1.0 Title Page

Statistical Analysis Plan for the Double-Blind Maintenance Study

Study M14-115

A Multicenter, Randomized, Double-Blind Study to Evaluate Higher Versus Standard Adalimumab Dosing Regimens for Induction and Maintenance Therapy in Subjects with Moderately to Severely Active Crohn's Disease and Evidence of Mucosal Ulceration

Date: 15 November 2019

Version 1.0

2.0 Table of Contents

1.0 Title Page 1
2.0 Table of Contents 2
3.0 Introduction 4
4.0 Study Objectives, Design and Procedures 4
4.1 Objectives 4
4.2 Study Design and Design Diagram 4
4.3 Sample Size 18
4.4 Interim Analysis 18
5.0 Analysis Populations 18
5.1 Definition for Analysis Populations 18
5.2 Variables Used for Stratification of Re-Randomization 18
6.0 Analysis Conventions 19
7.0 Demographics, Baseline Characteristics, Medical History, and Previous/Concomitant Medications 31
7.1 Demographic and Induction Baseline Characteristics 31
7.2 Medical History 33
7.3 Previous Treatment and Concomitant Medications 34
8.0 Patient Disposition 34
9.0 Study Drug Exposure and Compliance 35
10.0 Efficacy Analysis 35
10.1 General Considerations 35
10.2 Efficacy Analyses 36
10.3 Handling of Multiplicity 41
11.0 Safety Analysis 41
11.1 General Considerations 41
11.2 Analysis of Adverse Events 41
11.2.1 Treatment-Emergent Adverse Events 41
11.2.2 Adverse Events of Special Interest 43
11.3 Analysis of Laboratory Data 45
11.3.1 Assessment of Shifts from Induction Baseline in Liver-Specific Laboratory Variables 46

11.3.2 Hy's Law Cases 47
11.4 Analysis of Vital Signs and Weight 48
12.0 Summary of Changes 49
12.1 Summary of Changes Between the Latest Version of Protocol and the Current SAP 49
12.2 Summary of Changes Between the Previous Version and the Current Version of the SAP 50
13.0 Appendix 50
14.0 References 50

List of Tables

Table 1. Adalimumab Dose Adjustment Criteria 8
Table 2. Study Activities 12
Table 3. Visit Windows for Analysis of CDAI and Vital Signs for Maintenance Study 21
Table 4. Visit Windows for Analysis of Chemistry, Hematology, Urinalysis, hs-CRP, Fecal Calprotectin, Abdominal Pain Rating Scale, IBDQ, EQ-5D, and WPAI for Maintenance Study 21
Table 5. Visit Windows for Analysis of SES-CD for Maintenance Study 22
Table 6. Visit Windows for Analysis of Fistulas and EIM for Maintenance Study 22
Table 7. An EQ-5D-5L Value Set for England 30
Table 8. Criteria for Potentially Clinically Significant Vital Sign Findings 49

List of Figures

Figure 1. Therapeutic Drug Monitoring (TDM) Regimen Schematic 7
Figure 2. Study Schematic 8

3.0 Introduction

This statistical analysis plan (SAP) describes the statistical analysis to be completed by the AbbVie Global Statistics Department for study Protocol M14-115 Amendment 6. 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 the 44-week Double-Blind Maintenance Study of Protocol M14-115.

This analysis plan describes the efficacy analysis as well as the safety analysis for the double-blind maintenance study.

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 9.4 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 Maintenance study is to assess the efficacy and safety of two adalimumab maintenance regimens in reducing signs and symptoms of Crohn's disease at Week 56 as well as to assess pharmacokinetics (PK) and immunogenicity of two adalimumab induction regimens following subcutaneous (SC) administration.

4.2 Study Design and Design Diagram

This is a randomized, double-blind multicenter study of two adalimumab induction and maintenance regimens in subjects with moderately to severely active CD with evidence of mucosal ulceration confirmed by central reading. No placebo arm is planned since there is well-documented efficacy of adalimumab in CD and because the purpose of this study

is to achieve better efficacy than the standard induction and maintenance regimens in terms of clinical remission and endoscopic response. Additionally, since subjects are required to have failed or been intolerant of standard therapies and have evidence of endoscopic damage confirmed by a central reader, it would be not medically acceptable to deny those subjects effective treatment or feasible to enroll subjects in the trial when adalimumab is available to be prescribed for CD.

Approximately 500 adult subjects with active Crohn's disease, defined as having a CDAI of ≥ 220 and ≤ 450 and evidence of mucosal ulceration defined as SES-CD ≥ 6, excluding the presence of narrowing component, or SES-CD ≥ 4, excluding the presence of narrowing component, for patients with disease limited to the ileum on screening endoscopy or endoscopy performed within 45 days before Baseline, confirmed by a centrally read endoscopy, are enrolled at approximately 150 sites worldwide.

Subjects who meet all of the inclusion criteria and none of the exclusion criteria are enrolled into the study and randomized in a 3:2 ratio at Baseline to receive a higher induction adalimumab regimen or standard induction adalimumab regimen during the double-blind Induction Study.

The randomization of subjects for the Induction Study is stratified by hs-CRP at Baseline (< 10 and ≥ 10 mg/L), using the Screening hs-CRP value, prior infliximab use, and Crohn's disease activity (CDAI ≤ 300, > 300), at Baseline. Enrollment of subjects with prior infliximab use is limited to 25% of the total study population.

Subjects assigned to the higher induction regimen receive blinded adalimumab 160 mg at Baseline, Week 1, Week 2, and Week 3. At Week 4, subjects receive 40 mg every other week (eow) through Week 12. Subjects assigned to the standard adalimumab induction regimen receive blinded adalimumab 160 mg at Baseline and matching placebo at Week 1. Subjects receive 80 mg and matching placebo at Week 2, and then matching placebo at Week 3. At Week 4, subjects receive 40 mg every other week through Week 12.

At Week 12, subjects are re-randomized in a 1:1 ratio to one of the two double-blinded exploratory treatment regimens (adalimumab clinically adjusted [CA] regimen and adalimumab therapeutic drug monitoring [TDM] regimen). The re-randomization at Week 12 is stratified by induction treatment regimen, clinical response (CR-70) status at Week 12, and decrease in SES-CD > 50% from Baseline per the site investigator reading at Week 12. Among subjects achieving decrease in SES-CD > 50% from Baseline at Week 12, the randomization is further stratified by achievement of an SES-CD ≤ 4 and at least a 2 point reduction versus Baseline and no subscore greater than 1 in any individual variable using the Week 12 SES-CD value provided by the site. No study drug is administered or injected at the final visit.

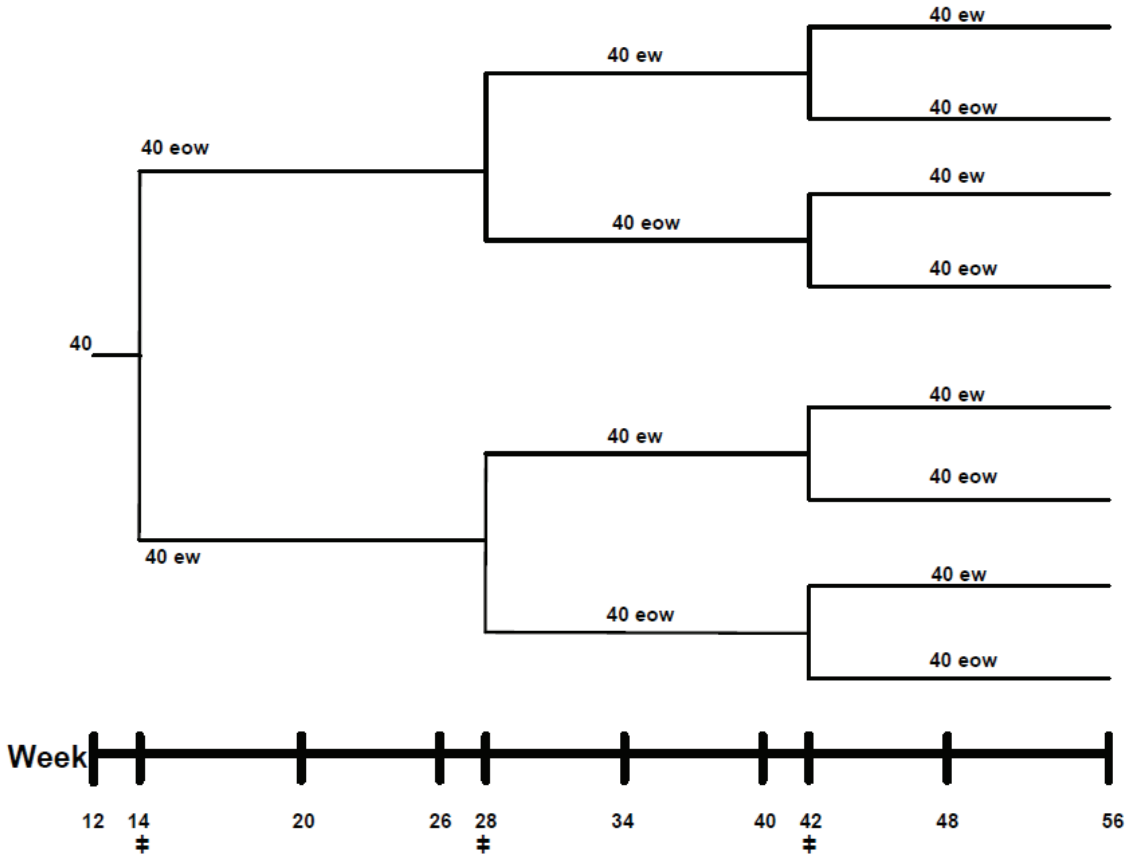
Clinically Adjusted (CA) Regimen:

Subjects randomized to the clinically adjusted regimen will receive 40 mg adalimumab every other week beginning at Week 12. The adalimumab dose will be escalated to every week starting as early as Week 14 if the subject's CDAI is ≥ 220 or hs-CRP ≥ 10 mg/L (using results from the prior or current visit). These subjects will also be allowed to escalate at unscheduled visits that may occur only on Weeks 16, 18, 22, 24, 30, 32, 36, 38, 44, 46, 50, 52 and 54. Once subjects in the CA regimen are escalated, they will remain on 40 mg ew dosing.

Therapeutic Drug Monitoring (TDM) Regimen:

At Weeks 14, 28 and 42, the adalimumab dose for subjects randomized to the TDM will be determined by the dose adjustment criteria table (Table 1). Doses will be determined using blinded serum concentrations at the prior visit (Weeks 12, 26 and 40, respectively) as well as the CDAI or hs-CRP values from the current or prior visit. For subjects who meet criteria for dose escalation at Weeks 14, 28 or 42, subjects will receive 40 mg weekly.

Figure 1. Therapeutic Drug Monitoring (TDM) Regimen Schematic



‡ Adjust dose based on labs taken 2 weeks prior

Table 1. Adalimumab Dose Adjustment Criteria

Adalimumab Dose Adjustment Criteria[a]			
ADA Serum Concentration[b] (ug/mL)	**CDAI[c]**	**hs-CRP[d] (mg/L)**	**Dose Change?**
Clinically Adjusted Regimen			
any	< 220	< 10	no
any	any	≥ 10	yes; dose escalate to ew
any	≥ 220	any	yes; dose escalate to ew
Therapeutic Drug Monitoring Regimen			
< 5	any	any	yes; dose escalate to ew
5 – 10	< 220	< 10	no
5 – 10	any	≥ 10	yes; dose escalate to ew
5 – 10	≥ 220	any	yes; dose escalate to ew
> 10	any	any	no

a. For subjects experiencing an active infection or those for whom the investigator feels dose escalation is not advisable, the investigator should contact the Study Designated Physician.

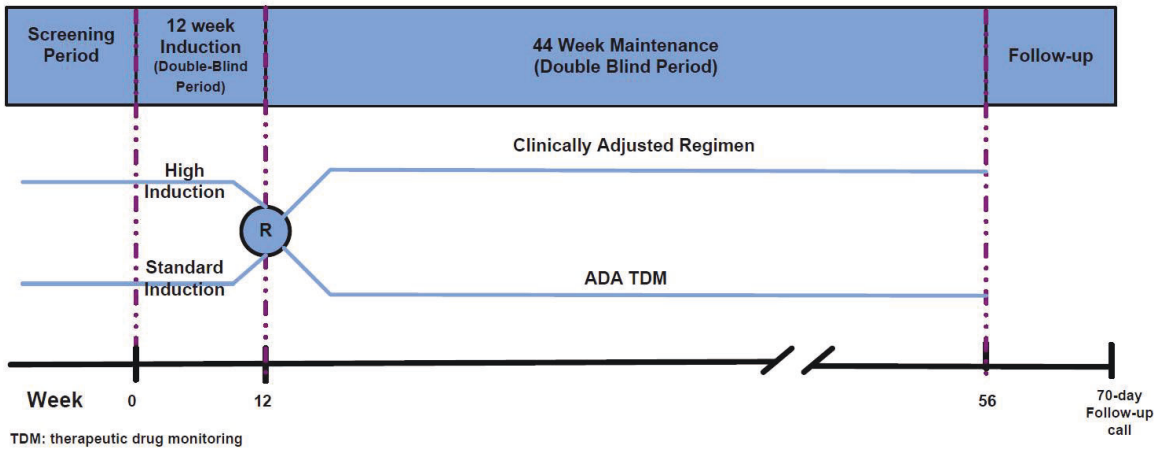
b. Measured from the serum concentration taken at the prior study visit.

c. Measured using hematocrit taken from prior study visit for CDAI calculation.

d. Measure using hs-CRP from the prior or current study visit.

A schematic of the study design is presented in Figure 2.

Figure 2. Study Schematic



Clinical evaluation is included at the following visits: Baseline, Weeks 2, 4, 6, 8, 12, 14, 20, 26, 28, 34, 40, 42, 48, and 56/PD. An electronic diary is dispensed at the Screening visit. In addition to routine physical examination, CDAI calculation, diary review, laboratory, adverse event, concomitant medication and vital sign assessments, the following are collected:

- Results of study questionnaires (inflammatory bowel disease questionnaire [IBDQ], European Quality of Life 5 dimensions [EQ-5D], work productivity and impairment [WPAI]) at Baseline, Week 4, Week 8, Week 12, Week 26, Week 40 and Week 56/PD.
- Calculation of the SFPS at Baseline, Week 2, Week 4, Week 6, Week 8, Week 12, Week 14, Week 20, Week 26, Week 28, Week 34, Week 40, Week 42, Week 48 and Week 56/PD. The Screening results serve as the Baseline value.
- Results of daily Bristol Stool Form Scale beginning at Baseline through Week 56/PD.
- Results of 11-point Abdominal Pain Rating Scale beginning at Baseline, Week 2, Week 4, Week 6, Week 8, Week 12, Week 26, Week 40, and Week 56/PD.
- Serum for measurement of adalimumab concentrations just prior to dosing at Baseline, Week 2, Week 4, Week 6, Week 8, Week 12, Week 26, Week 40, and Week 56/PD.
- Serum for measurement of Anti-Adalimumab Antibodies (AAA) just prior to dosing at Baseline, Week 4, Week 12, Week 26, Week 40, and Week 56/PD.
- Serum for measurement of infliximab serum levels and Human Anti-Chimeric Antibodies (HACA) just prior to dosing at Baseline.
- Serological biomarkers/mRNA at Baseline, Week 2, Week 4, Week 8, Week 12, Week 26, Week 40, and Week 56/PD.
- Stool samples for analysis of fecal calprotectin during Screening, at Week 4, Week 12, Week 26, and Week 56/PD.

- Stool samples for microbiota metagenomic analyses during Screening, at Week 4, Week 12, Week 26, and Week 56/PD. The stool samples should be taken before starting bowel preparations for endoscopy.
- Endoscopic evaluations, eligibility confirmed by central reader, are done at Screening, Week 12, and Week 56/PD.
- An optional pharmacogenetic sample should be drawn at Baseline and Week 12.

Throughout the study, subjects are only allowed to change the dosage of CD-specific concomitant medications as specified below:

- At Week 4, subjects who are taking corticosteroid therapy at Baseline have their corticosteroid therapy tapered according to a tapering schedule specified in the clinical study protocol. If the Investigator feels that the steroid taper is not advisable for a particular subject at Week 4, the Study Designated Physician (SDP) should be consulted for evaluation and approval.
- Subjects taking corticosteroids at Baseline who have a loss of satisfactory clinical response per the Investigator's judgment after the steroid taper has been initiated may have their corticosteroid dose increased per the Investigator's discretion during the study. Subjects in whom the maximum steroid dose equivalent exceeds the dose used at Baseline will be censored for efficacy assessments (i.e., will be considered non-responders for categorical endpoints and will have last non-missing values carried forward for non-categorical assessments) from that point forward. These subjects will continue to be evaluated in the safety population.
- Immunosuppressant doses may be decreased or terminated in the event of moderate-to-severe treatment-related toxicities.

The study was designed to enroll 500 subjects to meet scientific and regulatory objectives without enrolling an undue number of subjects in alignment with ethical considerations. Therefore, if the target number of subjects has been enrolled, there is a possibility that additional subjects in screening will not be enrolled.

Re-Screening

Subjects that initially screen fail for the study may be permitted to re-screen following re-consent. All screening procedures with the possible exceptions noted below will be repeated. The subject must meet all 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 In-Tube test or T SPOT TB test) chest x-ray (if applicable) and electrocardiogram (ECG), these tests will not be required to be repeated for re-screening provided the conditions noted in Protocol Section 5.3.1.1 are met and no more than 90 days have passed. If an endoscopy was performed, this will not be required to be repeated for re-screening provided the conditions noted in protocol Section 5.3.1 are met and no more than 45 days have passed between the endoscopy date and the Baseline date. As appropriate, sites are encouraged to contact the AbbVie Study Designated Physician (SDP) to confirm if subjects should or should not be re-screened.

Subjects may discontinue adalimumab treatment at any time during study participation. Subjects that end study participation early will have a Premature Discontinuation Visit. All subjects who discontinue from the study prematurely will have a follow-up phone call approximately 70 days after the last administration of study drug to obtain information on any new or ongoing adverse events (AEs). The 70-day follow-up phone call will not be required for any subject that initiates commercial adalimumab.

The study activities are presented in Table 2.

Table 2. Study Activities

Activity	Screening Period (30 Days)[a]	12-Week Double-Blind Induction Study						44-Week Double-Blind Maintenance Study									Unscheduled Visit[v]	70-Day Follow-Up[t]
	Screening	Baseline	Week 2	Week 4	Week 6	Week 8	Week 12 (Re-Randomization)	Week 14	Week 20	Week 26	Week 28	Week 34	Week 40	Week 42	Week 48	Week 56/Premature Discontinuation		
Informed Consent	X																	
Inclusion/Exclusion[b]	X	X																
Medical/Surgery History[b]	X	X																
Previous and Concomitant Medication[b]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Vital Signs[c]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Endoscopy[d]	X						X									X		
Physical Examination[e]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
TB Screening[f]	X																	
Chest X-Ray[g]	X																	
ECG[h]	X																	
Chemistry and Hematology[i]	X	X	X	X		X	X			X			X			X	X	
Urinalysis[i,j]	X	X	X	X	X	X	X			X			X			X	X	
Pregnancy Tests[k]	X[k]	X					X									X		
Hepatitis B Screen and HIV[l]	X																	

Table 2. Study Activities (Continued)

Activity	Screening Period (30 Days)[a]	12-Week Double-Blind Induction Study						44-Week Double-Blind Maintenance Study									Unscheduled Visit[v]	70-Day Follow-Up[t]
	Screening	Baseline	Week 2	Week 4	Week 6	Week 8	Week 12 (Re-Randomization)	Week 14	Week 20	Week 26	Week 28	Week 34	Week 40	Week 42	Week 48	Week 56/Premature Discontinuation		
hs-CRP	X		X	X	X	X	X			X			X			X	X	
C.-difficile toxin	X																	
Antinuclear antibody (ANA)/Anti-double-stranded DNA (anti dsDNA)[m]	X																	
Stool Sample (microbiota metagenomic analyses and fecal calprotectin)[n]	X			X			X			X			X			X		
Human Antichimeric Antibodies (HACA)/Infliximab Concentrations		X																
Adalimumab Concentration[o]		X	X	X	X	X	X			X			X			X	X	
AAA Concentration[o]		X		X			X			X			X			X	X	
Pharmacogenetic Marker[p]		X					X											
Serological Biomarkers/mRNA		X	X	X		X	X			X			X			X		

Table 2. Study Activities (Continued)

Activity	Screening Period (30 Days)[a]	12-Week Double-Blind Induction Study						44-Week Double-Blind Maintenance Study									Unscheduled Visit[v]	70-Day Follow-Up[t]
	Screening	Baseline	Week 2	Week 4	Week 6	Week 8	Week 12 (Re-Randomization)	Week 14	Week 20	Week 26	Week 28	Week 34	Week 40	Week 42	Week 48	Week 56/Premature Discontinuation		
Crohn's Disease Activity Index (CDAI) CDAI components "Number of liquid or very soft stools" and "Abdominal pain" (Stool [liquid/soft] Frequency + Abdominal Pain Score; SFPS)	X	X[u]	X	X	X	X	X	X	X	X	X	X	X	X	X	X		
SES-CD Score	X						X									X		
Inflammatory Bowel disease Questionnaire (IBDQ)		X		X		X	X			X			X			X		
European Quality of Life 5 dimensions (EQ‑5D)		X		X		X	X			X			X			X		
Work Productivity and Impairment Questionnaire (WPAI)		X		X		X	X			X			X			X		
Abdominal Pain Rating Scale		X	X	X	X	X	X			X			X			X		
Corticosteroid Taper[q]				X	X	X	X	X										
Monitor Adverse Events[r]		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Study Drug Dispensing/Administration[s]		X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	

Table 2. Study Activities (Continued)

Activity	Screening Period (30 Days)[a]	12-Week Double-Blind Induction Study						44-Week Double-Blind Maintenance Study									Unscheduled Visit[v]	70-Day Follow-Up[t]
	Screening	Baseline	Week 2	Week 4	Week 6	Week 8	Week 12 (Re-Randomization)	Week 14	Week 20	Week 26	Week 28	Week 34	Week 40	Week 42	Week 48	Week 56/Premature Discontinuation		
Dispense Subject Diary	X																	
Subject Diary Review		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Intestinal Biopsies[p]	X						X									X		

a. The Screening period will be a minimum of 7 days for CDAI calculation. The CDAI calculated at Screening will serve as the Baseline CDAI. Baseline visit date will serve as the reference for all subsequent visits. A ± 3-day window is permitted around all study visits.

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

c. Height will be measured at Screening only (with shoes off, and then adding 1 inch or 2.5 cm).

d. An ileocolonoscopy will be performed during Screening or one performed within 45 days before Baseline, Week 12, and at Week 56/PD. A biopsy will be performed at Screening, Week 12, and Week 56/PD if a suspicious lesion or suspected malignancy, in the assessment of the Investigator, is observed, and evaluated by the local pathologist. Site staff should schedule the Week 12 Endoscopy during the Baseline visit, if possible. Appropriate documentation of biopsy results consistent with the diagnosis of CD, in the assessment of the Investigator, must be available in order to confirm the subject's eligibility for the study. If this documentation is not available, a diagnostic biopsy from the most affected observed area of the ileum/colon must be performed during the Screening endoscopy and evaluated by a qualified local pathologist and the results reviewed by the Investigator. Biopsies to evaluate suspicious lesions and to rule out malignancy may be taken during any study endoscopy per the Investigator's discretion and evaluated by the local pathologist.

Table 2. Study Activities (Continued)

e. Physical examination performed at Screening, Week 12 and Week 56/Premature Discontinuation 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 at all other visits are symptom based and must include a count of the number of cutaneous fistulas.

f. Subjects with negative PPD test and/or QuantiFERON-TB Gold test within 90 days of Screening will not require a repeat skin test, if documentation is available. PPD skin test is to be read 48 to 72 hours after placement.

g. Chest x-ray includes posterior-anterior (PA) and lateral views. Subjects with normal chest x-ray within 3 months of Screening would not require a repeat chest x-ray, if documentation is available.

h. Subjects with normal ECG within 90 days of Screening would not require a repeat ECG, if documentation is available.

i. Lab assessments will only need to be repeated at Baseline if the time between Screening and Baseline is greater than 14 days, or if the subject's health status has changed to warrant a repeat test.

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. Further explanations of these tests are provided in the laboratory manual.

k. Serum pregnancy test will be performed on all women of childbearing potential at Screening. Urine pregnancy test will be performed locally at Baseline Visit, and at the Week 12 and Week 56/PD for all women of child-bearing potential. The frequency can be increased up to every visit as per local regulations. If any urine pregnancy test is positive, a serum pregnancy test will be performed by the central laboratory.

l. Subjects will be tested for the presence of the Hepatitis B Virus (HBV) at Screening. A positive result for the Hepatitis B surface antigen (HBs Ag) will be exclusionary. Please refer to hepatitis procedures for details on testing requirements. If required by country regulatory authorities to confirm eligibility, subjects will be tested for HIV and documented that the test has been performed. This testing is to be done at a local laboratory. A subject will not be eligible for study participation if test results indicate a positive HIV infection. AbbVie will not receive results from the testing and not be made aware of any positive result.

m. Anti-dsDNA performed if ANA result is positive.

n. If a sample cannot be obtained during the site visit, the site will give instructions and a stool sample supply kit (supplies will be provided at the time points indicated). The stool from which these samples are prepared should be scored using the Bristol stool chart by ePRO. All stool samples for metagenomic analysis should be collected before any bowel preparation for endoscopy is started and should be returned to the site within 3 days of collection.

   Note: The Screening stool sample may be taken anytime during the Screening period but should be collected prior to any bowel prep.

o. Blood samples for the measurement of adalimumab and AAA concentrations will be collected prior to dosing. Testing of the adalimumab and AAA concentrations must not be performed locally. All pharmacokinetic results will remain blinded to the investigator, study site personnel and the subject throughout the study.

Table 2. Study Activities (Continued)

p. Pharmacogenetic Marker and intestinal biopsy samples are optional. Separate consents must be signed prior to the sample draw. If the pharmacogenetic sample is not collected at Baseline, preferably it should be collected at the next study visit.

q. Subject will begin corticosteroid taper at Week 4. If the Investigator feels that the steroid taper is not advisable for a particular subject at Week 4, the SDP should be consulted for evaluation and approval.

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

s. Administration of drug will be performed after all assessments and examinations scheduled for that day have been completed.

t. Subjects will be contacted 70 days following study drug discontinuation for an assessment of any new or ongoing AEs, except those subjects that continue on adalimumab therapy after the end of study participation.

u. The Screening results of the SFPS will serve as the Baseline value.

v. Visits for dispensing new study drug in case of temperature excursion, loss or damage are not considered an Unscheduled Visit. In addition, visits to retest a lab will not be considered an Unscheduled Visit. Unscheduled Visits according to this table are for purposes when the subject is coming in for a visit for evaluation, assessment and potential dose escalation.

4.3 Sample Size

There is no sample size calculation for the Maintenance study.

4.4 Interim Analysis

No interim analysis is planned for the Maintenance study.

5.0 Analysis Populations

5.1 Definition for Analysis Populations

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

Modified Intent-to-Treat (mITT) set includes all ITT subjects who achieve clinical response (CR-70) at Week 12. The mITT set is the primary population for the efficacy analysis for the Maintenance Study.

The safety set consists of all subjects who received at least one injection of study drug during maintenance study. The safety set will be analyzed as treated, according to treatment the subject actually received. The safety set will be used only for safety analysis.

5.2 Variables Used for Stratification of Re-Randomization

At Week 12, subjects are re-randomized in a 1:1 ratio to one of two double-blinded exploratory treatment regimens (clinically adjusted [CA] regimen and adalimumab therapeutic drug monitoring [TDM] regimen). The re-randomization at Week 12 is stratified by induction treatment regimen, clinical response (CR-70) status at Week 12, and decrease in SES-CD > 50% from Baseline per the site investigator reading at Week 12. Among subjects achieving decrease in SES-CD > 50% from Baseline at Week 12, the randomization is further stratified by achievement of an SES-CD $\leq$ 4 and at least a 2 point reduction versus baseline and no subscore greater than 1 in any individual variable using the Week 12 SES-CD value provided by the site.

The randomization schedule was prepared by the Statistics Department of AbbVie.

6.0 Analysis Conventions

Definition of Induction Baseline

The Induction Baseline visit date is the date when the first dose of study drug is received and referred to as Day 1 or Week 0. If a subject has not taken any study drug during Induction Study, the randomization date will be used as Day 1. 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 Maintenance Baseline

The Maintenance Baseline visit date is the date when the first dose of study drug during Maintenance Study is received and referred to as Day 1 of Maintenance Study or Week 0 of Maintenance Study. If a subject has not taken any study drug during Maintenance Study, the re-randomization date will be used as Day 1. The Maintenance Baseline value for a variable is defined as the last non-missing value on or before Day 1 of Maintenance Study, and after the Induction Baseline.

Definition of Maintenance Final Observation

Final Observation in the 56-week period (44-week double-blind Maintenance Study) is defined as the last non-missing observation collected within 70 days following the last dose of study drug who are re randomized.

Definition of Rx Days (Days Relative to the First Dose of Maintenance 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 of Maintenance Study, while the day prior to the date of first dose is defined as Rx Day –1 of Maintenance Study (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, 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 one with later time will be used if collection time is also available, or 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. If more than one assessment is collected on the same day, then the average of those assessments will be used in analyses.

Table 3. Visit Windows for Analysis of CDAI and Vital Signs for Maintenance Study

Scheduled Week	Nominal Day	Time Window (Rx Day Range)
Week 12 (Week 0 of Maintenance Study)	1[a]	≤ 1
Week 14 (Week 2 of Maintenance Study)	15	2 – 36
Week 20 (Week 8 of Maintenance Study)	57	37 – 78
Week 26 (Week 14 of Maintenance Study)	99	79 – 106
Week 28 (Week 16 of Maintenance Study)	113	107 – 134
Week 34 (Week 22 of Maintenance Study)	155	135 – 176
Week 40 (Week 28 of Maintenance Study)	197	177 – 204
Week 42 (Week 30 of Maintenance Study)	211	205 – 232
Week 48 (Week 36 of Maintenance Study)	253	233 – 281
Week 56 (Week 44 of Maintenance Study)	309	282 – 379

Rx Day = date of visit – date of first study drug injection in double-blind maintenance study + 1

a. Day of first dose of study drug in double-blind maintenance study. If a subject has not taken any study drug during Maintenance Study, the randomization date will be use as Day 1 of Maintenance Study.

Table 4. Visit Windows for Analysis of Chemistry, Hematology, Urinalysis, hs-CRP, Fecal Calprotectin, Abdominal Pain Rating Scale, IBDQ, EQ-5D, and WPAI for Maintenance Study

Scheduled Week	Nominal Day	Time Window (Rx Day Range)
Week 12 (Week 0 of Maintenance Study)	1[a]	≤ 1
Week 26 (Week 14 of Maintenance Study)	99	2 – 148
Week 40 (Week 28 of Maintenance Study)	197	149 – 253
Week 56 (Week 44 of Maintenance Study)	309	254 – 379

Rx Day = date of visit – date of first study drug injection in double-blind maintenance study + 1

a. Day of first dose of study drug in double-blind maintenance study. If a subject has not taken any study drug during Maintenance Study, the randomization date will be use as Day 1 of Maintenance Study.

Table 5. Visit Windows for Analysis of SES-CD for Maintenance Study

Scheduled Week	Nominal Day	Time Window (Rx Day Range)
Week 12 (Week 0 of Maintenance Study)	1[a]	≤ 1
Week 56 (Week 44 of Maintenance Study)	309	2 – 379

Rx Day = date of visit – date of first study drug injection in double-blind maintenance study + 1

a. Day of first dose of study drug in double-blind maintenance study. If a subject has not taken any study drug during Maintenance Study, the randomization date will be use as Day 1 of Maintenance Study.

Table 6. Visit Windows for Analysis of Fistulas and EIM for Maintenance Study

Scheduled Week	Nominal Day	Time Window (Rx Day Range)
Week 12 (Week 0 of Maintenance Study)	1[a]	≤ 1
Week 56 (Week 44 of Maintenance Study)	309	2 – 379

Rx Day = date of visit – date of first study drug injection in double-blind maintenance study + 1

a. Day of first dose of study drug in double-blind maintenance study. If a subject has not taken any study drug during Maintenance Study, the randomization date will be use as Day 1 of Maintenance Study.

CD-Related Concomitant Corticosteroid Therapy

Subjects in whom the CD-related corticosteroids (systemic or rectal corticosteroids) that were not being taken at Induction Baseline and are initiated during the study or who have equivalent dose of these medications increased to greater than the dose taken at Induction Baseline will be censored for efficacy assessments (i.e., will be considered non-responders for categorical endpoints and will have last non-missing 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.

Subjects in whom systemic corticosteroids were used at Induction Baseline and initiated rectal corticosteroids during the study will be censored for efficacy assessment, regardless of rectal corticosteroid dose.

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. Maintenance Baseline and Pre-Maintenance Baseline values will not be used to impute the missing Post-Maintenance 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 Maintenance Baseline, then the LOCF value will be missing.

Mixed-Effect Model Repeated Measure (MMRM)

The MMRM model will be used for continuous efficacy variables with longitudinal data as a sensitivity analysis. The MMRM model includes the Induction baseline values as covariate; randomization strata, region (US versus non-US), 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.

Observed Case (OC)

Observed case analysis will be performed such that missing values will not be imputed. Subjects in whom the following CD-related systemic or rectal corticosteroids that were not being taken at Induction Baseline and are initiated during the study or who have dosages of these medications increased to greater than the dose taken at Induction Baseline will be censored for efficacy assessments (i.e., will be considered non-responders for categorical endpoints and the observed non-categorical assessments will be replaced by last non-missing values prior to censoring) from that point through the end of the study.

Subjects in whom systemic corticosteroids were used at Induction Baseline and initiated rectal corticosteroids during the study will be censored for efficacy assessment, regardless of rectal corticosteroid dose.

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

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.

Sensitivity analyses will be performed to make use of all available data based on the above rule if more than 5% of subjects have less than 4 days of diary data.

Definition of Clinical Remission

CDAI < 150.

Definition of Clinical Response

Decrease in CDAI ≥ 70 points from Induction Baseline.

Definition of Enhanced Clinical Response

Decrease in CDAI ≥ 100 points from Induction Baseline.

Definition of Endoscopic Remission

Simplified Endoscopic Score for Crohn's Disease (SES-CD) ≤ 4 and at least 2 point reduction versus Induction baseline and no subscore greater than 1 in any individual variable, as scored by central reviewer.

Definition of Endoscopic Response

Decrease in SES-CD > 50% from Induction Baseline (or for a Induction Baseline SES-CD of 4, at least a 2 point reduction from Induction Baseline), as scored by central reviewer.

Stool Frequency + Abdominal Pain Score (SFPS)

SFPS is the CDAI components "Number of liquid or very soft stools" and "Abdominal pain" (Stool [liquid/soft] Frequency + Abdominal Pain Score). The sub-total score for each component is calculated as the sum of the 7 day diary score multiplied by the corresponding multiplier.

Definition of Time to Dose Escalation

Time to Dose Escalation is defined as number of days from Study Day 1 in Maintenance Study to the date of Dose Escalation. It is calculated as (date of dose escalation – Study Day 1 date in Maintenance Study). Subjects who discontinue prior to dose escalation are censored at the time of discontinuation. Subjects who are not experiencing dose escalation by Week 56 are censored at the Week 56 visit.

Definition of Symptomatic 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). The analysis will be only performed on subjects with induction baseline average daily stool frequency score ≥ 4.0 or average daily abdominal pain score ≥ 2.0.

Definition of Symptomatic 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. The analysis will be only performed on subjects with induction baseline average daily stool frequency score ≥ 4.0 or average daily abdominal pain score ≥ 2.0.

Discontinued Corticosteroid Use at Each Visit during Maintenance Study Among Subjects Who Used Corticosteroids at Induction Baseline

Discontinued steroid use at each visit is based on the steroid use on the date of CDAI measurement for that visit. If the subject has no CDAI for the visit, the date of vital signs will be used. If no vital signs date, then the steroid use on the nominal day for that visit will be evaluated.

If a subject discontinued the study, 'discontinued corticosteroid use' will be classified as No from that point through the end of the study.

If a subject met censoring criteria for CD-related concomitant therapy, 'discontinued steroid use' will be classified as No from that point through the end of the study.

SES-CD Scoring

All videotaped colonoscopies will undergo central review providing the investigator's SES-CD score meets the entry criteria. Two primary central reviewers will evaluate the

videotaped colonoscopies separately and provide their SES-CD scores to Parexel. A third central reviewer will adjudicate between the two initial reviewers' SES-CD scores if there is discrepancy in any SES-CD variable. The adjudicator will select the final SES-CD score that he/she most agrees with from those provided by the two primary reviewers, 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 two primary central reviewers, that 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 that serves as the SES-CD for the study's efficacy analyses, the following imputation rules will be applied:

The missing SES-CD individual variables at Baseline will be imputed as 0. The Week 12 or Week 56 missing SES-CD individual variables will be imputed as 0 if there are 8 or fewer missing individual variables. If more than 8 individual variables are missing, the total SES-CD will be missing for that visit.

Bristol Stool Chart Score

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. Bristol Stool Form Scale is not a continuous variable. The new variable Bristol Stool Chart score is defined 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 decimals.

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

If a subject achieved reduction from Baseline $\geq$ 50%, this subject will be classified as Bristol Stool Chart Responder.

Hs-CRP

From November 2014 to December 2015, Roche hs-CRP kit lot 604450 was in use at ICON. In December 2015 ICON altered the study team of a bias in the lot of the Roche Diagnostic C-creative protein high sensitive reagent. ICON then informed the study team of the extent of the under recovery in March 2016. This lot was "retrospectively expired" as the low standard beads agglutinated causing specimen results that are below 5 mg/L to be up to 25% lower than the accurate value. ICON later determined with Roche that the affected samples were those tested from May 18, 2015 to December 2015.

In the Study M14-115 approximately 423 samples were affected according to ICON. AbbVie obtained the final listing of affected samples/results from ICON on August 10, 2016. A correction factor developed by Roche and ICON will be applied to all affected specimens with a concentration between 0.51 to 4.99 mg/L and will replace the initially provided hs-CRP values in the study database.

For Study M14-115, both Roche and Abbott platform are used to measure hs-CRP. The study team has decided to use Abbott platform for CSR analysis. If hs-CRP was measured by Roche platform only, the value will be converted to Abbott equivalent by a conversion factor.

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 7).

Table 7. 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 Induction Baseline Characteristics

For the subjects in mITT and safety analysis set, demographic information and Induction 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)
- Respiratory Rate (RPM)
- Temperature (°C)
- IBDQ score

- CDAI
- SFPS
- Total SES-CD score
- Crohn's Disease Duration (years)
- hs-CRP mg/L
- WPAI and its components
- EQ-5D
- Abdominal Pain Rating Scale
- Bristol Stool Chart score
- Fecal calprotectin

Categorical Variables:

- Sex (male, female)
- Race
- Ethnicity
- Age (≤ median, > median)
- Age (< 40, 40 to < 65, 65 to < 75, ≥ 75)
- Baseline fecal calprotectin (≤ median, > median)
- 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)
- Baseline fecal calprotectin (≤ median, > median)
- Baseline fecal calprotectin (≤ 250 µg/g, > 250 µg/g)
- Crohn's disease severity (CDAI ≤ 300, > 300) at Baseline
- Baseline CDAI (≤ median, > median)
- Baseline SES-CD (≤ median, > median)
- Prior infliximab use (or prior anti-TNF use for subjects randomized under original protocol) (yes, no)

- Weight (≤ median, > median)
- Baseline albumin (≤ median, > median)
- Disease duration (≤ median, > median)
- 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, will be obtained from each subject during the Screening period. Medical history will be summarized using body system and preferred term by treatment group. No statistical tests will be performed.

Chest X-Ray Results: All subjects undergo a standard chest 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 or 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. No statistical tests will be performed.

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

Prior medications are those medications taken prior to the first dose of study drug. This includes medications with a start date before the first study drug administration date, regardless of the end date of these medications. Medications taken on the day of the first dose of study drug are not counted as prior medications.

Concomitant medications are those medications, other than study drug, taken after the first dose of study drug in Maintenance, which include medications with a start date prior to the Maintenance Baseline date which are continuing after Baseline and all medications with a start date between the Maintenance Baseline date and last study drug administration + 14 days.

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 mITT 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 [0, ≤ 10 mg, ≤ 20 mg, or ≤ 30 mg prednisone equivalents], aminosalicylates, immunosuppressants, and antibiotics) at the Induction Baseline will be tabulated. In addition, the number and percent of subjects using any Crohn's disease specific medications including infliximab prior to Induction 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 44-week Maintenance study
- Subject disposition including the number and percent of subjects who prematurely discontinued the Maintenance 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 (total patient years) will be summarized by treatment group.

The number of injections received at each schedule time point will be summarized with number and percent for each treatment group.

Study Drug Exposure (in Days) during the Maintenance Study:

- Date of last dose of Maintenance Study drug – Date of first dose of Maintenance Study drug + 14 days

Study Drug Compliance:

Treatment compliance will be summarized for each treatment group in the mITT population. The treatment compliance is defined as the number of injections actually taken by the subject divided by the number of injections planned to be taken by the subject during the Maintenance Study.

10.0 Efficacy Analysis

10.1 General Considerations

All statistical tests will be two-sided with the significance level of 0.05. 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® Version 9.4 or later (SAS Institute Inc., Cary, NC, USA).

10.2 Efficacy Analyses

This section provides the details of the efficacy analysis for the Maintenance Study.

For these endpoints which are categorical variables, NRI method of imputation will be used for the missing values. LOCF will be used as a sensitivity analysis wherever applicable.

Both LOCF and OC analyses will be performed for continuous endpoints. The LOCF analysis is considered primary for inferential purposes. In additional, Mixed-Effect Model Repeated Measure (MMRM) will be applied as a sensitivity analysis, wherever appropriate, for the longitudinal continuous endpoints.

For endoscopy related endpoints, there will be no LOCF imputation as there is only one post-baseline endoscopy during Maintenance. OC will be used as a sensitivity analysis.

The difference in proportions of subjects between treatment groups will be analyzed using the two-sided Cochran-Mantel-Haenszel (CMH) test adjusted for induction treatment regimen, and decrease in SES-CD > 50% from Induction Baseline per central reading at Week 12. Additionally, the CMH-based two-sided 95% confidence interval (CI) for the difference in the proportions between the treatment groups will be calculated.

The difference in change from Induction Baseline between treatment groups will be analyzed using an ANCOVA model including factors of treatment, induction treatment regimen, decrease in SES-CD > 50% from Induction Baseline per central reading at Week 12 and Induction Baseline values. Parameter estimates with 95% confidence interval and *P*-value will be provided.

All efficacy endpoints for Maintenance Study are non-ranked.

- Proportion of subjects who achieve endoscopic response (SES-CD > 50% from Induction Baseline [or for a Induction Baseline SES-CD of 4, at least a 2 point

reduction from Induction Baseline]) at Week 56 among subjects with endoscopic response at Week 12

- Proportion of subjects who achieve endoscopic remission (SES-CD ≤ 4 and at least a 2 point reduction versus Induction Baseline and no subscore greater than 1 in any individual variable) at Week 56 among subjects with endoscopic remission at Week 12.
- Proportion of subjects who achieve endoscopic remission (SES-CD ≤ 4 and at least a 2 point reduction versus Induction Baseline and no subscore greater than 1 in any individual variable) at Week 56 among subjects with endoscopic response at Week 12.
- Proportion of subjects who achieve sustained clinical remission, CDAI < 150 at Week 56 among subjects with CDAI < 150 at Week 12.
- Proportion of subjects who achieve clinical remission (CDAI < 150) at Week 56.
- Proportion of subjects who achieve endoscopic remission (SES-CD ≤ 4 and at least 2 point reduction versus Induction Baseline and no subscore greater than 1 in any individual variable) at Week 56.
- Proportion of subjects who achieve SES-CD ≤ 2 at Week 56.
- Proportion of subjects with deep remission, CDAI < 150 at Week 56 and SES-CD ≤ 4 and at least 2 point reduction versus Induction Baseline and no subscore greater than 1 in any individual variable at Week 56.
- Proportion of subjects who discontinued corticosteroid use and achieved clinical remission (CDAI < 150) at Week 56 among subjects taking corticosteroid at Induction Baseline.
- Proportion of subjects with endoscopic response (decrease > 50% SES-CD from Induction Baseline [or for a Induction Baseline SES-CD of 4, at least a 2 point reduction from Induction Baseline]) at Week 56.
- Change from Induction Baseline in fecal calprotectin level at Week 56.
- Proportion of subjects with hs-CRP < 5 mg/L and fecal calprotectin < 250 μg/g at Week 56.
- Proportion of subjects with CDAI < 150, hs-CRP < 5 mg/L, and fecal calprotectin < 250 μg/g at Week 56.

- Proportion of subjects with CDAI < 150, hs-CRP < 5 mg/L, SES-CD ≤ 4 and at least 2 point reduction versus Induction Baseline and no subscore greater than 1 in any individual variable, and fecal calprotectin < 250 μg/g at Week 56.
- Proportion of subjects with endoscopic response, and ≥ 50% decrease from Induction Baseline in hs-CRP or fecal calprotectin at Week 56.
- Proportion of subjects with endoscopic remission, and ≥ 50% decrease from Induction Baseline in hs-CRP or fecal calprotectin at Week 56.
- Proportion of subjects with clinical remission, and ≥ 50% decrease from Induction Baseline in hs-CRP or fecal calprotectin at Week 56.
- Proportion of subjects with clinical response (decrease in CDAI ≥ 70 points from Induction Baseline) at Week 56.
- Proportion of subjects with clinical response (decrease CDAI ≥ 70 points from Induction Baseline) at each scheduled visit in Maintenance Study.
- Proportion of subjects with enhanced clinical response (decrease CDAI ≥ 100 points from Induction Baseline) at each scheduled visit in Maintenance Study.
- Proportion of subjects who discontinue corticosteroid use at each scheduled visit in Maintenance Study among subjects taking corticosteroid at Induction Baseline.
- Proportion of subjects who achieve a composite subtotal score of SFPS < 50 at Week 56 who had an SFPS ≥ 100 at Induction Baseline.
- Proportion of subjects who achieve SES-CD ≤ 3 and at least a 2 point reduction versus Induction Baseline and no subscore greater than 1 in any individual variable at Week 56.
- Proportion of subjects with SES-CD = 0 at Week 56.
- Change from Induction Baseline in fecal calprotectin level at each scheduled visit in Maintenance Study.
- Change from Induction Baseline in hs-CRP at each scheduled visit in Maintenance Study.

- Change in IBDQ total score and individual IBDQ domain scores (bowel, emotional, social, systemic) from Induction Baseline at each scheduled visit in Maintenance Study.
- Change in WPAI from Induction Baseline at each scheduled visit in Maintenance Study.
- Change in EQ-5D from Induction Baseline at each scheduled visit in Maintenance Study.
- Change in CDAI from Induction Baseline at each scheduled visit in Maintenance Study.
- Change in SFPS from Induction Baseline at each scheduled visit in Maintenance Study.
- Change in Abdominal Pain Rating Scale score from Induction Baseline at each scheduled visit in Maintenance Study.
- Change in Bristol Stool Chart score from Induction Baseline at each scheduled visit in Maintenance Study.
- Proportion of subjects who achieve CDAI remission (CDAI < 150) at each scheduled visit in Maintenance Study.
- Proportion of subjects who achieve SFPS remission (SFPS < 50) at each scheduled visit in Maintenance Study.
- Proportion of subjects with major CD related event (e.g., hospitalization, bowel surgery, abscess drainage) in Maintenance Study.
- Proportion of subjects with no draining fistulas at Week 56 among subjects with draining fistula at Induction Baseline.
- Proportion of subjects in each treatment group with > 50% reduction from Induction Baseline in the number of draining fistulas at Week 56 among subjects with draining fistula at Induction Baseline.
- Resolution of extra-intestinal manifestations at each scheduled visit in Maintenance Study.
- Proportion of subjects with an SES-CD decrease of ≥ 3 points compared to Induction Baseline at Week 56.
- Proportion of subjects who achieve symptomatic remission, defined as 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), at each scheduled visit in Maintenance Study among subjects with Induction Baseline SF ≥ 4.0 and/or AP ≥ 2.0.

- Proportion of subjects who achieve symptomatic response, defined as 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, at each scheduled visit in Maintenance Study among subjects with Induction Baseline SF ≥ 4.0 and/or AP ≥ 2.0.
- Time to dose escalation in Maintenance Study.
- Proportion of subjects with Inflammatory Bowel Disease Questionnaire (IBDQ) response (increase ≥ 16 points from Induction Baseline) at each scheduled visit in Maintenance Study.
- Proportion of subjects with IBDQ remission (IBDQ ≥ 170 points) at each scheduled visit in Maintenance Study.
- Proportion of subjects achieving response in IBDQ Bowel Symptom domain (increase of IBDQ bowel symptom domain score ≥ 8) at each scheduled visit in Maintenance Study.
- Proportion of subjects achieving response in IBDQ fatigue item (increase of IBDQ fatigue item score ≥ 1) at each scheduled visit in Maintenance Study.
- Proportion of subjects requiring dose escalation to weekly dosing during Maintenance Study.
- Proportion of subjects who achieve clinical remission (CDAI < 150) at Week 56 among subjects requiring dose escalation to weekly dosing during Maintenance Study.
- Proportion of subjects who achieve endoscopic response (decrease > 50% SES-CD from Induction Baseline [or for an Induction Baseline SES-CD of 4, at least a 2 point reduction from Induction Baseline]) at Week 56 among subjects requiring dose escalation to weekly dosing during Maintenance Study.
- Proportion of subjects who achieve endoscopic remission (SES-CD ≤ 4 and at least a 2 point reduction versus Induction Baseline and no subscore greater

than 1 in any individual variable) at Week 56 among subjects requiring dose escalation to weekly dosing during Maintenance Study.

Major CD related events will be identified by review of serious adverse events (SAEs) according to the hospitalization adjudication charter. *P* values from both Log-rank test and Cox proportional hazards model will be provided.

10.3 Handling of Multiplicity

No adjustment for multiplicity will be done.

11.0 Safety Analysis

11.1 General Considerations

All safety analyses will be performed on the safety analysis set. 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 AEs during Maintenance Study are defined as events that begin or worsen either on or after the first dose of the study drug and within 70 days after the last dose of the study drug.

The 70-day follow-up phone call will not be required for any subject that initiates commercial adalimumab.

An overview of treatment-emergent AEs, including AEs of special interest such as adverse events leading to death and adverse events leading to premature discontinuation, AEs by Medical Dictionary for Drug Regulatory Activities (MedDRA version 21.0 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 percent.

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.
- Any Deaths

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.

All Adverse Events identified to AbbVie from the 70-day follow-up phone call will be collected as source data to be evaluated and reported. Thus, all SAEs and nonserious AEs as defined by AbbVie, reported during the 70-day follow-up phone call must be captured in the clinical database. The end of trial is the last subject contact, i.e., the 70-day follow-up call.

11.2.2 Adverse Events of Special Interest

The following AEs of special interest will be summarized by number and percentage of subjects experiencing an AE of interest. The AEs of interest will be summarized and presented using primary MedDRA system organ classes (SOCs) and preferred terms (PTs) for the following AE categories:

- Any Infections AE

- Any Serious Infection AE
- Any Legionella Infection AE
- Any Diverticulitis AE
- Any Opportunistic Infection AE (Excluding Oral Candidiasis and TB)
- Any Oral Candidiasis
- Any Tuberculosis AE
- Any Active Tuberculosis
- Any Latent Tuberculosis
- Any Parasitic Infection AE
- Any Reactivation of Hepatitis B
- Any Progressive Multifocal Leukoencephalopathy (PML) AE
- Any Malignancy AE
- Any Lymphoma AE
- Any Hepatosplenic T-Cell Lymphoma AE (HSTCL)
- Any Non-Melanoma Skin Cancer (NMSC) AE
- Any Melanoma AE
- Any Leukemia AE
- Any Other Malignant AE (Excluding NMSC, Melanoma, Lymphoma, HSTCL, and Leukemia)
- Any Allergic Reaction (Including Angioedema/Anaphylaxis)
- Any Lupus-Like Reactions and Systemic Lupus Erythematosus
- Any Vasculitis AE
- Any Cutaneous Vasculitis AE
- Any Non-Cutaneous Vasculitis AE
- Any Sarcoidosis AE
- Any Autoimmune Hepatitis
- Any Myocardial Infarction Related AE
- Any Cerebrovascular Accident Related AE
- Any Congestive Heart Failure Related AE

- Any Pulmonary Embolism Related AE
- Any Interstitial Lung Disease AE
- Any Intestinal Perforation AE
- Any Pancreatitis AE
- Any Stevens-Johnson Syndrome AE
- Any Erythema Multiforme Related AE
- Any Worsening/New Onset of Psoriasis
- Any Demyelinating Disorder
- Any Amyotrophic Lateral Sclerosis AE
- Any Reversible Posterior Leukoencephalopathy Syndrome (RPLS) AE
- Any Hematologic Disorders AE (Including Pancytopenia)
- Any Liver Failure and Other Liver Event AE
- Any Humira Administration Related Medication Errors AE
- Any Injection Site Reaction AE
- Any Intestinal Stricture
- Any AE Leading to Death
- Any AE Leading to Discontinuation of Study Drug.
- Any Deaths

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

11.3 Analysis of Laboratory Data

All laboratory data collected after first dose of maintenance study drug will be included in this analysis.

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.

Cross (Shift) tables from Baseline to the minimum, maximum, and 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 3.0 of Grade ≥ 3 will be provided. 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.

The laboratory data using Abbott platform will be used for this analysis. If the data measured using Abbott platform are not available, the Abbott equivalent (the values measured using Roche platform that were converted to Abbott equivalent) will be used.

11.3.1 Assessment of Shifts from Induction Baseline in Liver-Specific Laboratory Variables

Induction Baseline and post-maintenance baseline livers-specific laboratory will be categorized as follows:

- < 1.5 times the upper limit of the reference range
- ≥ 1.5 to < 3 the upper limit of the reference range
- ≥ 3 to < 5 the upper limit of the reference range
- ≥ 5 to < 8 the upper limit of the reference range
- ≥ 8 the upper limit of the reference range

For each variable, shift tables will be generated as deemed appropriate:

- Category of the induction baseline value versus category of the final value,
- Category of the induction baseline value versus post-maintenance baseline high/low category

Note that the maximum category is used, rather than the category of the maximum values. The two may be different due to variation in the reference range.

No statistical tests will be performed for this analysis.

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

ALT ≥ 2.5 × ULN, or

AST ≥ 2.5 × ULN, or

Alkaline Phosphatase ≥ 2.5 × ULN, or

Total Bilirubin ≥ 1.5 × ULN.

11.3.2 Hy's Law Cases

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

All vital signs and weight data collected after first dose of maintenance study drug will be included in this analysis.

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

- Body Weight (kg)
- Blood Pressure (Systolic/Diastolic) (mmHg)
- Pulse (bpm)
- Respiratory Rate (RPM)
- 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 number and percentage of subjects meeting the criteria for potentially clinically significant (PCS) vital sign values will also be summarized.

The criteria for potentially clinically significant vital sign findings are presented in Table 8.

Table 8. Criteria for Potentially Clinically Significant Vital Sign Findings

Vital Sign	Category	Criteria for Potential Clinically Significant Vital Signs
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 mmHg from Baseline
	High	Value ≥ 120 mmHg and/or increase ≥ 15 mmHg from Baseline

Vital sign results meeting the criteria for PCS findings will be identified in a listing.

12.0 Summary of Changes

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

Under Section 10.2 Efficacy Analyses, there was a typo for the endpoint "Proportion of subjects with CDAI < 150 at Week 4 and SES-CD ≤ 4 and at least 2 point reduction versus baseline and no subscore greater than 1 in any individual variable at Week 56." It should be Week 56 for CDAI < 150, not Week 4. The endpoint has been updated in the SAP to "Proportion of subjects with CDAI < 150 at Week 56 and SES-CD ≤ 4 and at least 2 point reduction versus baseline and no subscore greater than 1 in any individual variable at Week 56."

Under Section 10.2 Efficacy Analyses, the following explanatory endpoints are added:

- Proportion of subjects who achieve endoscopic remission (SES-CD ≤ 4 and at least a 2 point reduction versus Induction Baseline and no subscore greater than 1 in any individual variable) at Week 56 among subjects with endoscopic response at Week 12.
- Proportion of subjects with endoscopic response, and ≥ 50% decrease from Induction Baseline in hs-CRP or fecal calprotectin at Week 56.
- Proportion of subjects with endoscopic remission, and ≥ 50% decrease from Induction Baseline in hs-CRP or fecal calprotectin at Week 56.

- Proportion of subjects with clinical remission, and ≥ 50% decrease from Induction Baseline in hs-CRP or fecal calprotectin at Week 56.

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

This is the first version of the SAP.

13.0 Appendix

14.0 References

1. Devlin NJ, Shah KK, Feng Y, et al. Valuing health-related quality of life: an EQ-5D-5L value set for England. Health Econ. 2018;27(1):7-22.

1.0 Title Page

Statistical Analysis Plan for the 12-Week Double-Blind Induction Study

Study M14-115

A Multicenter, Randomized, Double-Blind Study to Evaluate Higher Versus Standard Adalimumab Dosing Regimens for Induction and Maintenance Therapy in Subjects with Moderately to Severely Active Crohn's Disease and Evidence of Mucosal Ulceration

Date: 22 Feb 2019

Version 4.0

2.0 Table of Contents

1.0 Title Page 1
2.0 Table of Contents 2
3.0 Introduction 4
4.0 Study Objectives, Design and Procedures 4
4.1 Objectives 4
4.2 Study Design and Design Diagram 5
4.3 Sample Size 17
4.4 Interim Analysis 18
5.0 Analysis Populations 18
5.1 Definition for Analysis Populations 18
5.2 Variables Used for Stratification of Randomization 19
6.0 Analysis Conventions 19
7.0 Demographics, Baseline Characteristics, Medical History, and Previous/Concomitant Medications 32
7.1 Demographic and Baseline Characteristics 32
7.2 Medical History 34
7.3 Previous Treatment and Concomitant Medications 34
8.0 Patient Disposition 35
9.0 Study Drug Exposure and Compliance 36
10.0 Efficacy Analysis 37
10.1 General Considerations 37
10.2 Primary Efficacy Analyses 37
10.3 Secondary Efficacy Analyses 38
10.4 Handling of Multiplicity 43
10.5 Efficacy Subgroup Analysis 44
11.0 Safety Analysis 44
11.1 General Considerations 44
11.2 Analysis of Adverse Events 45
11.2.1 Treatment-Emergent Adverse Events 45
11.2.2 Adverse Events of Special Interest 47
11.3 Analysis of Laboratory Data 49

11.3.1 Assessment of Shifts from Baseline in Liver-Specific Laboratory Variables 50
11.3.2 Hy's Law Cases 51
11.4 Analysis of Vital Signs and Weight 51
12.0 Summary of Changes 53
12.1 Summary of Changes Between the Latest Version of Protocol and the Current SAP 53
12.2 Summary of Changes Between the Previous Version and the Current Version of the SAP 53
13.0 Appendix 54
14.0 References 54

List of Tables

Table 1. Study Activities 11
Table 2. Visit Windows for Analysis of Efficacy Variables (Except IBDQ, EQ-5D, WPAI, and SES CD), hs-CRP, Urinalysis and Vital Signs for Induction Study 21
Table 3. Visit Windows for Analysis of IBDQ, EQ-5D, and WPAI 22
Table 4. Visit Windows for Analysis of Chemistry and Hematology for Induction Study 22
Table 5. Visit Windows for Analysis of SES-CD, Fistulas, and EIM for Induction Study 23
Table 6. Visit Windows for Analysis of Fecal Calprotectin for Induction Study 23
Table 7. An EQ-5D-5L Value Set for England 31
Table 8. Criteria for Potentially Clinically Significant Vital Sign Findings 52

List of Figures

Figure 1. Study Schematic 7

3.0 Introduction

This statistical analysis plan (SAP) describes the statistical analysis to be completed by the AbbVie Global Statistics Department for study Protocol M14-115 Amendment 6. 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 forth version of the SAP for the 12-Week Double-Blind Induction Study of Protocol M14-115.

This analysis plan describes the primary and secondary efficacy analyses as well as the safety analysis for the 12-week double-blind induction study.

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 9.4 or later (SAS Institute Inc., Cary, NC 27513) under the UNIX operating system.

4.0 Study Objectives, Design and Procedures

4.1 Objectives

The primary objective of this study is to assess the efficacy and safety of two adalimumab induction regimens in achieving clinical remission (CDAI < 150) at Week 4 and endoscopic response defined as decrease in SES-CD > 50% from Baseline (or for a Baseline SES-CD of 4, at least a 2 point reduction from Baseline) at Week 12, in subjects with moderately to severely active Crohn's disease and evidence of mucosal ulceration at Baseline.

Additional objectives include:

- Assessing the efficacy and safety of two adalimumab induction regimens in reducing signs and symptoms of Crohn's disease at Week 12.
- Assessing the efficacy and safety of two adalimumab maintenance regimens in reducing signs and symptoms of Crohn's disease at Week 56.
- Assessing pharmacokinetics (PK) and immunogenicity of two adalimumab induction regimens following subcutaneous (SC) administration.

4.2 Study Design and Design Diagram

This is a randomized, double-blind multicenter study of two adalimumab induction and maintenance regimens in subjects with moderately to severely active CD with evidence of mucosal ulceration confirmed by central reading. No placebo arm is planned since there is well-documented efficacy of adalimumab in CD and because the purpose of this study is to achieve better efficacy than the standard induction and maintenance regimens in terms of clinical remission and endoscopic response. Additionally, since subjects will be required to have failed or been intolerant of standard therapies and have evidence of endoscopic damage confirmed by a central reader, it would be not medically acceptable to deny those subjects effective treatment or feasible to enroll subjects in the trial when adalimumab is available to be prescribed for CD.

Approximately 500 adult subjects with active Crohn's disease, defined as having a CDAI of ≥ 220 and ≤ 450 and evidence of mucosal ulceration defined as SES-CD ≥ 6, excluding the presence of narrowing component, or SES-CD ≥ 4, excluding the presence of narrowing component, for patients with disease limited to the ileum on screening endoscopy or endoscopy performed within 45 days before Baseline, confirmed by a centrally read endoscopy, will be enrolled at approximately 150 sites worldwide.

Subjects who meet all of the inclusion criteria and none of the exclusion criteria will be enrolled into the study and randomized in a 3:2 ratio at Baseline to receive a higher

induction adalimumab regimen or standard induction adalimumab regimen during the double-blind Induction Study.

The randomization of subjects for the Induction Study will be stratified by hs-CRP at Baseline (< 10 and ≥ 10 mg/L), using the Screening hs-CRP value, prior infliximab use, and Crohn's disease activity (CDAI ≤ 300, > 300), at Baseline. Enrollment of subjects with prior infliximab use will be limited to 25% of the total study population.

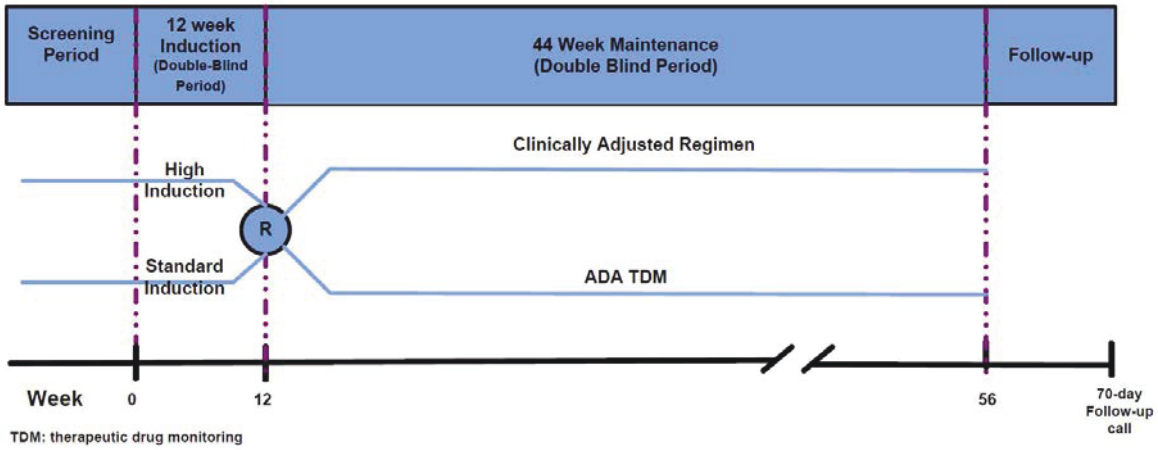
Subjects assigned to the higher induction regimen will receive blinded adalimumab 160 mg at Baseline, Week 1, Week 2, and Week 3. At Week 4, subjects will receive 40 mg every other week (eow) through Week 12. Subjects assigned to the standard adalimumab induction regimen will receive blinded adalimumab 160 mg at Baseline and matching placebo at Week 1. Subjects will receive 80 mg and matching placebo at Week 2, and then matching placebo at Week 3. At Week 4, subjects will receive 40 mg every other week through Week 12.

At Week 12, subjects will be re-randomized in a 1:1 ratio to one of the two double-blinded exploratory treatment regimens (adalimumab clinically adjusted [CA] regimen and adalimumab therapeutic drug monitoring [TDM] regimen). The re-randomization at Week 12 will be stratified by induction treatment regimen, clinical response (CR-70) status at Week 12 and decrease in SES-CD > 50% from Baseline per the site investigator reading at Week 12. Among subjects achieving decrease in SES-CD > 50% from Baseline at Week 12, the randomization will be further stratified by achievement of an SES-CD ≤ 4 and at least a 2 point reduction versus Baseline and no subscore greater than 1 in any individual variable using the Week 12 SES-CD value provided by the site. No study drug will be administered or injected at the final visit.

Further details on the maintenance treatment regimen are provided in the protocol Section 5.1 Overall Study Design and Plan.

A schematic of the study design is presented in Figure 1.

Figure 1. Study Schematic



Clinical evaluation will be at visits occurring at Baseline, Weeks 2, 4, 6, 8, 12, 14, 20, 26, 28, 34, 40, 42, 48, and 56/PD. An electronic diary will be dispensed at the Screening visit. In addition to routine physical examination, CDAI calculation, diary review, laboratory, adverse event, concomitant medication and vital sign assessments, the following will be collected:

- Results of study questionnaires (inflammatory bowel disease questionnaire [IBDQ], European Quality of Life 5 dimensions [EQ-5D], work productivity and impairment [WPAI]) at Baseline, Week 4, Week 8, Week 12, Week 26, Week 40 and Week 56/PD.
- Calculation of the SFPS at Baseline, Week 2, Week 4, Week 6, Week 8, Week 12, Week 14, Week 20, Week 26, Week 28, Week 34, Week 40, Week 42, Week 48 and Week 56/PD. The Screening results will serve as the Baseline value.
- Results of daily Bristol Stool Form Scale beginning at Baseline through Week 56/PD.
- Results of 11-point Abdominal Pain Rating Scale beginning at Baseline, Week 2, Week 4, Week 6, Week 8, Week 12, Week 26, Week 40, and Week 56/PD.

- Serum for measurement of adalimumab concentrations just prior to dosing at Baseline, Week 2, Week 4, Week 6, Week 8, Week 12, Week 26, Week 40, and Week 56/PD.
- Serum for measurement of Anti-Adalimumab Antibodies (AAA) just prior to dosing at Baseline, Week 4, Week 12, Week 26, Week 40, and Week 56/PD.
- Serum for measurement of infliximab serum levels and Human Anti-Chimeric Antibodies (HACA) just prior to dosing at Baseline.
- Serological biomarkers/mRNA at Baseline, Week 2, Week 4, Week 8, Week 12, Week 26, Week 40, and Week 56/PD.
- Stool samples for analysis of fecal calprotectin during Screening, at Week 4, Week 12, Week 26, and Week 56/PD.
- Stool samples for microbiota metagenomic analyses during Screening, at Week 4, Week 12, Week 26, and Week 56/PD. The stool samples should be taken before starting bowel preparations for endoscopy.
- Endoscopic evaluations, eligibility confirmed by central reader, will be done at Screening, Week 12, and Week 56/PD.
- An optional pharmacogenetic sample should be drawn at Baseline and Week 12.

Throughout the study, subjects will only be allowed to change the dosage of CD-specific concomitant medications as specified below:

- At Week 4, subjects who are taking corticosteroid therapy at Baseline will have their corticosteroid therapy tapered according to a tapering schedule specified in the clinical study protocol. If the Investigator feels that the steroid taper is not advisable for a particular subject at Week 4, the Study Designated Physician (SDP) should be consulted for evaluation and approval.
- Subjects taking corticosteroids at Baseline who have a loss of satisfactory clinical response per the Investigator's judgment after the steroid taper has been initiated may have their corticosteroid dose increased per the Investigator's discretion during the study. Subjects in whom the maximum steroid dose equivalent exceeds the dose used at Baseline will be censored for efficacy assessments (i.e., will be considered non-responders for categorical

endpoints and will have last non-missing values carried forward for non-categorical assessments) from that point forward. These subjects will continue to be evaluated in the safety population.

- Immunosuppressant doses may be decreased or terminated in the event of moderate-to-severe treatment-related toxicities.

The study was designed to enroll 500 subjects to meet scientific and regulatory objectives without enrolling an undue number of subjects in alignment with ethical considerations. Therefore, if the target number of subjects has been enrolled, there is a possibility that additional subjects in screening will not be enrolled.

Re-Screening

Subjects that initially screen fail for the study may be permitted to re-screen following re-consent. All screening procedures with the possible exceptions noted below will be repeated. The subject must meet all 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 In-Tube test or T SPOT TB test) chest x-ray (if applicable) and electrocardiogram (ECG), these tests will not be required to be repeated for re-screening provided the conditions noted in Protocol Section 5.3.1.1 are met and no more than 90 days have passed. If an endoscopy was performed, this will not be required to be repeated for re-screening provided the conditions noted in protocol Section 5.3.1 are met and no more than 45 days have passed between the endoscopy date and the Baseline date. As appropriate, sites are encouraged to contact the AbbVie Study Designated Physician (SDP) to confirm if subjects should or should not be re-screened.

Subjects may discontinue adalimumab treatment at any time during study participation. Subjects that end study participation early will have a Premature Discontinuation Visit. All subjects who discontinue from the study prematurely will have a follow-up phone call

approximately 70 days after the last administration of study drug to obtain information on any new or ongoing adverse events (AEs). The 70-day follow-up phone call will not be required for any subject that initiates commercial adalimumab.

The study activities are presented in Table 1.

Table 1. Study Activities

Activity	Screening Period (30 Days)[a]	12-Week Double-Blind Induction Study						44-Week Double-Blind Maintenance Study									Unscheduled Visit[v]	70-Day Follow-Up[t]
	Screening	Baseline	Week 2	Week 4	Week 6	Week 8	Week 12 (Re-Randomization)	Week 14	Week 20	Week 26	Week 28	Week 34	Week 40	Week 42	Week 48	Week 56/ Premature Discontinuation		
Informed Consent	X																	
Inclusion/Exclusion[b]	X	X																
Medical/Surgery History[b]	X	X																
Previous and Concomitant Medication[b]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Vital Signs[c]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Endoscopy[d]	X						X									X		
Physical Examination[e]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
TB Screening[f]	X																	
Chest X-Ray[g]	X																	
ECG[h]	X																	
Chemistry and Hematology[i]	X	X	X	X		X	X			X			X			X	X	
Urinalysis[i,j]	X	X	X	X	X	X	X			X			X			X	X	
Pregnancy Tests[k]	X[k]	X					X									X		
Hepatitis B Screen and HIV[l]	X																	

Table 1. Study Activities (Continued)

Activity	Screening Period (30 Days)[a]	12-Week Double-Blind Induction Study						44-Week Double-Blind Maintenance Study									Unscheduled Visit[v]	70-Day Follow-Up[t]
	Screening	Baseline	Week 2	Week 4	Week 6	Week 8	Week 12 (Re-Randomization)	Week 14	Week 20	Week 26	Week 28	Week 34	Week 40	Week 42	Week 48	Week 56/ Premature Discontinuation		
hs-CRP	X		X	X	X	X	X			X			X			X	X	
C. difficile toxin	X																	
Antinuclear antibody (ANA)/Anti-double-stranded DNA (anti dsDNA)[m]	X																	
Stool Sample (microbiota metagenomic analyses and fecal calprotectin)[n]	X			X			X			X			X			X		
Human Antichimeric Antibodies (HACA)/Infliximab Concentrations		X																
Adalimumab Concentration[o]		X	X	X	X	X	X			X			X			X	X	
AAA Concentration[o]		X		X			X			X			X			X	X	
Pharmacogenetic Marker[p]		X					X											
Serological Biomarkers/mRNA		X	X	X		X	X			X			X			X		

Table 1. Study Activities (Continued)

Activity	Screening Period (30 Days)[a]	12-Week Double-Blind Induction Study						44-Week Double-Blind Maintenance Study									Unscheduled Visit[v]	70-Day Follow-Up[t]
	Screening	Baseline	Week 2	Week 4	Week 6	Week 8	Week 12 (Re-Randomization)	Week 14	Week 20	Week 26	Week 28	Week 34	Week 40	Week 42	Week 48	Week 56/ Premature Discontinuation		
Crohn's Disease Activity Index (CDAI) CDAI components "Number of liquid or very soft stools" and "Abdominal pain" (Stool [liquid/soft] Frequency + Abdominal Pain Score; SFPS)	X	X[u]	X	X	X	X	X	X	X	X	X	X	X	X	X	X		
SES-CD Score	X						X									X		
Inflammatory Bowel disease Questionnaire (IBDQ)		X		X		X	X			X			X			X		
European Quality of Life 5 dimensions (EQ‑5D)		X		X		X	X			X			X			X		
Work Productivity and Impairment Questionnaire (WPAI)		X		X		X	X			X			X			X		
Abdominal Pain Rating Scale		X	X	X	X	X	X			X			X			X		
Corticosteroid Taper[q]				X	X	X	X	X										
Monitor Adverse Events[r]		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Study Drug Dispensing/Administration[s]		X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	

Table 1. Study Activities (Continued)

Activity	Screening Period (30 Days)[a]	12-Week Double-Blind Induction Study						44-Week Double-Blind Maintenance Study									Unscheduled Visit[v]	70-Day Follow-Up[t]
	Screening	Baseline	Week 2	Week 4	Week 6	Week 8	Week 12 (Re-Randomization)	Week 14	Week 20	Week 26	Week 28	Week 34	Week 40	Week 42	Week 48	Week 56/ Premature Discontinuation		
Dispense Subject Diary	X																	
Subject Diary Review		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Intestinal Biopsies[p]	X						X									X		

a. The Screening period will be a minimum of 7 days for CDAI calculation. The CDAI calculated at Screening will serve as the Baseline CDAI. Baseline visit date will serve as the reference for all subsequent visits. A ± 3-day window is permitted around all study visits.

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

c. Height will be measured at Screening only (with shoes off, and then adding 1 inch or 2.5 cm).

d. An ileocolonoscopy will be performed during Screening or one performed within 45 days before Baseline, Week 12, and at Week 56/PD. A biopsy will be performed at Screening, Week 12, and Week 56/PD if a suspicious lesion or suspected malignancy, in the assessment of the Investigator, is observed, and evaluated by the local pathologist. Site staff should schedule the Week 12 Endoscopy during the Baseline visit, if possible. Appropriate documentation of biopsy results consistent with the diagnosis of CD, in the assessment of the Investigator, must be available in order to confirm the subject's eligibility for the study. If this documentation is not available, a diagnostic biopsy from the most affected observed area of the ileum/colon must be performed during the Screening endoscopy and evaluated by a qualified local pathologist and the results reviewed by the Investigator. Biopsies to evaluate suspicious lesions and to rule out malignancy may be taken during any study endoscopy per the Investigator's discretion and evaluated by the local pathologist.

Table 1. Study Activities (Continued)

e. Physical examination performed at Screening, Week 12 and Week 56/Premature Discontinuation 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 at all other visits are symptom based and must include a count of the number of cutaneous fistulas.

f. Subjects with negative PPD test and/or QuantiFERON-TB Gold test within 90 days of Screening will not require a repeat skin test, if documentation is available. PPD skin test is to be read 48 to 72 hours after placement.

g. Chest x-ray includes posterior-anterior (PA) **and** lateral views. Subjects with normal chest x-ray within 3 months of Screening would not require a repeat chest x-ray, if documentation is available.

h. Subjects with normal ECG within 90 days of Screening would not require a repeat ECG, if documentation is available.

i. Lab assessments will only need to be repeated at Baseline if the time between Screening and Baseline is greater than 14 days, or if the subject's health status has changed to warrant a repeat test.

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. Further explanations of these tests are provided in the laboratory manual.

k. Serum pregnancy test will be performed on all women of childbearing potential at Screening. Urine pregnancy test will be performed locally at Baseline Visit, and at the Week 12 and Week 56/PD for all women of child-bearing potential. The frequency can be increased up to every visit as per local regulations. If any urine pregnancy test is positive, a serum pregnancy test will be performed by the central laboratory.

l. Subjects will be tested for the presence of the Hepatitis B Virus (HBV) at Screening. A positive result for the Hepatitis B surface antigen (HBs Ag) will be exclusionary. Please refer to hepatitis procedures for details on testing requirements. If required by country regulatory authorities to confirm eligibility, subjects will be tested for HIV and documented that the test has been performed. This testing is to be done at a local laboratory. A subject will not be eligible for study participation if test results indicate a positive HIV infection. AbbVie will not receive results from the testing and not be made aware of any positive result.

m. Anti-dsDNA performed if ANA result is positive.

n. If a sample cannot be obtained during the site visit, the site will give instructions and a stool sample supply kit (supplies will be provided at the time points indicated). The stool from which these samples are prepared should be scored using the Bristol stool chart by ePRO. All stool samples for metagenomic analysis should be collected before any bowel preparation for endoscopy is started and should be returned to the site within 3 days of collection.

   Note: The Screening stool sample may be taken anytime during the Screening period but should be collected prior to any bowel prep.

o. Blood samples for the measurement of adalimumab and AAA concentrations will be collected prior to dosing. Testing of the adalimumab and AAA concentrations must not be performed locally. All pharmacokinetic results will remain blinded to the investigator, study site personnel and the subject throughout the study.

Table 1. Study Activities (Continued)

p. Pharmacogenetic Marker and intestinal biopsy samples are optional. Separate consents must be signed prior to the sample draw. If the pharmacogenetic sample is not collected at Baseline, preferably it should be collected at the next study visit.

q. Subject will begin corticosteroid taper at Week 4. If the Investigator feels that the steroid taper is not advisable for a particular subject at Week 4, the SDP should be consulted for evaluation and approval.

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

s. Administration of drug will be performed after all assessments and examinations scheduled for that day have been completed.

t. Subjects will be contacted 70 days following study drug discontinuation for an assessment of any new or ongoing AEs, except those subjects that continue on adalimumab therapy after the end of study participation.

u. The Screening results of the SFPS will serve as the Baseline value.

v. Visits for dispensing new study drug in case of temperature excursion, loss or damage are not considered an Unscheduled Visit. In addition, visits to retest a lab will not be considered an Unscheduled Visit. Unscheduled Visits according to this table are for purposes when the subject is coming in for a visit for evaluation, assessment and potential dose escalation.

4.3 Sample Size

The calculated Week 4 remission rate for higher dose regimen are based on the average of the PK-PD Linear model and E_{max} model and is 56% [(70% [Linear] + 42% [E_{max}]) / 2] for infliximab naïve subjects and 45% [(60% [Linear] + 30% [E_{max}]) / 2] for infliximab experienced subjects. Assuming the ratio in the proposed study of infliximab naïve subjects versus infliximab experienced is 75% versus 25%, respectively, the weighted Week 4 remission rate will be approximately 50% for higher induction dose regimen. For the standard dose regimen, the Week 4 clinical remission rate will be approximately 30%. Subjects who meet all of the inclusion criteria and none of the exclusion criteria will be enrolled into the study and randomized in a 3:2 ratio at Baseline to receive a higher induction adalimumab regimen or standard induction adalimumab regimen during the double-blind period. A sample size of 500 (300 higher dose and 200 standard dose) will be adequate to detect at least a 20% treatment difference in clinical remission rates at Week 4 between the dosing regimens using Fisher's exact test with 99% power at a 0.05 two-sided significant level.

Since exposure/endoscopic relationships are not currently available for adalimumab, clinical remission will be based on PK/PD modeling which predicts an expected ~50% increase in the proportion of subjects with endoscopic response compared to the standard adalimumab induction regimen. The observed rate of endoscopic response (decrease > 50% SES-CD from Baseline [or for a Baseline SES-CD of 4, at least a 2 point reduction from Baseline]) in Study M05-769 was 44% at Week 12 with the currently approved dosing regimen. Thus, by extrapolation, the expected rate for the higher induction regimen is 66%. A sample size of 500 subjects (300 higher induction regimen and 200 standard induction regimen) will be adequate to detect at least a 22% treatment difference in endoscopic response (decrease > 50% SES-CD from Baseline [or for a Baseline SES-CD of 4, at least a 2 point reduction from Baseline]) rates at Week 12 between the dosing regimens using Fisher's exact test with approximately 99% power at a 0.05 two-sided significance level.

Since the clinical remission at Week 4 endpoint and the endoscopic response at Week 12 endpoint are likely not independent of each other, the power for the co-primary endpoints for the study is expected to be > 98%.

4.4 Interim Analysis

An interim analysis of the primary endpoint and ranked secondary efficacy variables for the Induction Study only as well as safety data collected from Baseline through double - blind Week 12 may be performed after the last subject in ITT population completes the 12-week double-blind Induction Study. A database cut will be performed and any discrepant data will be clarified before the lock. Since this interim analysis is the only and final analysis of the co-primary efficacy endpoints of the Induction Study, multiplicity adjustment is deemed not necessary.

5.0 Analysis Populations

5.1 Definition for Analysis Populations

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

Intent-to-Treat (ITT) set includes all subjects who are randomized. ITT subjects will be analyzed as randomized. ITT set is the primary population for the efficacy analysis during the induction double-blind period.

If there is > 10% difference in number of subjects with actual treatment received and the treatment randomized, a sensitivity analysis for the co-primary endpoints based on the actual treatment received will be performed.

The safety set consists of all subjects who received at least one injection of study drug. The safety set will be analyzed as treated, according to treatment the subject actually received. The safety set will be used only for safety analysis.

5.2 Variables Used for Stratification of Randomization

At Baseline, subjects will be randomized 3:2 to one of two double-blinded adalimumab Induction Study regimens (higher induction regimen or standard induction regimen). The randomization will be stratified by hs-CRP at Baseline (< 10 and ≥ 10 mg/L) using Screening hs-CRP values, prior infliximab use (Yes, No), Crohn's disease severity (CDAI ≤ 300, > 300) at Baseline.

At Week 12, subjects will be re-randomized in a 1:1 ratio to one of two double-blinded exploratory treatment regimens (clinically adjusted [CA] regimen and adalimumab therapeutic drug monitoring [TDM] regimen). The re-randomization at Week 12 will be stratified by induction treatment regimen, clinical response (CR-70) status at Week 12, and decrease in SES-CD > 50% from Baseline per the site investigator reading at Week 12. Among subjects achieving decrease in SES-CD > 50% from Baseline at Week 12, the randomization will be further stratified by achievement of an SES-CD ≤ 4 and at least a 2 point reduction versus baseline and no subscore greater than 1 in any individual variable using the Week 12 SES-CD value provided by the site.

The randomization schedule will be prepared by the Statistics Department of AbbVie.

6.0 Analysis Conventions

Definition of Induction Baseline

The Induction Baseline visit date is the date when the first dose of study drug is received and referred to as Day 1 or Week 0. If a subject has not taken any study drug during Induction Study, the randomization date will be used as Day 1. 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 Induction Final Observation

Final Observation in the 12-week double-blind Induction Study is defined as the last non-missing observation collected within 70 days following the last dose of study drug for

subjects who are not re-randomized nor enrolled in Study M14-347, and on or before the day of the first dose of study drug in the 44-week double-blind Maintenance Study for subjects who entered the Maintenance Study or the day of the first dose of study drug in Study M14-347 if the subjects are enrolled in Study M14-347. Of note, the efficacy, laboratory, and vital sign evaluations performed on the day of first dose of study drug in the double-blind Maintenance Study or Study M14-347 will be included in the efficacy and safety analyses for the Induction Study. AEs with onset on the date of the first dose of the Maintenance Study/M14-347 will be attributed to the Maintenance Study/M14-347.

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, 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 later time one will be used if there is time or 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. 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 IBDQ, EQ-5D, WPAI, and SES CD), hs-CRP, Urinalysis and Vital Signs for Induction Study

Scheduled Week	Nominal Day	Time Window (Rx Day Range)
Week 0	1[a]	≤ 1
Week 2	15	2 – 22
Week 4	29	23 – 36
Week 6	43	37 – 50
Week 8	57	51 – 71
Week 12	85	72 – 98[b]

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

a. Day of first dose of study drug in double-blind induction study. If a subject has not taken any study drug during Induction Study, the randomization date will be use as Day 1.

b. Rx Day 98, or first dose of maintenance, or first dose of study drug in Study M14-347, whichever earlier.

Table 3. Visit Windows for Analysis of IBDQ, EQ-5D, and WPAI

Scheduled Week	Nominal Day	Time Window (Rx Day Range)
Week 0	1[a]	≤ 1
Week 4	29	2 – 43
Week 8	57	44 – 71
Week 12	85	72 – 98[b]

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

a. Day of first dose of study drug in double-blind induction study. If a subject has not taken any study drug during Induction Study, the randomization date will be use as Day 1.

b. Rx Day 98, or first dose of maintenance, or first dose of study drug in Study M14-347, whichever earlier.

Table 4. Visit Windows for Analysis of Chemistry and Hematology for Induction Study

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 – 98[b]

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

a. Day of first dose of study drug in double-blind induction study. If a subject has not taken any study drug during Induction Study, the randomization date will be use as Day 1.

b. Rx Day 98, or first dose of maintenance, or first dose of study drug in Study M14-347, whichever earlier.

Table 5. Visit Windows for Analysis of SES-CD, Fistulas, and EIM for Induction Study

Scheduled Week	Nominal Day	Time Window (Rx Day Range)
Week 0	1[a]	≤ 1
Week 12	85	2 – 168[b]

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

a. Day of first dose of study drug in double-blind induction study. If a subject has not taken any study drug during Induction Study, the randomization date will be use as Day 1.

b. Rx Day 168, or first dose of maintenance + 7 days, or first dose of study drug in Study M14-347 + 7 days, whichever earlier. No + 7 days for Fistulas and EIM.

Table 6. Visit Windows for Analysis of Fecal Calprotectin for Induction Study

Scheduled Week	Nominal Day	Time Window (Rx Day Range)
Week 0	1[a]	≤ 1
Week 4	29	2 – 57
Week 12	85	58 – 112[b]

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

a. Day of first dose of study drug in double-blind induction study. If a subject has not taken any study drug during Induction Study, the randomization date will be use as Day 1.

b. Rx Day 112, or first dose of maintenance, or first dose of study drug in Study M14-347, whichever earlier.

CD-Related Concomitant Corticosteroid Therapy

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 last non-missing 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.

Mixed-Effect Model Repeated Measure (MMRM)

The MMRM model will be used for continuous efficacy variables with longitudinal data as a sensitivity analysis. The MMRM model includes the baseline values as covariate; randomization strata, region (US versus non-US), 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.

Observed Case (OC)

Observed case analysis will be performed such that missing values will not be imputed. Subjects in whom the following CD-related systemic or rectal corticosteroids that were not being taken at Baseline and are 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 the observed non-categorical assessments will be replaced by last non-missing values prior to censoring) from that point through the end of the study.

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

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.

Sensitivity analyses will be performed to make use of all available data based on the above rule if more than 5% of subjects have less than 4 days of diary data.

Definition of Clinical Remission

CDAI < 150.

Definition of Clinical Response

Decrease in CDAI ≥ 70 points from Baseline.

Definition of Enhanced Clinical Response

Decrease in CDAI ≥ 100 points from Baseline.

Definition of Endoscopic Remission

Simplified Endoscopic Score for Crohn's Disease (SES-CD) ≤ 4 and at least 2 point reduction versus baseline and no subscore greater than 1 in any individual variable, as scored by central reviewer.

Definition of Endoscopic Response

Decrease in SES-CD > 50% from Baseline (or for a Baseline SES-CD of 4, at least a 2 point reduction from Baseline), as scored by central reviewer.

Stool Frequency + Abdominal Pain Score (SFPS)

SFPS is the CDAI components "Number of liquid or very soft stools" and "Abdominal pain" (Stool [liquid/soft] Frequency + Abdominal Pain Score). The sub-total score for each component is calculated as the sum of the 7 day diary score multiplied by the corresponding multiplier.

Definition of Time to Event (Clinical Remission, Clinical Response, or Hospitalization)

Time to event is defined as number of days from Study Day 1 to the date of first occurrence of an event. It is calculated as (date of first occurrence of Event – Study Day 1 date + 1). Subjects who discontinue prior to experiencing the event are censored at the time of discontinuation. Subjects who are not experiencing the event by Week 12 are censored at the Week 12 visit.

Definition of Symptomatic Remission

Average daily stool frequency ≤ 2.8 (and not worse than baseline) and average daily abdominal pain ≤ 1.0 (and not worse than baseline). The analysis will be only performed on subjects with baseline average daily stool frequency score ≥ 4.0 or average daily abdominal pain score ≥ 2.0.

Definition of Symptomatic Response

Average daily stool frequency at least 30% reduction from baseline and average daily abdominal pain not worse than baseline or average daily abdominal pain at least 30% reduction from baseline and average daily stool frequency not worse than baseline. The analysis will be only performed on subjects with baseline average daily stool frequency score ≥ 4.0 or average daily abdominal pain score ≥ 2.0.

Discontinued Corticosteroid Use at Each Visit during Induction Among Subjects Who Used Corticosteroids at Baseline

Discontinued steroid use at each visit is based on the steroid use on the date of CDAI measurement for that visit. If the subject has no CDAI for the visit, the date of vital signs will be used. If no vital signs date, then the steroid use on the nominal day for that visit will be evaluated.

If a subject discontinued the study, 'discontinued corticosteroid use' will be classified as No from that point through the end of the study.

If a subject met censoring criteria for CD-related concomitant therapy, 'discontinued steroid use' will be classified as No from that point through the end of the study.

SES-CD Scoring

All videotaped colonoscopies will undergo central review providing the investigator's SES-CD score meets the entry criteria. Two primary central reviewers will evaluate the videotaped colonoscopies separately and provide their SES-CD scores to Parexel. A third central reviewer will adjudicate between the two initial reviewers' SES-CD scores if there is discrepancy in any SES-CD variable. The adjudicator will select the final SES-CD score that he/she most agrees with from those provided by the two primary reviewers, 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 two primary central reviewers, that 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 that serves as the SES-CD for the study's efficacy analyses, the following imputation rules will be applied:

The missing SES-CD individual variables at Baseline will be imputed as 0. The post-Baseline missing SES-CD individual variables will be imputed as 0 if there are 8 or fewer missing individual variables. If more than 8 individual variables are missing, the total SES-CD will be missing for that visit.

Bristol Stool Chart Score

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. Bristol Stool Form Scale is not a continuous variable. The new variable Bristol Stool Chart score is defined 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 score will not be calculated and will be considered missing.

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

Hs-CRP

From November 2014 to December 2015, Roche hs-CRP kit lot 604450 was in use at ICON. In December 2015 ICON altered the study team of a bias in the lot of the Roche Diagnostic C-creative protein high sensitive reagent. ICON then informed the study team of the extent of the under recovery in March 2016. This lot was "retrospectively expired" as the low standard beads agglutinated causing specimen results that are below 5 mg/L to be up to 25% lower than the accurate value. ICON later determined with Roche that the affected samples were those tested from May 18, 2015 to December 2015.

In the Study M14-115 approximately 423 samples were affected according to ICON. AbbVie obtained the final listing of affected samples/results from ICON on August 10, 2016. A correction factor developed by Roche and ICON will be applied to all affected specimens with a concentration between 0.51 to 4.99 mg/L and will replace the initially provided hs-CRP values in the study database.

For Study M14-115, both Roche and Abbott platform are used to measure hs-CRP. The study team has decided to use Abbott platform for CSR analysis. If hs-CRP was measured by Roche platform only, the value will be converted to Abbott equivalent by a conversion factor.

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 7).

Table 7. 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)
- Respiratory Rate (RPM)
- Temperature (°C)
- IBDQ score
- CDAI
- SFPS
- Total SES-CD score

- Crohn's Disease Duration (years)
- hs-CRP mg/L
- WPAI and its components
- EQ-5D
- Abdominal Pain Rating Scale
- Bristol Stool Chart score
- Fecal calprotectin

Categorical Variables:

- Sex (male, female)
- Race
- Ethnicity
- Age (≤ median, > median)
- Age (< 40, 40 to < 65, 65 to < 75, ≥ 75)
- Baseline fecal calprotectin (≤ median, > median)
- 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)
- Baseline fecal calprotectin (≤ median, > median)
- Baseline fecal calprotectin (≤ 250 µg/g, > 250 µg/g)
- Crohn's disease severity (CDAI ≤ 300, > 300) at Baseline
- Baseline CDAI (≤ median, > median)
- Baseline SES-CD (≤ median, > median)
- Prior infliximab use (or prior anti-TNF use for subjects randomized under original protocol) (yes, no)
- Weight (≤ median, > median)
- Baseline albumin (≤ median, > median)
- Disease duration (≤ median, > median)

- 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, will be obtained from each subject during the Screening period. Medical history will be summarized using body system and preferred term by treatment group. No statistical tests will be performed.

Chest X-Ray Results: All subjects undergo a standard chest 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 or 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. No statistical tests will be performed.

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

Prior medications are those medications taken prior to the first dose of study drug. This includes medications with a start date before the first study drug administration date,

regardless of the end date of these medications. Medications taken on the day of the first dose of study drug are not counted as prior medications.

Concomitant medications are those medications, other than study drug, taken after the first dose of study drug. For this interim analysis, concomitant medications include medications with a start date prior to the Baseline date which are continuing after Baseline and all medications with a start date between the Baseline date and last study drug administration + 14 days or until first dose of maintenance study drug/Study M14-347 study drug, whichever is earlier.

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 [0, ≤ 10 mg, ≤ 20 mg, or ≤ 30 mg prednisone equivalents], aminosalicylates, immunosuppressants, and antibiotics at the Baseline will be tabulated. In addition, the number and percent of subjects using any Crohn's disease specific medications including infliximab 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 12-week Induction study
- Subject disposition including the number and percent of subjects who prematurely discontinued the Induction study (on or before Week 12) 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 (total patient years) will be summarized by treatment group.

The number of injections received at each schedule time point will be summarized with number and percent for each treatment group.

Study Drug Exposure (in Days) during the induction period:

For subjects who discontinued from the Induction Study:

- Date of last dose of Induction Study drug – Date of first dose of Induction Study drug + 14 days

For subjects who completed the Induction Study:

- Date of first dose of Maintenance Study drug or date of first dose study drug in Study M14-347 – Date of first dose of Induction Study drug + 1 day,
  Or
- Date of last dose of Induction Study drug + 14 – Date of first dose of Induction Study drug + 1 day, whichever is shorter.

The date of last dose of Induction Study drug is defined as the date of last study drug injection before the date of re-randomization into Maintenance Study. The date of first dose of Maintenance Study drug is defined as the date of first study drug injection on or after the date of re-randomization into Maintenance Study.

Study Drug Compliance:

Treatment compliance will be summarized for each treatment group in the ITT population. The treatment compliance is defined as the number of injections actually

taken by the subject divided by the number of injections planned to be taken by the subject during the induction period.

10.0 Efficacy Analysis

10.1 General Considerations

All statistical tests will be two-sided with the significance level of 0.05. 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 Primary Efficacy Analyses

This section provides the details of the primary efficacy analysis for the Induction Study.

Co-Primary Efficacy Variables:

- Proportion of subjects who achieve clinical remission at Week 4 (CDAI < 150).
- Proportion of subjects with decrease in SES-CD > 50% from Baseline (or for a Baseline SES-CD of 4, at least a 2 point reduction from Baseline) at Week 12.

Analysis Data Set for the Primary Efficacy Analysis:

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

Imputation Method Used for the Primary Efficacy Analysis:

Missing CDAI at Week 4 or SES-CD at Week 12 will be imputed using the non-responder imputation (NRI) approach.

Statistical Method of the Primary Efficacy Analysis:

The comparisons between treatment groups on the two co-primary efficacy variables will be performed using the Cochran-Mantel-Haenszel (CMH) adjusted for hs-CRP at Baseline (< 10 and ≥ 10 mg/L), prior infliximab use (or prior anti-TNF use for subjects randomized under original protocol), and Crohn's disease severity (CDAI ≤ 300, > 300) at Baseline. A CMH based two-sided 95% confidence interval for the difference between treatment groups will be calculated. Breslow-Day test for homogeneity of odds ratios will be performed as well. An exact CMH test will be used to handle zero cells.

For Week 4 clinical remission (CDAI < 150), LOCF, OC and multiple imputation methods will be used as sensitivity analyses. For Week 12 endoscopic response, OC will be used as sensitivity analysis.

In addition, as sensitivity analyses, the logistic regression including treatment, the randomization stratification factors and additional clinically important factors such as corticosteroid use at Baseline, IMM use at Baseline, Baseline fecal calprotectin, Baseline SES-CD, and Region (US versus ex-US) will also be performed for the co-primary endpoints.

The co-primary endpoints will be tested prior to the ranked secondary endpoints.

10.3 Secondary Efficacy Analyses

For the analysis purpose, the secondary efficacy endpoints for the Induction Study are divided into two groups: (1) ranked secondary endpoints and (2) additional, non-ranked endpoints.

Ranked Secondary endpoints:

1. Proportion of subjects with sustained clinical remission (CDAI < 150) at both Weeks 4 and 12.

2. Proportion of subjects with CDAI < 150 at Week 4 and decrease in SES-CD > 50% from Baseline (or for a Baseline SES-CD of 4, at least a 2 point reduction from Baseline) at Week 12.
3. Proportion of subjects with clinical remission (CDAI < 150) at Week 12.
4. Proportion of subjects who discontinued corticosteroid use and achieved clinical remission (CDAI < 150) at Week 12 among subjects taking corticosteroids at Baseline.
5. Proportion of subjects with endoscopic remission (SES-CD ≤ 4 and at least a 2 point reduction versus baseline and no subscore greater than 1 in any individual variable) at Week 12.
6. Change from Baseline in fecal calprotectin level at Week 4.
7. Proportion of subjects with hs-CRP < 5 mg/L and fecal calprotectin < 250 μg/g at Week 4.
8. Proportion of subjects with CDAI < 150, hs-CRP < 5 mg/L, and fecal calprotectin < 250 μg/g at Week 4.
9. Proportion of subjects with CDAI < 150, hs-CRP < 5 mg/L, SES-CD ≤ 4 and at least a 2 point reduction versus baseline and no subscore greater than 1 in any individual variable, and fecal calprotectin < 250 μg/g at Week 12.
10. Proportion of subjects who achieve an SES-CED ≤ 2 at Week 12.
11. Proportion of subjects with clinical response (decrease in CDAI ≥ 70 points from Baseline) at Week 4.
12. Proportion of subjects with clinical response (decrease in CDAI ≥ 70 points from Baseline) at Week 12.
13. Proportion of subjects achieving response in IBDQ Bowel Symptom domain (increase of IBDQ bowel symptom domain score ≥ 8) at Week 4.
14. Proportion of subjects achieving response in IBDQ Bowel Symptom domain (increase of IBDQ bowel symptom domain score ≥ 8) at Week 12.

15. Proportion of subjects achieving response in IBDQ fatigue item (increase of IBDQ fatigue item score ≥ 1) at Week 12.

For these endpoints which are categorical variables, NRI method of imputation will be used for the missing values. LOCF will be used as a sensitivity analysis wherever applicable.

Both LOCF and OC analyses will be performed for continuous endpoints. The LOCF analysis is considered primary for inferential purposes. In additional, Mixed-Effect Model Repeated Measure (MMRM) will be applied as a sensitivity analysis, wherever appropriate, for the longitudinal continuous endpoints.

For endoscopy related endpoints, there will be no LOCF imputation as there is only one post-baseline endoscopy during Induction. Only OC will be used as a sensitivity analysis.

The difference in proportions of subjects between treatment groups will be analyzed using the two-sided Cochran-Mantel-Haenszel (CMH) test adjusted for hs-CRP at Baseline (< 10 and ≥ 10 mg/L), prior infliximab use (or prior anti-TNF use for subjects randomized under original protocol), and Crohn's disease severity (CDAI ≤ 300, > 300) at Baseline. Additionally, the CMH-based two-sided 95% confidence interval (CI) for the difference in the proportions between the treatment groups will be calculated. An exact CMH test will be used to handle zero cells.

The difference in change from Baseline between treatment groups will be analyzed using an ANCOVA model including factors of treatment, hs-CRP at Baseline, prior infliximab use (or prior anti-TNF use for subjects randomized under original protocol), and Crohn's disease severity at Baseline, and Baseline values. Parameter estimates with 95% confidence interval and *P*-value will be provided.

All additional efficacy endpoints will be non-ranked.

- Proportion of subjects with clinical response (decrease CDAI ≥ 70 points from Baseline) at each scheduled visit in Induction Study.

- Proportion of subjects with enhanced clinical response (decrease CDAI ≥ 100 points from Baseline) at each scheduled visit in Induction Study.
- Proportion of subjects who discontinue corticosteroid use at each scheduled visit in Induction Study among subject taking corticosteroids at Baseline.
- Proportion of subjects who achieve a composite subtotal score of SFPS < 50 at Week 12 who had an SFPS ≥ 100 at Baseline.
- Proportion of subjects who achieve a composite subtotal score of SFPS < 50 at Week 4 who had an SFPS ≥ 100 at Baseline.
- Proportion of subjects who achieve SES-CD ≤ 3 and at least a 2 point reduction versus baseline and no subscore greater than 1 in any individual variable at Week 12.
- Proportion of subjects with SES-CD = 0 at Week 12.
- Change from Baseline in fecal calprotectin level at each scheduled visit in Induction Study.
- Change from Baseline in hs-CRP at each scheduled visit in Induction Study.
- Change in IBDQ total score and individual IBDQ domain scores (bowel, emotional, social, systemic) from Baseline at each scheduled visit in Induction Study.
- Proportion of subjects with Inflammatory Bowel Disease Questionnaire (IBDQ) response (increase ≥ 16 points from Baseline) at each scheduled visit in Induction Study.
- Proportion of subjects with IBDQ remission (IBDQ ≥ 170 points) at each scheduled visit in Induction Study.
- Proportion of subjects achieving response in IBDQ fatigue item (increase of IBDQ fatigue item score ≥ 1) at Week 4
- Change in WPAI from Baseline at each scheduled visit in Induction Study.
- Change in EQ-5D from Baseline at each scheduled visit in Induction Study.
- Change in CDAI from Baseline at each scheduled visit in Induction Study.
- Change in SFPS from Baseline at each scheduled visit in Induction Study.
- Change in Abdominal Pain Rating Scale score from Baseline at each scheduled visit in Induction Study.

- Change in Bristol Stool Chart score from Baseline at each scheduled visit in Induction Study.
- Proportion of subjects who achieve Bristol Stool Chart response at each scheduled visit in Induction Study.
- Proportion of subjects who achieve CDAI remission (CDAI < 150) at each visit in Induction Study.
- Proportion of subjects who achieve SFPS remission (SFPS < 50) at each scheduled visit in Induction Study.
- Change in each CDAI component subscore (number of liquid or very soft stools, abdominal pain rating, general well-being, CD related complications, anti-diarrhea use, abdominal mass, hematocrit, body weight) from Baseline at each scheduled visit in Induction Study.
- Time to clinical remission (CDAI < 150) in Induction Study.
- Time to clinical response (CR –70) in Induction Study.
- Time to all cause Hospitalization in Induction Study.
- Time to CD-related Hospitalization in Induction Study.
- Proportion of subjects with no draining fistulas at Week 12 among subjects with draining fistula at Baseline.
- Proportion of subjects in each treatment group with > 50% reduction from Baseline in the number of draining fistulas at Week 12 among subjects with draining fistula at Baseline.
- Resolution of extra-intestinal manifestations over time at each scheduled visit in Induction Study.
- Proportion of subjects with an SES-CD decrease of ≥ 3 points compared to Baseline at Week 12.
- Proportion of subjects who achieve symptomatic remission, defined as average daily stool frequency ≤ 2.8 (and not worse than baseline) and average daily abdominal pain ≤ 1.0 (and not worse than baseline), at each scheduled visit in Induction Study among subjects with Baseline SF ≥ 4.0 and/or AP ≥ 2.0.
- Proportion of subjects who achieve symptomatic response, defined as average daily stool frequency at least 30% reduction from baseline and average daily abdominal pain not worse than baseline or average daily abdominal pain at

least 30% reduction from baseline and average daily stool frequency not worse than baseline, at each scheduled visit in Induction Study among subjects with Baseline SF ≥ 4.0 and/or AP ≥ 2.0.

The non-ranked secondary efficacy endpoints will be analyzed similarly to the ranked secondary efficacy endpoints.

Events of hospitalization and surgery will be identified by review of serious adverse events (SAEs) according to the hospitalization adjudication charter. All-cause hospitalization is defined as SAEs resulting in admission to the hospital or prolongation of an existing hospitalization for any reason. CD-related hospitalizations are defined as hospital admissions or prolongation of an existing hospitalization due to AEs that are related to CD. *P* values from both Log-rank test and Cox proportional hazards model will be provided.

10.4 Handling of Multiplicity

No adjustment for multiplicity will be done. In order to claim the study success, both co-primary endpoints need to be statistically significant at significance level 0.05 (two-sided). The co-primary endpoints will be tested prior to the ranked secondary endpoints; and the ranked secondary endpoints are tested for statistical significance only if both the co-primary endpoints meet statistical significance at a two-sided significance level of 0.05.

The statistical test for the ranked secondary endpoints will be carried out in the hierarchical order as specified in Section 10.3. This means that statistical significance for the remaining ranked secondary endpoints is assessed at 0.05 (2-sided) in the above order until the significance level exceeds 0.05. No additional statistically significant treatment differences could be declared after the ranked endpoint fails to achieve statistical significance at a two-sided significance level of 0.05.

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 immunosuppressants 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]
- Prior infliximab use (YES, NO)
- Weight (≤ median, > median)
- Baseline albumin (≤ median, > median)
- Disease duration (≤ 3 years, > 3 years)
- Disease duration (≤ median, > median)
- Region (US, ex-US)

11.0 Safety Analysis

11.1 General Considerations

All safety analyses will be performed on the safety analysis set. 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 AEs during Induction Study are defined as events that begin or worsen either on or after the first dose of the study drug and within 70 days after the last dose of the study drug for subjects who do not participate in the OLE (Study M14-347) or Maintenance study, or until first dose of maintenance study drug, or until first dose of study drug in the OLE study if the subject who enrolled under Amendment #3 or earlier, is a study completer and is enrolled in the OLE.

The 70-day follow-up phone call will not be required for any subject that initiates commercial adalimumab.

An overview of treatment-emergent AEs, including AEs of special interest such as adverse events leading to death and adverse events leading to premature discontinuation, AEs by Medical Dictionary for Drug Regulatory Activities (MedDRA version 21.0 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 percent.

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.
- Any Deaths

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.

All Adverse Events identified to AbbVie from the 70-day follow-up phone call will be collected as source data to be evaluated and reported. Thus, all SAEs and nonserious AEs as defined by AbbVie, reported during the 70-day follow-up phone call must be captured in the clinical database. The end of trial is the last subject contact, i.e., the 70-day follow-up call.

11.2.2 Adverse Events of Special Interest

The following AEs of special interest will be summarized by number and percentage of subjects experiencing an AE of interest. The AEs of interest will be summarized and presented using primary MedDRA system organ classes (SOCs) and preferred terms (PTs) for the following AE categories:

- Any Infections AE
- Any Serious Infection AE
- Any Legionella Infection AE
- Any Diverticulitis AE
- Any Opportunistic Infection AE (Excluding Oral Candidiasis and TB)
- Any Oral Candidiasis
- Any Tuberculosis AE
- Any Active Tuberculosis
- Any Latent Tuberculosis
- Any Parasitic Infection AE
- Any Reactivation of Hepatitis B
- Any Progressive Multifocal Leukoencephalopathy (PML) AE
- Any Malignancy AE
- Any Lymphoma AE

- Any Hepatosplenic T-Cell Lymphoma AE (HSTCL)
- Any Non-Melanoma Skin Cancer (NMSC) AE
- Any Melanoma AE
- Any Leukemia AE
- Any Other Malignant AE (Excluding NMSC, Melanoma, Lymphoma, HSTCL, and Leukemia)
- Any Allergic Reaction (Including Angioedema/Anaphylaxis)
- Any Lupus-Like Reactions and Systemic Lupus Erythematosus
- Any Vasculitis AE
- Any Cutaneous Vasculitis AE
- Any Non-Cutaneous Vasculitis AE
- Any Sarcoidosis AE
- Any Autoimmune Hepatitis
- Any Myocardial Infarction Related AE
- Any Cerebrovascular Accident Related AE
- Any Congestive Heart Failure Related AE
- Any Pulmonary Embolism Related AE
- Any Interstitial Lung Disease AE
- Any Intestinal Perforation AE
- Any Pancreatitis AE
- Any Stevens-Johnson Syndrome AE
- Any Erythema Multiforme Related AE
- Any Worsening/New Onset of Psoriasis
- Any Demyelinating Disorder
- Any Amyotrophic Lateral Sclerosis AE
- Any Reversible Posterior Leukoencephalopathy Syndrome (RPLS) AE
- Any Hematologic Disorders AE (Including Pancytopenia)
- Any Liver Failure and Other Liver Event AE
- Any Humira Administration Related Medication Errors AE

- Any Injection Site Reaction AE
- Any Intestinal Stricture
- Any AE Leading to Death
- Any AE Leading to Discontinuation of Study Drug.
- Any Deaths

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

11.3 Analysis of Laboratory Data

All laboratory data collected after first dose of maintenance study drug, or after first dose of study drug for Study M14-347 will be excluded from this analysis.

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.

Cross (Shift) tables from Baseline to the minimum, maximum, and 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 3.0 of Grade ≥ 3 will be provided. 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.

The laboratory data using Abbott plarform will be used for this analysis. If the data measured using Abbott plarform are not available, the Abbott equivalent (the values measured using Roche platform that were converted to Abbott equivalent) will be used.

11.3.1 Assessment of Shifts from Baseline in Liver-Specific Laboratory Variables

Baseline and post-baseline livers-specific laboratory will be categorized as follows:

- < 1.5 times the upper limit of the reference range
- ≥ 1.5 to < 3 the upper limit of the reference range
- ≥ 3 to < 5 the upper limit of the reference range
- ≥ 5 to < 8 the upper limit of the reference range
- ≥ 8 the upper limit of the reference range

For each variable, shift tables will be generated that cross tabulate the subjects' as deemed appropriate:

- Category of the baseline value versus category of the final value,
- Category of the baseline value versus maximum category

Note that the maximum category is used, rather than the category of the maximum values. The two may be different due to variation in the reference range.

No statistical tests will be performed for this analysis.

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

ALT ≥ 2.5 × ULN, or

AST ≥ 2.5 × ULN, or

Alkaline Phosphatase ≥ 2.5 × ULN, or

Total Bilirubin ≥ 1.5 × ULN.

11.3.2 Hy's Law Cases

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

All vital signs and weight data collected after first dose of maintenance study drug, or after first dose of study drug for Study M14-347 will be excluded from this analysis.

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

- Body Weight (kg)
- Blood Pressure (Systolic/Diastolic) (mmHg)
- Pulse (bpm)
- Respiratory Rate (RPM)
- 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 number and percentage of subjects meeting the criteria for potentially clinically significant (PCS) vital sign values will also be summarized.

The criteria for potentially clinically significant vital sign findings are presented in Table 8.

Table 8. Criteria for Potentially Clinically Significant Vital Sign Findings

Vital Sign	Category	Criteria for Potential Clinically Significant Vital Signs
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 mmHg from Baseline
	High	Value ≥ 120 mmHg and/or increase ≥ 15 mmHg from Baseline

Vital sign results meeting the criteria for PCS findings will be identified in a listing.

12.0 Summary of Changes

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

None.

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

Under Section 6.0 Analysis Conventions: Only CD related systemic or rectal corticosteroids will be used for censoring purpose, and will have last non-missing value prior to censoring carried forward for non-categorical assessments) from that point forward. This change is made to be consistent with protocol Amendment 6.

In addition, SES-CD scoring rule is updated regarding missing individual variables. The missing SES-CD individual variables at Baseline will be imputed as 0. The post-Baseline missing SES-CD individual variables will be imputed as 0 if there are 8 or fewer missing individual variables. If more than 8 individual variables are missing, the total SES-CD will be missing for that visit. Both changes are based on newly created Analysis Convention for Crohn's Disease and Ulcerative Colitis Clinical Studies. The new imputation rule is also used by current Abbvie IBD registrational trials.

Bristol Stool Form Scale is not a continuous variable. The new variable Bristol Stool Chart score is defined 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 achieved reduction from Baseline ≥ 50%, this subject will be classified as Bristol Stool Chart Responder. This change is made to be consistent with protocol Amendment 6.

The Week 12 visit window for SES-CD is updated so that no post-baseline Induction data will be discarded. In addition, endoscopy data collected prior to or within 7 days of the fist dose of maintenance study or OLE Study M14-347 will be considered for the Week 12 visit window for SES-CD; thus ensuring endoscopy data collected is not ignored

while also reflecting the efficacy based on the study treatment during the 12 week induction period.

Under Section 10.2, one of the co-primary endpoints, SES-CD ≤ 4 and at least a 2 point reduction versus baseline and no subscore greater than 1 in any individual variable) at Week 12 is replaced by decrease in SES-CD > 50% from Baseline (or for a Baseline SES-CD of 4, at least a 2 point reduction from Baseline) at Week 12. This change is made to be consistent with protocol Amendment 6.

Under Section 10.3, 13th (Change in IBDQ from Baseline at Week 4) and 14th (Change in IBDQ from Baseline at Week 12) ranked endpoints have been replaced by new 13th (Proportion of subjects achieving response in IBDQ Bowel Symptom domain (increase of IBDQ bowel symptom domain score ≥ 8) at Week 4) and new 14th (Proportion of subjects achieving response in IBDQ Bowel Symptom domain (increase of IBDQ bowel symptom domain score ≥ 8) at Week 12) endpoints. Proportion of subjects achieving response in IBDQ fatigue item (increase of IBDQ fatigue item score ≥ 1) at Week 12 is added as 15th ranked endpoint. The above changes were made to be consistent with protocol Amendment 6.

In addition, some of the non-ranked endpoints are also updated to be consistent with protocol Amendment 6.

13.0 Appendix

14.0 References

1. Devlin NJ, Shah KK, Feng Y, et al. Valuing health-related quality of life: an EQ-5D-5L value set for England. Health Econ. 2018 Jan;27(1):7-22.
Epub 2017 Aug 22