale)		X				X	X								
Monitor Adverse Events[r]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Study Drug Dispensing/ Administration		X	X	X	X	X	X	X	X	X	X				
Dispense Subject Diary	X														
Subject Diary Review		X	X	X	X	X	X	X	X	X	X	X	X	X	

Table 1. Study Activities (Continued)

Activity	Screening[a]	Baseline (Week 0)	Wk 2	Wk 4	Wk 8	Wk 12	Wk 16	Wk 20	Wk 28	Wk 36	Wk 44	Wk 52	Early Termination (ET) Visit	Unscheduled Visit	30-Day Follow-Up Visit
Additional Samples Collection (Optional)															
Pharmacogenetic[s]		X													
Serologic markers[t]		X					X								
mRNA[t]		X					X								
Biopsy for gene expression[u]	X					X	X								

Wk = Week

a. The Screening period will be a minimum of 7 days but no more than 35 days. Baseline Visit date will serve as the reference for all subsequent visits. A ± 3-day window is permitted around scheduled study visits.

b. Update inclusion/exclusion, prior and concomitant therapy, and medical/surgical history information to assure subject eligibility.

c. Physical examination performed at Screening, Baseline, Week 16 and Week 52/Early Termination Visits are full physical examinations which must include an assessment of extra-intestinal manifestations (EIMs) and a count of the number of cutaneous fistulas. Physical examinations and those performed at all other visits are symptom based but must include a count of the number of cutaneous fistulas.

d. Colonoscopy at/during the screening period or within 15 days of the Screening Visit will be used to calculate the SES-CD score at Baseline. Subject will be randomly assigned at Baseline to have endoscopy performed at Week 12 or Week 16.

e. Biopsy may be done when performing the endoscopy. Biopsy to confirm CD diagnosis may be done when performing the endoscopy if appropriate documentation for confirmation of the diagnosis does not exist. Biopsies to rule out dysplasia and colon cancer may be taken at the investigator's discretion.

f. Blood pressure, pulse rate, temperature, respiratory rate and weight should be performed before blood draws are performed. Height will be measured at Screening only (with shoes off and then adding 1 inch or 2.5 cm).

g. Subjects with normal ECG within 90 days of Screening would not require a repeat ECG, if documentation is available. Subjects can have a repeat ECG at any time during the study as warranted based on the opinion of the Investigator.

Table 1. Study Activities (Continued)

h. Chest x-ray includes posterior-anterior (PA) and lateral views. Obtain chest x-ray for subjects with TB risk factors as identified by the TB risk factor questionnaire or for subjects living in areas endemic for TB or for subjects with positive PPD or QuantiFERON-TB Gold. Subjects can have a chest x-ray anytime during the study as warranted based on the opinion of the Investigator.

i. PPD skin test is to be read 48 to 72 hours after placement.

j. Dipstick urinalysis will be completed by the sites at all required visits. A microscopic analysis will be performed by the central laboratory, in the event the dipstick results show protein, ketones or blood greater than negative or glucose greater than normal.

k. Serum pregnancy test will be performed on all women of childbearing potential at Screening and at Week 52/ET. Urine pregnancy test will be performed locally at every visit for all women of childbearing potential. If any urine pregnancy test is positive, a serum pregnancy test will be performed by the central laboratory.

If required by country regulatory authorities, monthly pregnancy tests will be performed throughout the study.

l. Subjects will be tested for the presence of the hepatitis B Virus (HBV) and hepatitis C Virus (HCV) at Screening. A positive result for the hepatitis B surface antigen (HBs Ag) or hepatitis C (HCV RNA detectable in any subject with anti-HCV Ab) will be exclusionary. For subjects who are negative for HBs Ag but are positive for core antibodies (HBc Ab), HBV DNA PCR will be performed and any result that meets or exceeds detection sensitivity will be exclusionary.

m. For all subjects, a PK sample will be collected at Week 2, within 1 to 3 hours after dosing, if possible. At Week 4, Week 8, Week 12, Week 16, Week 20, Week 28, Week 36, Week 44 and Week 52/ET, a blood sample for PK will be collected at any time during the visit, and preferably within 1 to 8 hours after the last dose. Patients can take the dose on visit days at their regular schedule and not necessarily at the clinic. The date and accurate time of the PK sample collection and the last ABT-494 dose will be recorded in the CRF to the nearest minute.

n. Stool sample will be collected at each time point indicated. For the visit that endoscopy will be conducted, stool sample should be collected prior to endoscopy.

o. A stool sample will be collected for fecal calprotectin analysis at each time point indicated. Subjects will be asked to provide a stool sample at the visit, if possible, or subjects will be sent home with instructions and stool sample supplies (supplies will be provided).

p. Subject is mandated to begin corticosteroid taper at Week 2.

q. Collection of SAEs begins the day the subject signs the informed consent.

r. Only if subject provides written consent to collect the pharmacogenetic sample; if the informed consent form is not signed, no pharmacogenetic sample can be collected. The sample is preferred to be collected at BL.

Table 1. Study Activities (Continued)

s. Only if subject provides written consent to collect the serum biomarkers and mRNA samples; if the informed consent form is not signed, no serum biomarkers and mRNA can be collected.

t. Only if subject provides written consent to collect biopsy samples for gene expression analysis; if the informed consent form is not signed, no biopsy for gene expression can be collected. Biopsies will be obtained at baseline and as the Week 12/16 endoscopy, based on the colonoscopy assignment for Week 12 or 16.

u. Optional tests, only if determined by the investigator and based on the medical assessment.

v. Laboratory tests performed during the Screening period can be repeated in case the abnormalities are considered to be transient by the investigator.

4.3 Sample Size

Approximately 210 subjects will be equally allocated to five treatment groups and the placebo group, representing a randomization ratio of 1:1:1:1:1:1. The sample size for this study is based on the expected proportion of subjects who achieve endoscopic remission at Week 12/16 and on the expected proportion of subjects who achieve clinical remission at Week 16. Assuming an endoscopic remission (as well as clinical remission) rate of 12% in the placebo arm and maximum of 35% in at least one of the ABT-494 BID treatment arms (3 mg BID, 6 mg BID, 12 mg BID and 24 mg BID) at Week 12/16, a sample size of 35 subjects per treatment group is sufficient to test for the presence of a dose response signal, to select the best dose response model for the observed data out of a prespecified set of candidate models, and to estimate target doses of interest (e.g., the minimum effective dose, MED) via modeling using MCP-Mod (Multiple comparison procedure and modeling) approach. This approach provides at least 80% average power to detect a dose effect at 5% level of significance (one-sided) with the linear, E_{max}, exponential, logistic, and sigEmax models pre-specified as likely candidates to characterize the dose-response for ABT-494 for the two co-primary endpoints of clinical and endoscopic remission.

5.0 Analysis Populations

5.1 Definition for Analysis Populations

The following populations will be used for analyses in the study:

The modified intent-to-treat (MITT) analysis set includes all randomized subjects who have taken at least one dose of study drug in the double-blind induction period. For the MITT analysis set, subjects are assigned to a treatment group based on the randomization schedule, regardless of the treatment actually received. The MITT population will also be referred as "All Randomized and Treated Subjects." The MITT analysis set will be used for all efficacy analysis for the double-blind induction period and baseline demographic and characteristics analyses.

The safety analysis set consists of all subjects who received at least one dose of study medication. For the safety analysis set, subjects are assigned to a treatment group based on the treatment actually received, regardless the treatment randomized.

5.2 Variables Used for Stratification of Randomization

At Baseline, subjects will be randomized in a 1:1:1:1:1:1 ratio to receive one of the six treatment groups (double-blind ABT-494 induction doses or matching placebo).

The following are the treatment groups:

Group 1: ABT-494 3 mg BID

Group 2: ABT-494 6 mg BID

Group 3: ABT-494 12 mg BID

Group 4: ABT-494 24 mg BID

Group 5: ABT-494 24 mg QD

Group 6: Placebo

The randomization at Baseline will be stratified by endoscopic disease severity (SES-CD < 15 and ≥ 15), prior anti-TNF use (naïve and experienced), and by participation in the substudy of the gene expression in intestinal biopsies (Yes and No). As part of the randomization at Baseline, subjects will be randomly assigned (1:1) to have their follow-up ileocolonoscopy done at Week 12 or Week 16.

At Week 16, subjects who have completed the induction period will be re-randomized in a ratio of 1:1:1 to one of the three double-blinded doses of ABT-494 3 mg BID, 6 mg BID or 12 mg BID. The re-randomization will be stratified by dose received during the first 16 weeks, and overall response (responder versus non-responder) at Week 16. The

subjects who were re-randomized at Week 16 to ABT-494 24 mg QD prior to Protocol Amendment 2 will continue to receive the same double-blind dose until Week 52/PD.

Overall response consists of both Endoscopic response and Clinical response. The central reader endoscopic score will be used for calculating the Endoscopic response for the evaluation of the efficacy endpoints. However, for stratification re-randomization, the endoscopic score at baseline from central reader and the endoscopic score at Week 12 or Week 16 from site local reader will be used in order to determine response status.

6.0 Analysis Conventions

Definition of Baseline for Induction Period

The Baseline visit date for Induction Period is the date when the first dose of study drug is received and referred to as Day 1 or Week 0. The Baseline value for a variable is defined as the last non-missing value on or before the date of the first dose of study drug.

Definition of Rx Days (Days Relative to the First Dose of Study Drug)

Rx Days are calculated for each time point of interest and it provides a quantitative measure of days between the event and the first dose date. That is, the Rx Day is calculated as the event date minus the date of first dose of study drug plus 1. The Rx Day will be a negative value when the time point of interest is prior to the date of first dose of study drug, and the Rx Day will be a positive value when the time point of interest is after the first dose date. By this calculation algorithm the first dose day is Rx Day 1, while the day prior to the date of first dose is defined as Rx Day –1 (there is no Rx Day 0).
Rx Days are used to map actual study visits to the protocol specified study visits.

Definition of Analysis Windows

Since subjects do not always adhere to the study visit schedule, the following rules will be applied to assign actual visits to protocol-specified visits including early termination visits. For each study visit mentioned in the protocol, a nominal or target day will be selected to represent the corresponding visit along with a window around the target day.

Windows will be selected in a non-overlapping fashion so that a date collected on the CRF does not correspond to multiple visit windows. Moreover, windows will not discard any Post-Baseline measurement recorded on the CRF. If a subject had two or more actual visits in one visit window, the visit closest to the schedule visit will be used as the study visit for that window. If two visits are equidistant from the target, then the later visit will be used for reporting. If more than one assessment is collected on the same day, then the average of those assessments will be used in analyses.

Table 2. Visit Windows for Analysis of Efficacy Variables (Except SES-CD, IBDQ, WPAI, Abdominal Pain Rating Scale), Laboratory Parameters and Vital Signs for 16-Week Double-Blind Induction Period

Scheduled Week	Nominal Day	Time Window (Rx Day Range)
Week 0	1[a]	≤ 1
Week 2	15	2 – 22
Week 4	29	23 – 43
Week 8	57	44 – 71
Week 12	85	72 – 99
Week 16	113	100 – 999[b]

Rx Day A = date of visit – date of first study drug in double-blind induction period + 1.

a. Day of first dose of study drug in double-blind induction period.

b. Rx Day 999 or first dose of extension period, whichever earlier.

Table 3. Visit Windows for Analysis of IBDQ, EQ5D, and WPAI for 16-Week Double-Blind Induction Period

Scheduled Week	Nominal Day	Time Window (Rx Day Range)
Week 0	1[a]	≤ 1
Week 8	57	2 – 85
Week 16	113	86 – 999[b]

Rx Day A = date of visit – date of first study drug in double-blind induction period + 1.

a. Day of first dose of study drug in double-blind induction period.

b. Rx Day 999 or first dose of extension period, whichever earlier.

Table 4. Visit Windows for Analysis of SES-CD and Abdominal Pain Rating Scale for 16-Week Double-Blind Induction Period

Scheduled Week	Nominal Day	Time Window (Rx Day Range)
Week 0	1[a]	≤ 1
Week 12	85	2 – 99
Week 16	113	100 – 999[b]

Rx Day A = date of visit – date of first study drug in double-blind induction period + 1.

a. Day of first dose of study drug in double-blind induction period.

b. Rx Day 999 or first dose of extension period, whichever earlier.

Table 5. Visit Windows for Analysis of Fecal Calprotectin for 16-Week Double-Blind Induction Period

Scheduled Week	Nominal Day	Time Window (Rx Day Range)
Week 0	1[a]	≤ 1
Week 4	29	2 – 71
Week 16	113	72 – 999[b]

Rx Day A = date of visit – date of first study drug in double-blind induction period + 1.

a. Day of first dose of study drug in double-blind induction period.

b. Rx Day 999 or first dose of extension period, whichever earlier.

CD-Related Corticosteroid Use

Subjects in whom the CD-related corticosteroids (systemic or rectal corticosteroids) that were not being taken at Baseline and are initiated during the study or who have equivalent dose of these medications increased to greater than the dose taken at Baseline will be censored for efficacy assessments (i.e., will be considered non-responders for categorical endpoints and will have Baseline values carried forward for non-categorical assessments) from that point through the end of the study. These subjects will continue to be evaluated in the safety population.

The equivalent steroid dose will be determined based on the table below:

Corticosteroid	Equivalent Dose (mg)
Cortisone	25
Hydrocortisone	20
Prednisone	5
Prednisolone	5
Triamcinolone	4.0
Methylprednisolone	4.0
Betamethasone	0.75
Dexamethasone	0.75
Budesonide	1
Beclomethasone	5
Beclometasone	5

Definition of Missing Data Imputation

The following imputation methods will be used to impute missing values in the efficacy analyses. In addition, an observed case analysis will be performed.

Non-Responder Imputation (NRI)

The NRI approach is used for binary efficacy variables. These variables can take values of 'Achieved' or 'Not Achieved' or may be missing for any reason including discontinuation from study. According to the NRI imputation approach, all missing values will be considered as 'Not Achieved.'

Last Observation Carried Forward (LOCF)

For all variables (categorical variables and continuous variables), the following rules will be used for the LOCF approach:

1. Baseline and Pre-Baseline values will not be used to impute the missing Post-Baseline values.

2. Missing values after Study Day 1 will be imputed using the latest non-missing values after Day 1 and prior to the missing value. If there are no non-missing values after Baseline, then the LOCF value will be missing.

Observed Case (OC)

Observed case analysis will be performed such that missing values will not be imputed.

However, all observed values will be replaced by Baseline value once subjects received CD-related corticosteroid therapy.

Mixed-Effect Model Repeated Measure (MMRM)

The MMRM model will be used for continuous efficacy variables with longitudinal data as the primary analysis. The MMRM model includes the baseline values as covariate; randomization strata, treatment, time point and treatment-by-time point interaction as fixed effects; and subjects within treatment as random effect. An unstructured (co)variance structure will be used to model the within-subject error. The comparison at a time point will be the contrast between treatments at that time point. Satterthwaite's approximation will be used to estimate denominator degrees of freedom.

Imputation of Missing Dates

For Baseline, efficacy, and safety parameters, if the day and/or month are missing, the following conventions will be used to impute the missing dates:

- 01 for missing start day
- End of month for missing end day
- January 1st for missing start month
- December 31st for missing end month

In case of partially missing AE start and stop dates, the dates will be imputed by comparing to first dose date of study medication so than the corresponding AEs will be

made treatment-emergent whenever possible. If the start date of an AE is partially missing and the month is the same as the start date of a new therapy, the AE will be made treatment emergent to the new therapy.

In case of missing or partially missing study drug dosing dates, the dates will not be imputed. Subjects will be treated as not receiving dose on that date.

Rule for CDAI Calculation

Up to 14 days of diary entries will be evaluated from the ePRO tool for the CDAI calculation for each visit. The diary entries on the days the subjects receive endoscopy preparation medications, the day of endoscopy procedure, and 2 days after endoscopy procedure will be excluded. For each CDAI subscore, the available scores from the most recent diary days (at least 4 days, up to 7 days) prior to actual day of the study visit will be summed, and then multiplied by the corresponding multiplier to get subtotal score. If available diary entries are fewer than 7 days, the subtotal score will be calculated as (summed total available score/number of days) × 7 × corresponding multiplier. The three subtotal scores that are based on ePRO (number of liquid/very soft stools, abdominal pain rating, and general well-being) will then be rounded to one decimal. The final CDAI is rounded to a whole number.

If a subject has less than 4 days of diary data, the total CDAI score will not be calculated and will be considered missing.

Same rule will be applied to average daily stool frequency and average daily abdominal pain.

Bristol Stool Chart

Up to 14 days of diary entries will be evaluated from the ePRO tool for the Bristol Stool Chart calculation. These assessments on the day of endoscopy procedure, the day before endoscopy procedure (due to preparation medications), and 2 days after endoscopy procedure are excluded. The 7 most recent non-missing assessments will be used for the

analysis. The Bristol Stool Chart is calculated as the number of days the subjects with Type 6 or Type 7 divided by the total number days with non-missing assessments, rounded to two decimal.

If a subject has less than 4 days of diary data, the Bristol Stool Chart will not be calculated and will be considered missing.

If a subject achieved reduction from Baseline ≥ 50%, this subject will be classified as Responder.

SES-CD Scoring

All colonoscopies shall be performed and recorded in video format. To be eligible for the study, the screening colonoscopy may be performed during the Screening Period (or up to 45 days before the Baseline visit if Amendment 2 is in place at the site). Colonoscopies will also be performed at the Week 12 or 16 visit. Investigators shall provide their SES-CD score assessment on the SES-CD score sheet for the baseline and Week 12 or 16 visits. The investigator SES-CD scores are to be entered in the appropriate eCRF, however neither the Investigator's Baseline nor Week 12 or 16 SES-CD scores will be used for the study's efficacy analyses.

All videotaped colonoscopies will undergo central review. One primary central reviewer will evaluate the videotaped colonoscopies and provide the SES-CD scores to Parexel. A second central reviewer will adjudicate between the investigator's and the central reviewer's SES-CD scores if there is discrepancy (more than 10% difference) in total SES-CD score. The adjudicator will select the final SES-CD score that he/she most agrees with from those provided by the investigator and the central reviewer, and this final SES-CD score will be entered into the study database and be used for the study's efficacy analyses. If there is no discrepancy between the investigator and the central reviewer, that central reviewer's score will be entered into the study database and serve as the SES-CD score to be used for the study's efficacy analyses.

If there is a missing SES-CD individual variable in the SES-CD score that serves as the SES-CD score for the study's efficacy analyses (for all variables except the "Presence of Narrowing" variable), the following imputation rules will be applied:

If the same individual variable is:

- Missing at both Baseline and Week 12 or 16: the missing value will be imputed as zero at both time points, and no change will be assumed.
- Present at Week 12 or 16 but missing at Baseline: the missing Baseline value will be imputed based on the value given at Week 12 or 16, and no change will be assumed.
- Present at Baseline but missing at Week 12 or 16: the missing Week 12 or 16 value will be imputed based on the value given it at Baseline and no change will be assumed.

If there is a missing SES-CD "Presence of Narrowing" individual variable in the SES-CD score that serves as the SES-CD score for the study's efficacy analyses, the following imputation rules will be applied:

If the Presence of Narrowing individual variable is:

- Missing at both Baseline and Week 12 or 16: the missing Presence of Narrowing individual variable will be imputed as zero and no change will be assumed.
- Present at Week 12 or 16 but missing at Baseline: the missing Baseline value will be imputed based on the individual variable given at Week 12 or 16 and no change will be assumed.

  However, if there is a Baseline individual variable (for any SES-CD subscore) for intestinal segments proximal to (above) the segment with the missing Presence of Narrowing individual variable, the maximum imputed value for the missing Presence of Narrowing individual variable will be 2.

- Present at Baseline but missing at Week 12 or 16: the missing Week 12 or 16 value will be imputed based on score given at Baseline and no change will be assumed.

  However, if there is a Week 12 or 16 individual variable is present for intestinal segments proximal to (above) the segment with the missing Presence of Narrowing variable, the maximum imputed value for the missing Presence of Narrowing individual variable will be 2.

(Gastrointestinal anatomy places the ileum proximal to the right colon, which is proximal to the transverse colon, which is proximal to the sigmoid and left colon, which is proximal to the rectum.)

A sensitivity analysis (Worst Case Imputation) will be performed by imputing all missing Baseline individual variables as '0' and all missing Week 12 or 16 variables as '3' except the Week 12 or 16 Presence of Narrowing individual variable. At Week 12 or 16, the worst values for Presence of Narrow individual variable will be 2 for Rectum, Sigmoid and Left Colon, Transverse Colon, and Right Colon, and 3 for Ileum.

If a subject only has Baseline SES-CD, and no post-Baseline SES-CD, all Baseline missing individual variables will be imputed as zero.

Isolated Ileal Crohn's Disease at Baseline

Isolated Ileal Crohn's disease at Baseline is defined as all individual variables for Rectum, Sigmoid and Left Colon, Transverse Colon, and Right Colon are zero at Baseline and subscore for ileal ≥ 4. If any of the individual variables is missing or not zero, this subject is classified as Not Isolated Ileal Crohn's disease at Baseline.

Definition of Endoscopic Remission

SES-CD ≤ 4 and at least two point reduction versus Induction baseline and no subscore > 1 in any individual variable.

Definition of Endoscopic Response

SES-CD at least 25% reduction from Induction baseline.

Definition of Clinical Remission

Average daily stool frequency ≤ 1.5 and not worse than baseline AND average daily abdominal pain ≤ 1.0 and not worse than Induction baseline.

Definition of Clinical Response

Average daily stool frequency at least 30% reduction from Induction baseline and average daily abdominal pain not worse than Induction baseline OR average daily abdominal pain at least 30% reduction from Induction baseline and average daily stool frequency not worse than Induction baseline.

Definition of Remission

Remission at Week 16 is defined as both Endoscopic remission at Week 12/16 AND Clinical remission at Week 16.

Definition of Response

Response at Week 16 is defined as both Endoscopic response at Week 12/16 AND Clinical response at Week 16.

Definition of Modified Clinical Remission

Average daily stool frequency ≤ 2.8 and not worse than baseline AND average daily abdominal pain ≤ 1.0 and not worse than Induction baseline. This endpoint will only be analyzed among subjects with baseline average daily stool frequency ≥ 4 or average daily abdominal pain ≥ 2.0.

EuroQol-5D-5L (EQ-5D)

EQ-5D is a standardized measure of health status developed by the EuroQol Group in order to provide a simple, generic measure of health for clinical and economic appraisal.

The EQ-5D consists of 2 pages. The first page measures 5 dimensions of the health status (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) with 5 levels per dimension (no problems, slight problems, moderate problems, severe problems, and unable corresponding to Level 1 to Level 5 respectively). The second page is an EQ Visual Analogue Scale (EQ VAS). EQ-5D health states, defined by the EQ-5D-5L descriptive system on the first page, may be converted into a single index value. The change from baseline of the index value and EQ VAS will be analyzed and reported. UK scoring algorithm will be used (Table 6).

Table 6. An EQ-5D-5L Value Set for England

	Central Estimate	Value for Health State 23245
Mobility		
No Problem	0	
Slight	0.051	0.051
Moderate	0.063	
Severe	0.212	
Unable	0.275	
Self-Care		
No Problem	0	
Slight	0.057	
Moderate	0.076	0.076
Severe	0.181	
Unable	0.217	
Usual Activities		
No Problem	0	
Slight	0.051	0.051
Moderate	0.067	
Severe	0.174	
Unable	0.190	
Pain/Discomfort		
No	0	
Slight	0.060	
Moderate	0.075	
Severe	0.276	0.276
Unable	0.341	
Anxiety/Depression		
Not	0	
Slight	0.079	
Moderate	0.104	
Severe	0.296	
Unable	0.301	0.301
The value for health state 23245: 1 – 0.9675 × (sum of the subscores)		**1 – 0.9675** × (0.051 + 0.076 + 0.051 + 0.276 + 0.301) = 0.270

The index value is 1 – 0.9675 × sum of 5 components based on central estimates.

The minimum index value is – 0.281 (health state 55555), and the maximum index value is 1 (health state 11111). (Reference: Devlin N et al. Valuing health-Related Quality of Life: Ann EQ-5D-5L Value Set for England, January 2016 – https://www.ohe.org/publications/valuing-health-related-quality-life-eq-5d-5l-value-set-england or http://www.euroqol.org/about-eq-5d/valuation-of-eq-5d/eq-5d-5l-value-sets.html).

If one of the 5 dimensions is missing, the EQ-5D index value will be missing.

7.0 Demographics, Baseline Characteristics, Medical History, and Previous/Concomitant Medications

7.1 Demographic and Baseline Characteristics

For the subjects in MITT and safety analysis set, demographic information and Baseline values will be summarized by descriptive statistics. Categorical data will be summarized by number and percent; and quantitative data will be presented by n, mean, standard deviation, minimum value, median, and maximum value.

In general, continuous variables will be analyzed using analysis of variance (using SAS procedure 'PROC GLM') with treatment group as factor. Categorical variable will be analyzed using chi-square test or Fisher's exact test if ≥ 20% of the cells have expected cell count < 5.

The following demographic and Baseline values will be summarized.

Continuous Variables:

- Age (years)
- Body weight (kg)
- Height (cm)
- Body Mass Index (kg/m^2)
- Blood Pressure (systolic/diastolic) (mmHg)
- Pulse (bpm)

- Temperature (°C)
- CDAI
- Average Daily Stool Frequency
- Average Daily Abdominal Pain
- Total SES-CD score
- Crohn's Disease Duration (years)
- hs-CRP mg/L
- Fecal calprotectin
- IBDQ score
- WPAI and its components
- EQ-5D
- Abdominal Pain Rating Scale
- Bristol Stool Chart Score

Categorical Variables:

- Sex (male, female)
- Race
- Ethnicity
- Age (< 40, 40 to < 65, 65 to < 75, ≥ 75)
- Baseline corticosteroid use (yes, no)
- Baseline immunosuppressant use (yes, no)
- hs-CRP at Baseline (< 10 and ≥ 10 mg/L)
- Baseline fecal calprotectin (≤ 250 µg/g, > 250 µg/g)
- Crohn's disease severity (CDAI ≤ 300, > 300) at Baseline
- Baseline SES-CD (< 15, ≥ 15)
- Disease duration (≤ 3 years, > 3 years)
- Tobacco use (user, ex-user, never used, unknown)
- Alcohol use (drinker, ex-drinker, non-drinker, unknown)
- Region (US, ex-US)

7.2 Medical History

Medical and Surgical History: A complete medical and surgical history (which includes CD-onset date), history of tobacco and alcohol use, and TB history will be obtained from each subject during the Screening period. Medical history will be summarized using body system and condition/diagnosis by treatment group. No statistical tests will be performed.

Chest X-Ray Results: All subjects undergo a standard X-ray of chest (including a posteroanterior [PA] and lateral view) at Screening period. Number and percent of subjects with presence or absence of finding for the previous TB infection, calcified granulomas, Pleural scarring/thickening, and other findings will be presented by treatment group. No statistical tests will be performed.

TB Test Results: Results of PPD skin test, QuantiFERON-TB Gold test at screening visit will be summarized. Induration will be summarized descriptively using n, mean, standard deviation, minimum values, median, and maximum values. The frequency distribution of induration ≥ 5 and < 5 will be provided. QuantiFERON-TB tests will be described as positive or negative. Indetermiante QuantiFERON-TB test results will be repeated. If the second QuantiFERON-TB test is positive or indeterminate, the final assessment will be considered positive. If the second QuantiFERON-TB test is negative, the final assessment will be considered negative. No statistical tests will be performed.

TB Prophylaxis: History of use of TB Prophylaxis or initiation of TB prophylaxis will be summarized.

ECG Results: ECG results at screening will be presented as frequency distribution showing results as Normal, Abnormal (Not clinically significant), Abnormal (Clinically significant) and Unable to evaluate/missing. No statistical tests will be performed.

7.3 Previous Treatment and Concomitant Medications

Based on generic medication names, these categories of medications used by subjects before and during the study will be summarized by number and percent for ITT and safety analysis sets for the treatment groups. No statistical tests will be performed.

The number and percent of subjects using Crohn's disease specific medications (including corticosteroids, aminosalicylates, immunosuppressants [defined as azathioprine, 6-mercaptopurine, or methotrexate], antibiotics) within past 90 days prior to the Baseline, and at the Baseline will be tabulated. In addition, the number and percent of subjects using Crohn's disease specific immunosuppressants, and biologic therapies at any time prior to Baseline will be tabulated.

8.0 Patient Disposition

Subject disposition will be presented for subjects in the MITT and safety analysis sets using the following information by treatment group:

- Number and percent of subjects in various analysis sets by treatment group and by investigator and/or site number
- Number and percent of subjects completing double-blind induction period and discontinuing on or before Week 16 visit of induction period
- Subject disposition including the number and percent of subjects who prematurely discontinued the double-blind induction period (on or before Week 16 of induction period) by primary reason and by any reason

Summary of protocol deviations will be provided.

9.0 Study Drug Exposure and Compliance

Study drug exposure and compliance will be summarized using the mean, standard deviation, minimum, median, and maximum. Exposure to study drug will be summarized by treatment group.

Compliance is defined as the number of capsules taken (i.e., the difference between the number of capsules dispensed and the number of capsules returned) divided by the number of capsules a subject is supposed to take each day times the length of time that the subject was in the Treatment Phase of the study (i.e., Final/Discontinuation Visit date during Treatment Phase – Day 1 [Baseline] Visit date +1). Subjects with missing data for the number of capsules returned will be excluded from the summary.

10.0 Efficacy Analysis

10.1 General Considerations

All statistical tests will be two-sided with the significance level of 0.100. Descriptive statistics will be provided. These include the number of observations, mean, standard deviation, minimum, median, and maximum for continuous variables; and number and percent for discrete variables. The analysis will be performed using SAS® (SAS Institute Inc., Cary, NC, USA).

Dose Response Modeling

Other than estimating the relative treatment effect of the investigational compound ABT-494 to placebo, one important goal of this Phase 2 study is to establish dose-response relationship to facilitate the dose selection for future Phase 3 trials. Multiple Comparison Procedure and dose-response Modeling (MCPMod) with a pre-defined group of candidates dose response curves will be tested against flat dose-response curve to best characterize the dose-response relationship.

Steps of MCPMod:

1. Choose a candidate set of ***S*** models.
2. Compute the optimum contrast for each model.
3. Use contrast test to find the significant ***T*** models while preserving FWER (one-sided significant level of 0.05).

4. Use AIC criteria to find the most significant model from the significant ***T*** models found from Step 3.
5. Use the model found from Step 4 to fit observed data from the study and make inference (e.g., to find Minimum Effective Dose (MED) or the dose achieving certain amount of maximum effect), or use all significant models to make inference about the weighted target dose of interest.

ADDPLAN or R will be used to evaluate different dose-response models and to make dose recommendation.

10.2 Primary Efficacy Analyses

This section provides the details of the primary efficacy analysis for the study.

Co-Primary Efficacy Variables:

- Proportion of subjects who achieve endoscopic remission at Week 12/16.
- Proportion of subjects who achieve clinical remission at Week 16.

Analysis Data Set for the Primary Efficacy Analysis:

The primary efficacy analysis will use the MITT analysis data set.

Imputation Method Used for the Primary Efficacy Analysis:

Missing average daily stool frequency or average daily abdominal pain at Week 16 or SES-CD at Week 12/16 will be imputed using the non-responder imputation (NRI) approach.

Statistical Method of the Primary Efficacy Analysis:

The dose-response relationships among the five ABT-494 treatment groups and placebo group will be characterized for the primary endpoints endoscopic remission at Week 12/16 and clinical remission at Week 16 using MCPMod approach. The following

models will be considered: linear, E_{max}, exponential, logistic, sigEmax, and quadratic. The MCPMod approach for trial analysis stage consists of two main steps: MCP and Mod step. The MCP step focuses on establishing evidence for a drug effect across the doses, i.e., detecting a statistically significant dose response signal for the clinical endpoint and patient population investigated in the study. This step will typically be performed using an efficient test for trend, adjusting for the fact that multiple candidate dose response models are being considered. If a statistically significant dose response signal has been established, one proceeds with determining a reference set of significant dose response models by discarding the non-significant models from the initial candidate set.

The response function will be the log odds (logit) of the proportion of subjects with endoscopic/clinical remission. The fitted curve will be shown graphically with confidence intervals for each dose. Estimates of the treatment differences in the response function and associated 95% confidences for each active dose against placebo will be calculated from the model. These results will be back-transformed to give point estimates of the difference in proportions and associated 95% confidence intervals.

Sensitivity Analysis of the Primary Efficacy Variables

The following sensitivity analyses for the primary endpoint of endoscopic remission will be conducted:

- A sensitivity analysis using observed cases, which excludes those subjects with missing post-baseline endoscopy, will also be done.

The following sensitivity analyses for the primary endpoint of clinical remission will be conducted:

- An analysis of observed cases, which excludes those subjects with missing SF or AP data at scheduled assessment visits.
- The primary analysis will be repeated using mixed-imputation. Subjects who discontinue prior to Week 16 due to lack of efficacy or adverse events will be

considered as "not achieved" for the clinical remission. Subjects who discontinue for other reasons will be categorized according to LOCF.

10.3 Secondary Efficacy Analyses

Secondary efficacy variables are divided into two groups. The first group includes double-blind induction treatment period secondary endpoints. Analyses for the first group of secondary endpoints will be performed using MITT analysis set. The second group includes double-blind extension phase secondary endpoints. Analyses for the second group of secondary endpoints will be performed using ITT-N and ITT-NR analysis sets. This SAP is for Induction period only.

The secondary endpoints *(Double-Blind Induction Treatment Period)* include:

- Proportion of subjects who achieve CDAI < 150 at Week 16.
- Proportion of subjects with decrease in CDAI ≥ 70 points from Baseline at Week 16.
- Proportion of subjects with decrease in CDAI ≥ 100 points from Baseline at Week 16.
- Proportion of subjects who achieve clinical remission at Week 12.
- Proportion of subjects who achieve remission at Week 16 (endoscopic remission at Week 12/16 and clinical remission at Week 16).
- Proportion of subjects who achieve response at Week 16 (endoscopic response at Week 12/16 and clinical response at Week 16).
- Proportion of subjects with endoscopic response at Week 12/16.
- Proportion of subjects who achieve clinical response at Week 16.
- Proportion of subjects with an average daily SF ≥ 2.5 AND average daily AP ≥ 2.0 at Baseline who achieve clinical remission at Week 16.
- Proportion of subjects who steroid-free achieve CDAI < 150 at Week 16 among subjects taking corticosteroids at Baseline.
- Proportion of subjects who achieve steroid-free remission at Week 16 among subjects taking corticosteroids at Baseline.

- Proportion of subjects who achieve steroid-free clinical remission at Week 16 among subjects taking corticosteroids at Baseline.
- Proportion of subjects who achieve steroid-free endoscopic remission at Week 12/16 among subjects taking corticosteroids at Baseline.
- Change from Baseline in fecal calprotectin level at Week 16.
- Change from Baseline in fecal calprotectin level at Week 4.
- Change from Baseline in hs-CRP at Week 16.
- Change from Baseline in IBDQ at Week 16.
- Change from Baseline in IBDQ at Week 8.
- Proportion of subjects who achieve remission at Week 16 among subjects with isolated ileal Crohn's disease at Baseline.
- Change from Baseline in abdominal pain rating scale at Week 12 and Week 16.
- The proportion of subjects who achieve > 50% reduction from Baseline in SES-CD or endoscopic remission at Week 12/16.
- The proportion of subjects who achieve modified clinical remission at Week 16 among subjects with an average daily SF $\geq$ 4.0 or average daily AP $\geq$ 2.0 at Baseline.

For categorical efficacy endpoints for Induction Period, the pairwise comparisons for the difference in proportions of subjects between treatment groups and placebo group will be analyzed using the CMH test adjusted for Crohn's disease severity (SES-CD < 15, $\geq$ 15) at Baseline. Additionally, the CMH-based 95% confidence interval for the difference in proportions will be provided. The 95% confidence interval for the difference in proportions within each Crohn's disease severity (SES-CD < 15, $\geq$ 15) at Baseline category will be calculated based on normal approximation.

The non-responder imputation will be used for subjects with missing data at the endpoint evaluated. The last observation carried forward (LOCF) method will also be used as the sensitivity analyses for the secondary endpoints.

In general, continuous secondary efficacy variables with repeated measurements will be analyzed using a Mixed Effect Repeated Measure (MMRM) model. The mixed model includes the categorical fixed effects of treatment, Crohn's disease severity (SES-CD < 15, ≥ 15) at Baseline, week and treatment-by-week interaction and the continuous fixed covariates of baseline measurement. An unstructured variance-covariance matrix will be used. If convergence is not achieved, other variance-covariance matrix will be explored until coverage. The parameter estimations will be based on assumption of data being missing at random and using the method of restrictive maximum likelihood (REML).

Continuous secondary efficacy variables will also be analyzed using an Analysis of Covariance (ANCOVA) model including factors for treatment group and Crohn's disease severity (SES-CD < 15, ≥ 15) at Baseline, and Baseline values as a covariate. The MMRM analysis is considered primary for inferential purposes.

The additional endpoints include:

- Proportion of subjects who achieve clinical remission over time in Induction Period.
- Proportion of subjects who achieve clinical response over time in Induction Period.
- Proportion of subjects who achieve CDAI < 150 over time in Induction Period.
- Proportion of subjects with decrease in CDAI ≥ 70 points from Induction Baseline over time in Induction Period.
- Proportion of subjects with decrease in CDAI ≥ 100 points from Induction Baseline over time in Induction Period.
- Proportion of subjects who are IBDQ responders (increase in IBDQ score ≥ 16 points from Induction Baseline) over time in Induction Period.
- Proportion of subjects with IBDQ score ≥ 170 over time in Induction Period.
- Proportion of subjects who achieve steroid-free CDAI < 150 over time in Induction Period among subjects taking corticosteroids at Baseline.
- Proportion of subjects who achieve steroid-free clinical remission over time in Induction Period among subjects taking corticosteroids at Baseline.

- Change from Induction Baseline in fecal calprotectin level over time in Induction Period.
- Change from Induction Baseline in hs-CRP over time in Induction Period.
- SES-CD endpoints without any missing individual variables.
- The proportion of subjects with no anal fissures at Week 16 among subjects with anal fissures at Baseline.
- The proportion of subjects with no draining fistulas at Week 16 among subjects with draining fistula at Baseline.
- The proportion of subjects with no non-draining fistulas at Week 16 among subjects with non-draining fistula at Baseline.
- The proportion of subjects with no any fistulas at Week 16 among subjects with any fistula at Baseline.
- Change from Induction Baseline in Bristol Stool Chart at Week 4 and Week 16.
- The proportion of subjects who achieve Bristol Stool Chart response at Week 4 and Week 16.
- Cross tabulation of Baseline and Week 16 in EIM.
- The proportion of subjects who achieve modified clinical remission at over time in Induction Period among subjects with an average daily SF ≥ 4.0 or average daily AP ≥ 2.0 at Baseline.
- Change from Induction Baseline in EQ-5D over time in Induction Period.
- Change from Induction Baseline in WPAI over time in Induction Period.

10.4 Handling of Multiplicity

The MCP step in MCPMod approach will typically be performed using an efficient test for trend, adjusting for the fact that multiple candidate dose response models are being considered.

10.5 Efficacy Subgroup Analysis

The subgroups listed below will be used in subgroup analyses of the co-primary endpoints.

- Sex (male, female)
- Age (≤ median, > median)
- Race (white, non-white)
- Baseline fecal calprotectin (≤ median, > median)
- Baseline fecal calprotectin (≤ 250 µg/g, > 250 µg/g)
- Baseline corticosteroid use (yes, no)
- Baseline immunosuppressant use (yes, no)
- hs-CRP at Baseline (< 10 and ≥ 10 mg/L)
- hs-CRP at Baseline (≤ median, > median)
- Crohn's disease severity (CDAI ≤ 300, > 300) at Baseline
- Baseline CDAI (≤ median, > median)
- Baseline SES-CD (≤ median, > median)
- Baseline SES-CD (< 15, ≥ 15)
- Weight (≤ median, > median)
- Baseline albumin (≤ median, > median)
- Disease duration (≤ 3 years, > 3 years)
- Disease duration (≤ median, > median)
- Region (US, ex-US)
- Prior anti-TNF status (prior anti-TNF, prior anti-TNF naïve)

11.0 Safety Analysis

11.1 General Considerations

Adverse events (AEs), laboratory data and vital signs are the primary safety parameters in this study. All safety comparisons will be performed between treatment groups using the

safety analysis set. The safety variable will be summarized by treatment according to the treatment a subject actually received. The differences between treatment groups and placebo in safety parameters will be evaluated using two--sided tests at the significance level of 0.100.

11.2 Analysis of Adverse Events

11.2.1 Treatment-Emergent Adverse Events

Treatment-emergent AEs for double-blind Induction period are defined as events that begin or worsen either on or after the first dose of the study medication and within 30 days after the last dose of the study medication for who subjects who discontinued study prior to double-blind Extension Period or up to the first dose of double-blind Extension Period for subjects who received at least a dose of study medication during double-blind extension period. An overview of treatment-emergent AEs, including AEs of special interest such as adverse events leading to death and adverse events leading to early termination, AEs by Medical Dictionary for Drug Regulatory Activities (MedDRA version 18.1 or later) preferred term and system organ class, AEs by maximum relationship to study drug, and AEs by maximum severity will be summarized by number and percentage. Treatment group differences (each ABT-494 dose group versus placebo group, as well as ABT-494 dose groups combined versus placebo group) in the overall incidence of treatment-emergent AEs will be assessed with Fisher's exact test for each preferred term for double-blind induction period.

Changes in laboratory data will be described using statistical characteristics and compared between treatment groups and placebo will be performed using a one-way Analysis of Variance (ANOVA). In addition, shift tables and listings will be provided for abnormal values, whereby the normal range of the analyzing laboratory will be used. Vital signs will be analyzed similarly.

The number and percent of subjects experiencing treatment-emergent adverse events will be summarized for the following adverse event categories. Comparisons of the percent of subjects experiencing an adverse event between treatment groups and placebo will be

performed using Fisher's exact tests. Only *P* values < 0.100 when rounded to three digits will be presented.

- Any treatment-emergent adverse event.
- Any treatment-emergent adverse event that was rated as possibly related to study drug by the investigator (Reasonable Possibility).
- Any treatment-emergent severe adverse event.
- Any treatment-emergent serious adverse event.
- Any treatment-emergent adverse event leading to discontinuation of study drug.
- Any treatment-emergent adverse event leading to death.
- Any treatment-emergent adverse event of special interest.

Treatment-emergent adverse events will be summarized as follows:

- Grouped by System Organ Class and Preferred Term.
- A by-subject listing will be provided.
- Grouped by System Organ Class, Preferred Term and Severity.
- Grouped by System Organ Class, Preferred Term and Relationship to Study Drug.
- Grouped by System Organ Class and Preferred Term with subject numbers.

In treatment-emergent AE tables, a subject who reports more than one treatment-emergent AE in different system organ classes will be counted only once in the overall total. A subject who reports two or more different preferred terms which are in the same SOC will be counted only once in the SOC total. A subject who reports more than one treatment AE with the same preferred term will be counted only once for that preferred term using the most extreme incident (i.e., most "severe" for the severity tables and most "related" for the relationship tables).

Adverse events will also be summarized by maximum severity. If a subject has an adverse event with unknown severity, then the subject will be counted in the severity

category of "unknown," even if the subject has another occurrence of the same adverse event with a severity present. The only exception is if the subject has another occurrence of the same adverse event with the most extreme severity – "Severe." In this case, the subject will be counted under the "Severe" category.

Adverse events will also be summarized by maximum relationship to study, as assessed by the investigator. If a subject has an adverse event with unknown relationship, then the subject will be counted in the relationship category of "unknown," even if the subject has another occurrence of the same adverse event with a relationship present. The only exception is if the subject has another occurrence of the same adverse event with a relationship assessment of "Reasonable Possibility." In this case, the subject will be counted under the "Reasonable Possibility" category, respectively.

Incidence rates per 100 patient years of exposure to study drug will be presented for AE overviews and for AEs by SOC and preferred term where the number of events will be used as the numerator.

11.2.2 Adverse Events of Special Interest

The Adverse Events of Special Interests (AESI) categories will be summarized and presented for each treatment group using SMQ/CMQ.

AESI	Type of MedDRA Query	Broad or Narrow Search	SMQ/CMQ Search Criteria
Infection	CMQ		"Infection"
Serious Infection	CMQ		"Infection" – Subset for SAEs
Opportunistic Infection	CMQ		"Opportunistic infection"
Herpes zoster	CMQ		"Herpes zoster"
Tuberculosis	CMQ		"Tuberculosis"
Malignancy	SMQ	Narrow	"Malignancies," "Malignant lymphomas"
Non-Melanoma Skin Cancer (NMSC)	SMQ	Narrow	"Skin neoplasms, malignant and unspecified"
Malignancy other than NMSC			
Hepatic Disorder	SMQ	Narrow	"Drug Related Hepatic disorders"
Gastrointestinal Perforations	SMQ	Narrow	"Gastrointestinal perforation"
Anemia	CMQ		"Non-Hemolytic and Non-Aplastic anemias"
Adjudicated Cardiovascular Events	Output from CAC		
Neutropenia	CMQ		"Hematological toxicity – Neutropenia"
Increased serum creatinine and renal dysfunction	SMQ	Narrow	"Acute renal failure"
Elevated creatine phosphokinase	SMQ	Narrow	"Rhabdomyolysis/myopathy"

Additional AEs may be considered for tabulation/summary based on recommendations from Clinical and Safety as deemed appropriate.

11.3 Analysis of Laboratory Data

Changes from Baseline in continuous laboratory parameters will be summarized by n, mean, standard deviation, minimum value, median, and maximum value for each treatment group. The mean change from Baseline will be compared between treatment groups at each time point using the ANOVA model with treatment as a factor. The differences between treatment groups and placebo for mean changes from Baseline will

be summarized using the mean, standard error, 95% confidence interval, and *P* value for the between-group difference.

Cross (Shift) tables from Baseline to the final value according to the normal range will be provided for each hematology, clinical chemistry parameter and urinalysis parameter except for the microscopic examination.

For selected laboratory parameter with Common Toxicity Criteria (CTC) a listing of all subjects with any laboratory determinations meeting CTC Version 4.03 (or later) of Grade ≥ 3 will be provided. For hemoglobin, an additional list will also be provided based on CTC Version 3.0 (or later) of Grade ≥ 2. For each of these subjects, the whole course of the parameter will be listed. For subjects with laboratory values with CTC ≥ 3, all of the laboratory parameters for those subjects will be listed.

11.3.1 Analysis of Liver Specific Laboratory Tests

The liver specific laboratory tests include the serum glutamic-pyruvic transaminase (ALT/SGPT), serum glutamic-oxaloacetic transaminase (AST/SGOT), alkaline phosphatase and total bilirubin. Each of these laboratory values will be categorized as follows:

1. < 1.5 ULN,
2. ≥ 1.5 × ULN TO < 3 × ULN,
3. ≥ 3 × ULN TO < 5 ULN,
4. ≥ 5 × ULN TO < 8 ULN, and
5. ≥ 8 × ULN,

where ULN is the upper normal limit.

Shift tables showing shift from Baseline to maximum and final values will be presented using these five categories.

A listing of potentially clinically significant liver function laboratory values will be provided. The listing will include all subjects who met any of the following 4 criteria:

1. ALT ≥ 2.5 × ULN, or
2. AST ≥ 2.5 × ULN, or
3. Alkaline Phosphatase ≥ 2.5 × ULN, or
4. Total Bilirubin ≥ 1.5 × ULN.

11.4 Analysis of Vital Signs and Weight

The following vital signs are measured at every visit during the study.

- Body Weight (kg)
- Blood Pressure (Systolic/Diastolic) (mmHg)
- Pulse (bpm)
- Temperature (°C)

Changes from Baseline in vital sign values will be summarized by n, mean, standard deviation, minimum value, median, and maximum value for each treatment group. The mean change from Baseline will be compared between treatment groups at each time point using the ANOVA model with treatment as a factor. Treatment group differences for mean changes from Baseline will be summarized using the mean, standard error, 95% confidence interval, and *P*-value for the between-group difference.

In additional, incidence of potential clinically significant results will be summarized.

Criteria for potentially clinically significant vital signs results:

Vital Sign	Category	Criteria for Potentially Clinically Significant Vital Sign
Systolic blood pressure	Low	Value ≤ 90 mmHg and/or decrease ≥ 20 mmHg from Baseline
	High	Value ≥ 180 mmHg and/or increase ≥ 20 mmHg from Baseline
Diastolic blood pressure	Low	Value ≤ 50 mmHg and/or decrease ≥ 15 mmHg from Baseline
	High	Value ≥ 105 mmHg and/or increase ≥ 15 mmHg from Baseline
Pulse	Low	Value ≤ 50 bpm and/or decrease ≥ 15 bpm from Baseline
	High	Value ≥ 120 bpm and/or increase ≥ 15 bpm from Baseline

12.0 Summary of Changes

12.1 Summary of Changes Between the Latest Version of Protocol and the Current SAP

12.2 Summary of Changes Between the Previous Version and the Current Version of the SAP

13.0 Appendix

14.0 References