1.0 Title Page

Statistical Analysis Plan

Study M13-375

A Multicenter, Randomized, Double-Blind, Study Comparing the Efficacy and Safety of Continuing Versus Withdrawing Adalimumab Therapy in Maintaining Remission in Subjects with Non-Radiographic Axial Spondyloarthritis

Date: 02 March 2017

2.0 Table of Contents

List of Tables

3.0 Introduction

This Statistical Analysis Plan (SAP) was created based on Study Protocol M13-375 Amendment 2 dated 15 January 2015 and Administrative Change 1 dated 29 May 2015.

This statistical analysis plan will describe the methods for summarizing and analyzing the efficacy and safety data obtained from Study M13-375. This document describes the populations that will be analyzed, the variables that will be analyzed, and the statistical methods that will be utilized. The intention of this plan is to provide the details of the planned statistical analyses outlined in the study protocol before the database lock for the Period 2 analysis. However, features of the study and the data collected may suggest that deviations from the approach presented in this analysis plan are appropriate, and exploratory analyses can be added. A detailed description of any substantive modifications to this SAP will be included in the final clinical study report (CSR). Serum adalimumab (ADA) concentration and anti-ADA antibodies (AAA) will be studied by pharmacokinetics group and a separate report will be provided.

A final CSR is planned after the study is completed.

Unless noted otherwise, Derived Data Sets will be programmed using SAS® Version 8.2 (SAS Institute, Inc., Cary, NC 27513) or later under the UNIX operating system. Analyses will be performed using SAS® Version 9.3 (SAS Institute, Inc., Cary, NC 27513) or later under the UNIX operating system.

4.0 Study Objectives, Design and Procedures

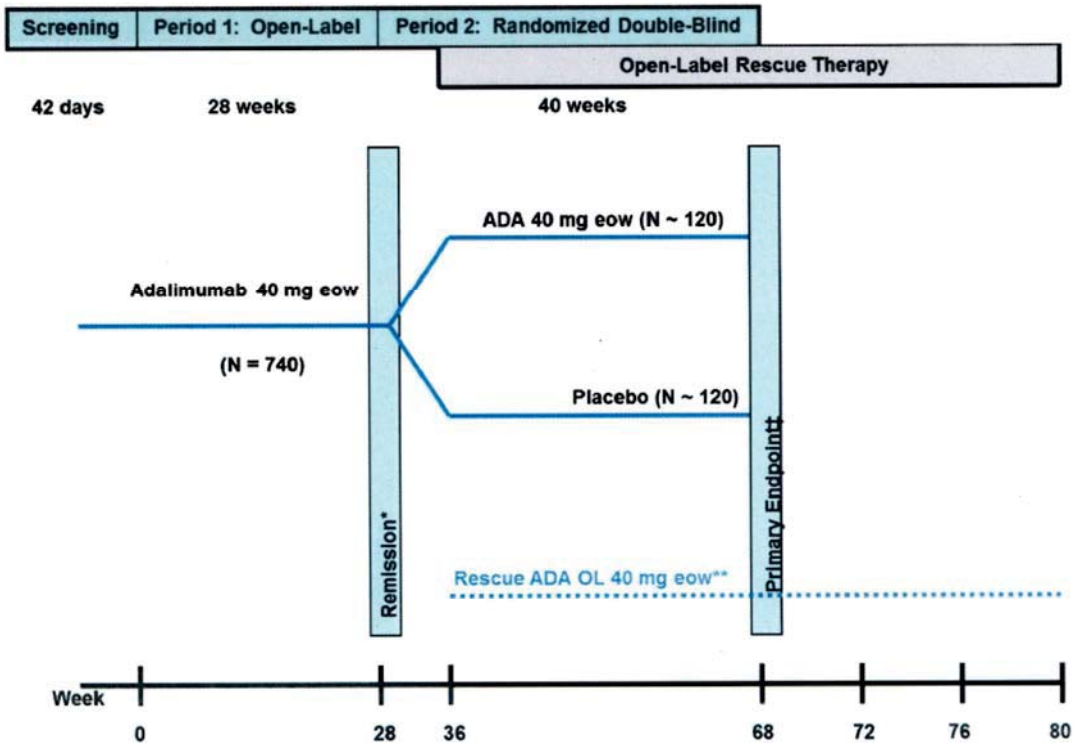
4.1 Objectives

The objective of this study is to evaluate the efficacy and safety of continuing versus withdrawing therapy with adalimumab 40 mg given eow SC in maintaining remission in subjects with nr-axSpA.

4.2 Design Diagram

The Study M13-375 study duration includes a 42 day Screening Period, a 28-week open-label (OL) 40 mg adalimumab eow treatment period (Period 1), a 40-week double-blind placebo-controlled treatment period (Period 2) with an opportunity to receive at least 12 weeks of rescue therapy (subjects who flare at Weeks 60, 64 or 68 will be allowed 12 weeks of rescue therapy and final visits will be at Weeks 72, 76 or 80 respectively), plus a 70 day follow-up phone call. The study design schematic is presented in Figure 1. Length of exposure will depend on remission or flare status (20 [first time Ankylosing Spondylitis Disease Activity Score (ASDAS) remission will be calculated] to 80 weeks of treatment).

Figure 1. Study Design Schematic



* Subjects who meet remission criteria (ASDAS inactive disease < 1.300) at Weeks 16, 20, 24 and 28 will be randomized to Period 2. Subjects not meeting remission criteria will be discontinued at Weeks 20, 24 or 28.

** Subjects who flare (ASDAS ≥ 2.100 for 2 consecutive study visits) during the double-blind period (Period 2) can receive rescue therapy with OL ADA 40 mg eow for at least 12 weeks and will continue through the duration of the subject's participation of this study. The earliest time a subject can meet the flare criteria would be at Week 36 since a subject would need to have 2 consecutive visits with ASDAS ≥ 2.100. Subjects who flare at Weeks 60, 64 or 68, rescue therapy with OL adalimumab will be provided for 12 weeks and final study visit will be at Weeks 72, 76 or 80 respectively.

‡ Primary Endpoints: Proportion of subjects who do not experience a flare by Week 68 of the study. Study participation will end at Week 68 for subjects who meet the primary endpoint.

Note: Study enrollment will be closed prior to reaching 740 subjects should a sufficient number of subjects fulfill remission criteria and be available for randomization at the end of Period 1 prior to full enrollment.

4.3 Sample Size

Assuming that 50% of subjects treated with adalimumab will not experience any flare up to Week 68, 120 subjects per arm are required for the double-blind period (Period 2) (with an adjustment for dropout rate of no more than 30%) and is targeted to result in at least 90% power for detecting a 25% difference at Week 68 for the primary endpoint between adalimumab and placebo based on a two-sided test with significance level 0.05.

It is assumed that approximately 33% of the subjects enrolled into the OL period (Period 1) will meet ASDAS Inactive Disease at Weeks 16, 20, 24 and Week 28; therefore, a total of approximately 740 subjects would be required to be enrolled in order to have a sufficient number of subjects (i.e., 120 subjects per arm) for the double-blind period. Study enrollment will be closed prior to reaching 740 subjects should a sufficient number of subjects fulfill remission criteria and be available for randomization at the end of Period 1 prior to full enrollment.

5.0 Analysis Populations

5.1 Definition for Analysis Populations

Full Analysis Set Population

The Full Analysis Set (FAS) population comprises of all subjects who enrolled into the OL period (Period 1) and received at least one dose of adalimumab.

The FAS is used for all analysis during the OL period (Period 1).

Modified Intent-to-Treat Population

The modified Intent-To-Treat (mITT) population comprises all subjects who were randomized into double-blind (DB) Period 2 and received at least one dose of DB study medication. The data from the mITT population will be analyzed by the treatment group assignment given at the time of randomization, even if the subject does not receive the correct treatment, is not compliant to the protocol procedures, or does not follow the protocol until completion.

The mITT population is used for all efficacy analysis of the DB period.

Per Protocol Population

The Per Protocol Population (PPP) consists of all mITT subjects who entered the DB period of the study (Period 2) and did not have any major protocol violation expected to influence the primary endpoint during the OL and DB periods. The determination of major protocol deviation will be finalized prior to the final database lock.

The PPP is used as a sensitivity analysis of the primary efficacy endpoint during the DB period.

Double-Blind Safety Population

The DB safety population consists of all subjects who received at least one dose of study medication during the DB period. For analyses of the DB safety population, each subject's treatment group will be the actual treatment received by the subject even if this differs from his/her randomized treatment assignment. If all subjects take the treatment to which they were randomly assigned, the safety population will be the same as the mITT population.

The DB safety population is used for all safety analysis of the DB period.

Rescued Population

The Rescued Population consists of all patients who received OL rescue treatment after the DB period.

The Rescued Population is used for analysis of selected efficacy and all safety data collected post-rescue.

Any Adalimumab Population

The Any Adalimumab (ADA) population consists of all subjects who received at least one dose of adalimumab any time during the study (including OL period, DB period and rescue period).

The Any ADA population is used for safety analysis under the exposure of Adalimumab treatment any time during the study.

5.2 Variables Used for Stratification of Randomization

No stratification was performed for this study.

6.0 Analysis Conventions

6.1 Definition of Baseline

The baseline value for Period 1 (OL period) will be the last available pre-treatment value recorded before the first dose of OL study drug. If multiple measurements are recorded on the same day, the measurement recorded at the latest time (if available) prior to dosing will be used as Baseline. The baseline value for Period 2 (DB period) will be determined by the last non-missing measurement recorded before the first DB dose of study drug. If multiple measurements are recorded on the same day, the measurement recorded at the latest time (if available) prior to dosing will be used as Baseline. The baseline value for the rescue period will be determined by the last non-missing measurement recorded before the first rescue dose of study drug. If multiple measurements are recorded on the same day, the measurement recorded at the latest time (if available) prior to dosing will be used as Baseline.

For efficacy analysis of the OL and DB period, all changes and percent changes will be calculated based on OL baseline. For efficacy analysis of the rescue period, changes and percent changes will be calculated based on the rescue baseline.

For safety, OL baseline is used for OL period analysis on FAS; DB Baseline is used for DB safety analysis on DB safety population and rescue baseline is used for rescue period analysis for rescued population. The OL Baseline is used for safety analysis for any ADA population.

6.2 Definition of Final Observation

Final observation for Period 1 is defined as the last non-missing observation post-baseline of Period 1. For subjects who were randomized to the DB period, the final observation for Period 1 is defined as the last non-missing observation on or prior to the date of the first dose of DB period.

Final observation for Period 2 is defined as the last non-missing observation post-baseline of Period 2. For those subjects who receive OL rescue medication the final observation for double-blind Period will be the last non-missing observation prior to receiving OL rescue medication. Final observation for rescue period is defined as the last non-missing observation after received the OL rescue medication.

6.3 Definition of Rx Days

Rx Days are calculated for each time point relative to visit. Rx days for OL Period 1 are defined as the number of days between the day of the first dose of study drug in OL Period 1 and the specific time point. Rx days for DB Period 2 are defined as the number of days between the day of the first dose of study drug for DB Period 2 and the specific time point. Rx days are negative values when the time point of interest is prior to the first study drug dose day. Rx days are positive values when the time point of interest is after the first study drug dose day. The day of the first dose of study drug is defined as Rx Day 1, while the day prior to the first study drug 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 for each study period as described in Definition of Analysis Windows.

6.4 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 case report form (CRF) does not correspond to multiple visit windows. Moreover, windows will not discard any post-baseline measurement recorded on the CRF. For efficacy, vital signs, or laboratory values, the final observations will be included in the tabulation as captured in the visit windows listed in Table 1, Table 2, Table 3, Table 4, Table 5, Table 6 and Table 7.

If a subject had 2 or more actual visits in one visit window, the visit closest to the target will be used as the study visit for that window. If 2 visits are equidistant from the target day, then the later visit will be used for reporting. For hs-CRP, the last observation in the window will be used.

Table 1 provides the analysis windows for the following parameters for OL Period 1:

- Vital Signs/Weight
- Anterior Uveitis
- Chemistry and Hematology (no Week 2 visit)
- BASDAI
- BASFI
- Physician's Global Assessment of Disease Activity NRS
- Patient's Assessment of Nocturnal Back Pain NRS
- Patient's Assessment of Total Back Pain NRS
- Patient's Global Assessment of Disease Activity NRS
- Patient's Global Assessment of Pain NRS
- hs-CRP (no Week 2 visit)
- ASDAS (no Week 2 visit)

Table 1. Analysis Windows for All Protocol-Specified Visits at Weeks 0, 2, 4, 8, 12, 16, 20, 24 and 28 (OL Period 1)

Protocol Specified Study Visit Week	Lower Bound	Open-Label Nominal Day	Upper Bound
0 (Baseline)	–999	1[a]	1
2	2	15	22
4	23[b]	29	43
8	44	57	71
12	72	85	99
16	100	113	127
20	128	141	155
24	156	169	183
28	184	197	211

a. Day of first dose of study drug during OL period.

b. For parameters without Week 2, the window for Week 4 lower bound is 2.

Table 2 provides the analysis windows for the following parameters for OL Period 1:

- Urinalysis
- TJC/SJC
- $BASMI_{lin}$
- MASES
- Plantar Fascia Enthesitis
- Dactylitis
- HAQ-S
- SF-36V2 Health Survey
- WPAI-Axial SpA
- EQ-5D

Table 2. Analysis Windows for All Protocol-Specified Visits at Weeks 0, 12, 20, 24 and 28 (OL Period 1)

Protocol Specified Study Visit Week	Lower Bound	Open-Label Nominal Day	Upper Bound
0 (Baseline)	–999	1[a]	1
12	2	85	99
20	100	141	155
24	156	169	183
28	184	197	211

a. Day of first dose of study drug during OL.

Table 3 provides the analysis windows for the following parameters for DB Period 2:

- Vital Signs/Weight
- Anterior Uveitis
- Chemistry and Hematology
- BASDAI
- BASFI
- Physician's Global Assessment of Disease Activity NRS
- Patient's Assessment of Nocturnal Back Pain NRS
- Patient's Assessment of Total Back Pain NRS
- Patient's Global Assessment of Disease Activity NRS
- Patient's Global Assessment of Pain NRS
- hs-CRP
- ASDAS

Table 3. Analysis Windows for All Protocol-Specified Visits at Weeks 28, 32, 36, 40, 44, 48, 52, 56, 60, 64, 68 (DB Period 2)

Protocol Specified Study Visit Week	Double-Blind Week	Lower Bound	Double-Blind Nominal Day	Upper Bound
28 (Baseline)	0	–999	1[a]	1
32	4	2	29	43
36	8	44	57	71
40	12	72	85	99
44	16	100	113	127
48	20	128	141	155
52	24	156	169	183
56	28	184	197	211
60	32	212	225	239
64	36	240	253	267
68	40	268	281	295

a. Day of first dose of study drug during DB.

Table 4 provides the analysis windows for the following parameters for DB Period 2:

- Urinalysis
- TJC/SJC
- $BASMI_{lin}$
- MASES
- Plantar Fascia Enthesitis
- Dactylitis
- HAQ-S
- SF-36V2 Health Survey
- WPAI-Axial SpA
- EQ-5D

Table 4. Analysis Windows for All Protocol-Specified Visits at Weeks 28, 36, 52, 68 (DB Period 2)

Protocol Specified Study Visit Week	Double-Blind Week	Lower Bound	Double-Blind Nominal Day	Upper Bound
28 (Baseline)	0	–999	1[a]	1
36	8	2	57	113
52	24	114	169	225
68	40	226	281	295

a. Day of first dose of study drug during DB.

Table 5 provides the analysis windows for the following parameters for DB Period 2:

- Chemistry and Hematology.

Table 5. Analysis Windows for All Protocol-Specified Visits at Weeks 28, 36, 44, 52, 68 (DB Period 2)

Protocol Specified Study Visit Week	Double-Blind Week	Lower Bound	Double-Blind Nominal Day	Upper Bound
28 (Baseline)	0	–999	1[a]	1
36	8	2	57	85
44	16	86	113	141
52	24	142	169	225
68	40	226	281	295

a. Day of first dose of study drug during DB.

Table 6 provides the analysis windows for the following parameters for DB Period 2:

- TB Screening (PPD Skin Test or IGRA)

Table 6. Analysis Windows for All Protocol-Specified Visits at Week 52 (DB Period 2)

Protocol Specified Study Visit Week	Double-Blind Week	Lower Bound	Double-Blind Nominal Day	Upper Bound
0 (Baseline)	–28	–999	1[a]	1
52	24	2	169	295

a. Day of first dose of study drug during OL.

Table 7 provides the analysis windows for the Rescued Population for all possible visits.

Table 7. Analysis Windows for Rescued Population Visits at Rescue Week 0, 4, 8, 12, 16, 20, 24, 28, 32

Rescue Visit Week	Lower Bound	Post-Rescue Nominal Day	Upper Bound
0	–999	1[a]	1
4	2	29	43
8	44	57	71
12	72	85	99
16	100	113	127
20	128	141	155
24	156	169	183
28	184	197	211
32	212	225	239

a. Last observation prior to first dose of rescue medication.

6.5 Missing Data Imputation Methods

There will be no missing data imputation for Period 1. Observed cases will be reported.

To account for missing data for the categorical efficacy endpoints, a non-responder imputation (NRI) approach will be used. That is, the missing value will be considered as a non-responder. A mixed-effect repeated measures model (MMRM) will be used for the analysis of continuous efficacy endpoints. In addition, an analysis using only the observed or reported data will be performed as a sensitivity analysis.

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

Demographics, baseline characteristics, prior and concomitant medications will be summarized as detailed in the following sections.

7.1 Demographic and Baseline Characteristics

The following demographic and baseline characteristics variables will be summarized.

Demographics

- Gender [male/female]
- Age [years]
- Age category (< 40 years, 40 to 65 years, > 65 years)
- Body weight [kg] – overall and by gender
- Body weight category (< 70 kg, ≥ 70 kg)
- Height [cm]
- Body mass index (BMI) [kg/m^2]
- BMI category [Underweight (< 18.5), Normal (18.5 – < 25), Overweight (25 – < 30.0), Obese (≥ 30.0)] based on WHO cutoffs
- Race
- Ethnicity
- Region (4 categories: North America (US/Canada), Central/South America (Mexico/Brazil), Europe, Australia/New Zealand)

Baseline Characteristics

- Duration (years) since SpA diagnosis: defined as the duration between date of first SpA diagnosis as collected in the Medical History for Axial Spondyloarthritis Case Report Form and date of first dose of OL study drug
- Duration (years) since first symptom of SpA: defined as the duration between date of first symptom of SpA as collected in the Medical History for Axial Spondyloarthritis Case Report Form and date of first dose of OL study drug

- HLA-B27 (negative, positive)
- OL Period 1: Contingency table of MRI of the SI joints (positive, negative) and spine (positive, negative), overall and by CRP (elevated, normal) category
- OL Period 1: Contingency table of N (%) fulfilling the ASAS criteria imaging arm (MRI of SI joint positive with at least 1 additional SpA feature) and clinical arm (HLA-B27 positive with at least 1 additional SpA feature)
- OL Period 1: Anteroposterior (AP) pelvis x-ray (presence, absence of sacroiliitis, presence = sacroiliitis grade ≥ 2 bilaterally or sacroiliitis grade 3 – 4 unilaterally).

Efficacy Measurements at Baseline

- Tender joint count (68 joints)
- Swollen joint count (66 joints)
- Patient's Global Assessment of Disease Activity NRS (0 – 10)
- BASDAI NRS (0 – 10)
- BASDAI categories (< 4, 4 to 6, > 6)
- Inflammation/Morning Stiffness (0 – 10 NRS) (the average of Items 5 and 6 of the BASDAI)
- BASFI NRS (0 to 10)
- Patient's Assessment of Total Back Pain NRS (0 – 10)
- Patient's Assessment of Total Back Pain categories (< 4, ≥ 4)
- Patient's Assessment of Nocturnal pain NRS (0 – 10)
- ASDAS
- ASDAS categories (inactive disease state, moderate, high and very high)
- hs-CRP (mg/L)
- hs-CRP categories (normal as ≤ 2.87 mg/L, above upper limit of normal as > 2.87 mg/L)
- Linear BASMI score
- MASES
- Physician Global Assessment of Disease Activity NRS (0 – 10)

- Patient's Global Assessment of Pain NRS (0 – 10)
- Dactylitis
- Plantar Fascia enthesitis
- OL Period 1: SPARCC MRI score for the spine
- OL Period 1: SPARCC MRI score for the SI joints
- SF-36 Version 2 Health Survey: physical component summary, and mental component summary
- HAQ-S
- WPAI: Axial SpA
- EQ-5D

Vital Signs at Baseline

- Systolic and diastolic blood pressure [mmHg]
- Pulse [bpm]
- Respiratory rate [rpm]
- Temperature [°C]

OL Period 1: Electrocardiogram (ECG)

- Overall ECG assessment at Baseline [Normal, Abnormal – not clinically significant, Abnormal – Clinically significant, Unable to evaluate]
- Abnormal ECG details [Rhythm, Conduction, Axis, Ventricular hypertrophy, Myocardial infarction, ST segment, T waves, Abnormal U waves, Corrected QT interval using Bazett Correction (QTcB), and Other]

OL Period 1: Chest X-Ray, Purified Protein Derivative (PPD), Tuberculosis (TB) Prophylaxis

- Chest x-ray at Baseline
  - Normal, Abnormal
  - Calcified granulomas [Absent, Present]
  - Pleural scarring [Absent, Present]

- Pleural thickening [Absent, Present]

- PPD Skin Test or IGRA [Positive, Negative, Missing] (Note that a positive or negative PPD skin test will be determined by the PPD mm induration result on the CRF. An induration of less than 5 mm is considered negative. IGRA may be indeterminate in which case the subject should have a repeat test – either another IGRA or a PPD. The results of the second test are then considered final. If a second IGRA is indeterminate they are considered to be positive). The number and percentage will be reported for PPD and IGRA separately, as well as combined (Positive by either test, both Negative)
- Enrolled under TB prophylaxis [Yes, No]
  - If enrolled [INH, Other]

OL Period 1: Tobacco/Nicotine and Alcohol Use

- Tobacco/Nicotine Use [user, ex-user, non-user, unknown]
- Alcohol Use [drinker, ex-drinker, non-drinker, unknown]

7.1.1 Analysis of Demographic Data and Baseline Characteristics

Demographic and baseline characteristics will be summarized for the FAS at OL baseline and mITT population at both OL and DB baseline. The number of observations, mean, standard deviation, median, minimum and maximum will be summarized for continuous variables. Categorical or discrete variables will be summarized via counts and percentages. Details of descriptive statistics for each demographic and baseline characteristic are provided in Table 8 below.

Table 8. List of Variables for Demographics and Baseline Characteristics

Variable	Analysis Method
1. Age 2. Body weight 3. Height 4. BMI 5. Duration (years) of Axial Spondyloarthritis 6. Duration (years) of Axial Spondyloarthritis Symptoms 7. TJC68 8. SJC66 9. MASES 10. Patient's Assessment of Nocturnal pain (NRS) 11. Patient's Global Assessment of Disease Activity (NRS) 12. Physician's Global Assessment of Disease Activity (NRS) 13. Patient's Global Assessment of Pain (NRS) 14. Patient's Assessment of Total Back Pain (NRS) 15. Inflammation/Morning Stiffness (NRS) 16. hs-CRP 17. BASDAI (NRS) 18. BASFI (NRS) 19. ASDAS 20. Linear BASMI 21. Dactylitis 22. Systolic and diastolic blood pressure 23. Pulse 24. Respiratory rate 25. Temperature 26. SF-36 v.2 27. WPAI-Axial SpA2 28. HAQ-S 29. EQ-5D 30. OL Period 1: SPARCC MRI score (spine) 31. OL Period 1: SPARCC MRI score (SI joints)	Items 1 through 31 are continuous measures and will be summarized and presented using the number of observations, mean, standard deviation, median, minimum and maximum. Treatment comparison will be performed utilizing ANOVA with treatment as a factor.

Table 8. List of Variables for Demographics and Baseline Characteristics (Continued)

Variable	Analysis Method
32. Gender 33. Age category 34. Body weight category 35. BMI category 36. Race 37. Ethnicity 38. Region 39. BASDAI categories 40. Patient's Assessment of Total Back pain categories 41. hs-CRP categories 42. ASDAS categories 43. OL Period 1: Anterior Uveitis presence (categorical) 44. Plantar fascia Enthesitis 45. OL Period 1: Overall ECG assessment 46. OL Period 1: Abnormal ECG details 47. OL Period 1: Chest X-ray 48. OL Period 1: Enrolled under TB prophylaxis 49. OL Period 1: PPD Skin Test 50. OL Period 1: IGRA Test 51. OL Period 1: PPD Skin Test and IGRA Test combined 52. HLA-B27 53. OL Period 1: MRI Spine Positive vs. Negative for inflammation 54. OL Period 1: MRI SI Joints Positive vs. Negative for inflammation 55. OL Period 1: MRI either SI Joints or Spine Positive vs. Negative for inflammation 56. OL Period 1: AP Pelvis X-ray grade fulfilling the modified New York criteria for AS (yes/no) 57. OL Period 1: Tobacco and Alcohol use	Items 32 through 57 will be summarized via counts and percentages. Treatment comparison will be performed using Fisher's exact test for categorical variables when presented by randomized treatment groups.

7.2 Medical History

Non-SpA medical history data will be summarized and presented using body systems and conditions/diagnoses as captured on the CRF for the FAS. The following body systems

will be reported (if no subject reported any pertinent information for a specific body system, then that category will not be displayed in the summary table):

A. Eye/Ear/Nose/Throat, **B.** Neurologic and Psychiatric System, **C.** Pulmonary, **D.** Cardiovascular, **E.** Musculoskeletal, **F.** Blood, **G.** Gastrointestinal, **H.** Metabolic, **I.** Genitourinary, **J.** Whole Body, **Z.** Other Body System

The body systems will be presented in alphabetical order and the conditions/diagnoses will be presented in alphabetical order within each body system. The number and percentage of subjects with a particular condition/diagnosis will be summarized for each treatment group and overall. Subjects reporting more than 1 condition/diagnosis within a body system will be counted only once for that body system. No statistical comparison will be performed on medical history.

Medical History will be presented by count and percentage of subjects broken down by Body System and Diagnosis.

7.3 Medical History for Axial Spondyloarthritis

Medical history for axial SpA data will be summarized and presented using the parameters as captured on the CRF for the FAS. The number and percentage of subjects with a particular parameter will be summarized.

Number of subjects in the following manifestation of disease categories will be summarized:

- MRI evidence of active inflammatory lesions of sacroiliac joints with definite bone marrow edema/osteitis, suggestive of sacroiliitis associated with SpA
- Inflammatory back pain meeting 4 of the 5 parameters:
  1. Age at onset < 40 yrs
  2. Insidious onset
  3. Improvement with exercise
  4. No improvement with rest

5. Night pain with improvement upon getting up

- Arthritis (past or present)
- Heel enthesitis (past or present)
- Anterior Uveitis confirmed by an ophthalmologist (past or present)
- Dactylitis (past or present)
- Crohn's disease or ulcerative colitis (past or present)
- Good prior response to NSAIDs (back pain is not present anymore or much better 24 to 48 hours after a full dose of a NSAID)
- Family history of SpA
- Elevated CRP (past or present)
- Psoriasis

7.4 Previous Treatment and Concomitant Medications

The number and percent of subjects who received prior or concomitant medication will be tabulated by the generic name assigned by the World Health Organization (WHO) Dictionary. A particular medication will be classified as "prior" or "concomitant" or "both" according to the rules provided in Table 9 below.

For OL Period 1, "FD" is the date of the first dose of study drug and "LD" is the date of the last dose of study drug during Period 1 plus minimum of 14 days or days between last dose of OL Period 1 and first dose of DB period, whichever is smaller.

For DB Period 2, "FD" is the date of the first dose of study drug during DB Period 2 and "LD" is the date of the last dose of study drug during Period 2 plus 14 days. Prior and Concomitant medications will be summarized by treatment group. Prior medication at baseline include medication that have a start date before first DB dose of study drug and are ongoing or have a stop before first DB dose of study drug. Concomitant medications at baseline include prior medications that have a start date before first DB dose of study drug and are ongoing or have a stop date after first DB dose of study drug.

Table 9. Rules for Prior and Concomitant Medication Classification

Medication Start Date	Ongoing?	Medication Stop Date					
		Missing	< FD	= FD	(FD, LD)	= LD	> LD
Missing	YES	1, 2					
	NO/Missing	1, 2	1	1, 2	1, 2	1, 2	1, 2
< FD	YES	1, 2					
	NO/Missing		1	1, 2	1, 2	1, 2	1, 2
= FD	YES	2					
	NO/Missing			2	2	2	2
(FD, LD)	YES	2					
	NO/Missing				2	2	2
= LD	YES	2					
	NO/Missing					2	2
> LD	YES	3					
	NO/Missing						3

1 = Prior; 2 = Concomitant; 3 = Post-treatment

7.4.1 Reporting Special Medications

OL Period 1

The number and percent of subjects who received prior medications will be tabulated separately by the generic name assigned by the World Health Organization (WHO) Drug Dictionary and will be summarized for the number and percentage of subjects using nonsteroidal anti-inflammatory drugs (NSAIDs), disease-modifying anti-rheumatic drugs (DMARDs), systemic corticosteroids, and other medications (excluding DMARDs, NSAIDs, and systemic corticosteroids).

The number and percent of subjects who received concomitant medications will be tabulated separately by the generic name assigned by the WHO Drug Dictionary and will be summarized for the number and percentage of subjects using NSAIDs, DMARDs, systemic corticosteroids, and other medications (excluding DMARDs, NSAIDs, and systemic corticosteroids).

The number and percent of subjects who received a prior medication for SpA will be tabulated by the generic name (assigned by the World Health Organization (WHO) Drug Dictionary) and for the three medication categories: nonsteroidal anti-inflammatory drugs (NSAIDs), disease-modifying antirheumatic drugs (DMARDs), and corticosteroids prior to BL or at the time of the first dose of study drug in the OL Period 1 (including ongoing at BL). The number of different DMARDs, NSAIDs, and overall medication taken per subject as prior medication for axial SpA (including ongoing at BL) will be reported.

DB Period 2

The number and percent of subjects who received prior medications will be tabulated separately by the generic name assigned by the World Health Organization (WHO) Drug Dictionary and will be summarized for the number and percentage of subjects using nonsteroidal anti-inflammatory drugs (NSAIDs), disease-modifying anti-rheumatic drugs (DMARDs), systemic corticosteroids, and other medications (excluding DMARDs, NSAIDs, and systemic corticosteroids).

The number and percent of subjects who received concomitant medications will be tabulated separately by the generic name assigned by the WHO Drug Dictionary and will be summarized for the number and percentage of subjects using NSAIDs, DMARDs, systemic corticosteroids, and other medications (excluding DMARDs, NSAIDs, and systemic corticosteroids).

The number of different DMARDS, NSAIDs taken per subject as concomitant medication regardless of indication for use (including medications ongoing at DB Period 2 BL and taken prior to the first dose of OL rescue therapy) will be reported. The same summary will be repeated for different DMARDS, NSAIDs taken per subject as concomitant medication for SpA.

7.5 Protocol Deviation

Number and percentage of subjects who reported at least 1 of the following protocol deviation categories will be provided:

- Inclusion/exclusion criteria deviations
- Developed withdrawal criteria but was not withdrawn
- Received wrong treatment or incorrect dose
- Received excluded concomitant treatment

7.6 Inclusion/Exclusion Criteria

For OL Period 1, number and percentage of subjects who did not meet either inclusion or exclusion criteria will be reported.

8.0 Patient Disposition

8.1 Patient Disposition

OL Period 1

The number of subjects will be tabulated by investigator site and overall for the following categories: discontinued study at Week 20 due to failure to achieve study defined remission, discontinued study at Week 24 due to failure to achieve study defined remission, discontinued study at Week 28 due to failure to achieve study defined remission, discontinued from OL Period 1 due to reasons other than failure to achieve study defined remission, and completed Week 28 of OL Period 1 and randomized into DB Period 2 for FAS. The number and percentage of subjects by windowed efficacy visits also will be tabulated.

DB Period 2

The number of subjects will be tabulated by investigator site and overall for the following categories: randomized, treated, included in mITT population, included in PP population, included in DB safety population, completed study, started rescue therapy with OL adalimumab, and discontinued for each treatment group and overall. The number and percentage of subjects by windowed efficacy visits also will be tabulated.

Reasons for discontinuation of study drug will be summarized as recorded on the CRF by the following categories: adverse event (AE), withdrew consent, lost to follow-up, lack of efficacy, and other, for all randomized subjects. Subjects may have more than one reason for discontinuing study drug, but they will be counted once for the total number of discontinuations.

9.0 Study Drug Exposure and Compliance

OL Period 1

Study drug exposure will be summarized as follows:

Exposure to any adalimumab in OL Period 1 = number of days since the first adalimumab dose in OL Period 1 through the last adalimumab dose date plus 14 days in OL Period 1 for subjects who did not enter double-blind period. For subjects who entered double-blind period, OL Period 1 exposure is first dose of double-blind dose – first adalimumab dose in OL Period 1.

These will be presented by the number of subjects, mean, standard deviation, median, minimum, and maximum. In addition, the number and percentage of subjects exposed to study drug will be summarized for the following exclusive duration intervals:

Any Adalimumab in OL Period 1

- ≥ 1 day
- ≥ 15 days
- ≥ 29 days
- ≥ 57 days
- ≥ 85 days
- ≥ 113 days
- ≥ 141 days
- ≥ 169 days
- ≥ 197 days

Compliance

Compliance (%) calculated as (number of injections received/number of injections planned) × 100%, rounded to 0.1%, prior to first dose in DB Period 2.

DB Period 2

Study drug exposure will be summarized as follows:

Exposure to study drug in DB Period 2 = number of days since the first dose in DB Period 2 through the last dose date plus 14 days prior to rescue in DB Period 2 for subject who did not enter rescue period. For subjects who entered rescue period, DB Period 2 exposure is first dose of rescue period – first dose of DB Period 2.

These will be presented by the number of subjects, mean, standard deviation, median, minimum, and maximum. In addition, the number and percentage of subjects exposed to study drug will be summarized by 12-week intervals:

- ≥ 12 weeks (85 days)
- ≥ 24 weeks (169 days)
- ≥ 36 weeks (253 days)
- ≥ 48 weeks (337 days)

Compliance

Compliance (%) calculated as (number of injections received/number of injections planned) × 100%, rounded to 0.1%, prior to rescue therapy in DB Period 2.

Rescue Period

Study drug exposure will be calculated as follows:

Exposure to study drug in rescue period = number of days since the first dose in rescue period through the last dose date + 14 days.

Any Adalimumab Exposure

Exposure to adalimumab is the last dose date – first dose in OL Period 1 + 14 days for subjects who received adalimumab in double-blind Period 2. For subjects who received placebo in double-blind Period 2 and did not enter rescue period, exposure to adalimumab is first dose of double-blind – first dose in OL Period 1 + 13 days. For subjects who received placebo in double-blind Period 2 and entered rescue period, exposure to adalimumab is the sum of (minimum of ((first dose in double-blind Period 2 – first dose in OL Period 1 + 13 days) or (first dose of rescue – first dose in OL Period 1)) and (last dose date – first dose of rescue period + 14 days)).

10.0 Efficacy Analysis

10.1 Efficacy Variables for DB Period 2

The following sections describe the efficacy variables that are collected and derived for this study's reporting purpose.

10.1.1 Ankylosing Spondylitis Disease Activity Score (ASDAS)

Parameters used for the calculation of ASDAS:

1. Patient's assessment of total back pain (BASDAI Question 2)
2. Patient global assessment of disease activity
3. Peripheral pain/swelling (BASDAI Question 3)
4. Duration of morning stiffness (BASDAI Question 6)
5. High-sensitivity C-reactive protein (hs-CRP) in mg/L.

Calculation of ASDAS:

$$ASDAS_{hs\text{-}CRP} = 0.121 \times \text{total back pain} + 0.110 \times \text{patient global} + 0.073 \times \text{peripheral pain/swelling} + 0.058 \times \text{duration of morning stiffness} + 0.579 \times \text{Ln(hs-CRP+1)}.$$

ASDAS score is categorized in to the following ASDAS Disease Activity States:

- ASDAS Inactive Disease: ASDAS < 1.3
- ASDAS Moderate Disease: 1.3 ≤ ASDAS < 2.1
- ASDAS High Disease: 2.1 ≤ ASDAS ≤ 3.5
- ASDAS Very High Disease: ASDAS > 3.5

ASDAS Response categories are defined as follows:

- ASDAS Major Improvement (a change from baseline ≤ –2.000)
- ASDAS Clinically Important Improvement (a change from baseline ≤ –1.100)

10.1.2 Bath Ankylosing Spondylitis Disease Activity Index (BASDAI)

BASDAI consists of the following 6 questions; Questions 1 through 5 have responses that can range from 0 (none) to 10 (very severe); Question 6 have response range from 0 (0 hr) to 10 (2 or more hrs), and 5 represents 1 hr.

Q1. How would you describe the overall level of fatigue/tiredness you have experienced?

Q2. How would you describe the overall level of AS neck, back or hip pain you have had?

Q3. How would you describe the overall level of pain/swelling in joints, other than neck, back or hips you have had?

Q4. How would you describe the overall level of discomfort you have had from any areas tender to touch or pressure?

Q5. How would you describe the overall level of morning stiffness you have had from the time you wake up?

Q6. How long does your morning stiffness last from the time you wake up?

Scoring of the BASDAI

BASDAI will be reported 0 to 10. The score has a maximum value of 10 and is calculated as follows:

BASDAI Score = 0.2 (Q1 + Q2 + Q3 + Q4 + Q5/2 + Q6/2)

If one of the 5 items (Questions 1 – Question 4, inflammation) is missing, then the score is the mean of the 4 non-missing items (total of 4 non-missing items divided by 4). If more than 1 of the 5 items is missing, then the BASDAI score is missing.

Note: Question 5 and Question 6 jointly constitute Item 5 (inflammation). If both Questions 5 and 6 are missing, and Questions 1 through 4 are non-missing, then only one item will be considered missing. The BASDAI score can still be calculated as the mean of Questions 1 – 4.

However, if, for example, both Question 6 and Question 1 are missing, then 2 items will be considered missing, as the inflammation calculation would be incomplete. The BASDAI score would then be considered missing in this case.

10.1.3 BASDAI50

BASDAI50 is a categorical response based on BASDAI that represents a 50% improvement in BASDAI.

10.1.4 ASAS20, ASAS40, ASAS 5/6, and ASAS Partial Remission

Parameters used for the ASAS responses:

1. Patient's Global Assessment – Represented by the PTGA-disease activity NRS score (0 to 10)
2. Pain – Represented by the patient's assessment of total back pain NRS score (0 to 10)
3. Function – Represented by the BASFI NRS score (0 to 10)

4. Inflammation – Represented by the mean of the 2 morning stiffness-related BASDAI NRS scores (mean of items 5 and 6 of the BASDAI [0 to 10])

ASAS20 Response

Improvement of ≥ 20% and absolute improvement of ≥ 1 unit (on a scale of 0 to 10) from Baseline in ≥ 3 of the above 4 domains, with no deterioration in the remaining domain (defined as a worsening of ≥ 20% and a net worsening of ≥ 1 unit).

ASAS40 Response

Improvement of ≥ 40% and absolute improvement of ≥ 2 units (on a scale of 0 to 10) from Baseline in ≥ 3 of the 4 domains above with no deterioration (defined as a net worsening of > 0 units) in the potential remaining domain.

ASAS Partial Remission

Absolute score of < 2 units for each of the 4 domains identified above.

ASAS 5/6 Response

20% improvement from Baseline in 5 out of the following 6 domains: BASFI, patient's assessment of total back pain, PTGA-disease activity, inflammation (mean of Items 5 and 6 of the BASDAI]) lateral lumbar flexion from BASMI, and hs-CRP.

10.1.5 Health Assessment Questionnaire Modified for the Spondyloarthropathies (HAQ-S)

The Health Assessment Questionnaire (HAQ-S) is a self-reported subject-oriented outcome measure. The Disability Index of HAQ-S for a subject is calculated as the mean of the following 8 category scores (range: 0 to 3): Dressing and Grooming, Rising, Eating, Walking, Hygiene, Reach, Grip, and Activities. Five additional items in the functional status measure have been included in the revised HAQ-S questionnaire, such as carrying heavy packages, sitting for long periods, ability to work at a flat topped table,

driving a car (if the patient has a driver's license or a car), ability to look in the rear view mirror, and ability to turn head to drive in reverse.

The highest score reported for any component question of the 8 categories determines the score for that category. A lower category score will be adjusted to 2 if aids and devices or help from another person were used in that category. A subject must have scores for at least 6 categories in order to compute the HAQ-S score.

10.1.6 Bath Ankylosing Spondylitis Functional Index (BASFI)

The Bath Ankylosing Spondylitis Functional Index (BASFI) consists of the following 10 questions, each with a response ranging from 0 (easy) to 10 (impossible).

1. Putting on your socks or tights without help or aids (e.g., sock-aid).
2. Bending forward from the waist to pick up a pen from the floor without an aid.
3. Reaching up to a high shelf without help or aids (e.g., helping hand).
4. Getting up out of an armless dining room chair without using your hands or any other help.
5. Getting up off the floor without help from lying on your back.
6. Standing unsupported for 10 minutes without discomfort.
7. Climbing 12 to 15 steps without using a handrail or walking aid. One foot on each step.
8. Looking over your shoulder without turning your body.
9. Doing physically demanding activities (e.g., physiotherapy, exercises, gardening, or sports).
10. Doing a full day's activities whether at home or at work.

Scoring of BASFI

The BASFI score will be derived based on the average of Questions 1 through 10. If up to 2 items are missing, corresponding scores will be replaced with the mean of the remaining non-missing items. If 3 or more items are missing, BASFI will be considered missing.

10.1.7 High Sensitivity C-Reactive Protein

The high-sensitivity C-reactive protein (hs-CRP) is a laboratory parameter and is considered an efficacy variable for the Axial Spondyloarthritis indication. It is a general marker of inflammation that is sensitive to acute changes in inflammatory response. The abnormal and normal values will be determined according to the normal ranges provided by the laboratory for hs-CRP.

10.1.8 Bath Ankylosing Spondylitis Metrology Index (Linear)

The Linear BASMI ($BASMI_{lin}$) composite score will be calculated using the BASMI components.

Scoring of $BASMI_{lin}$

The table below presents the components of $BASMI_{lin}$ and assessment ranges for score.

	Score		
	0	**Between 0 and 10**	**10**
Lateral Lumbar flexion (cm)	$A \geq 21.1$	$(21.1 – A)/2.1$	$A \leq 0.1$
Tragus to wall distance (cm)	$A \leq 8$	$(A – 8)/3$	$A \geq 38$
Lumbar flexion (modified Schober) (cm)	$A \geq 7.4$	$(7.4 – A)/0.7$	$A \leq 0.4$
Intermalleolar distance (cm)	$A \geq 124.5$	$(124.5 – A)/10$	$A \leq 24.5$
Cervical rotation (°)	$A \geq 89.3$	$(89.3 – A)/8.5$	$A \leq 4.3$

$BASMI_{lin}$ = Assessment measurements for tragus to wall, cervical rotation and lateral lumbar flexion are the means of the left and right measurements

A = assessment measurement

Scores for each assessment range from 0 to 10, and the $BASMI_{lin}$ total score will be the average of the 5 assessment scores. If 1 item is missing, the $BASMI_{lin}$ will be calculated as the mean of remaining 4 items. Hence, the range of the $BASMI_{lin}$ total score should be between 0 and 10. If 2 or more items are missing, then the $BASMI_{lin}$ score will be considered missing.

10.1.9 Maastricht Ankylosing Spondylitis Enthesitis Score (MASES)

The Maastricht Ankylosing Spondylitis Enthesitis Score (MASES) will be measured at the designated study visits listed in Table 15 of Appendix D. To assess the presence or absence of enthesitis at 13 different sites, noting the subjects' responses.

Scoring of MASES

The following left and right locations will be graded for presence (1) or absence (0) of tenderness: 1st Costochondral joint, 7th Costochondral joint, Posterior Superior Iliac Spine, Anterior Superior Iliac Spine, Iliac Crest, Proximal Insertion of Achilles tendon; the 5th Lumbar Spinous process will also be graded for tenderness yielding a total score ranging 0 – 13. If one or more locations are missing, the score may be calculated using available data.

10.1.10 Plantar Fascia Enthesitis

Evaluation for plantar fascia enthesitis will be conducted at the designated study visits listed in Table 15 of Appendix D to assess the presence or absence of plantar fascia enthesitis on the right and left feet. Frequency and percentage of answers to the absence will be summarized.

10.1.11 Dactylitis

The Dactylitis score will be the total count of the presence and absence of dactylitis of each digit (range 0 – 20). If dactylitis was not measured at one or more locations, the score may be calculated using available data.

10.1.12 Anterior Uveitis

The frequency and percentage of anterior uveitis presence and absence responses will be summarized for the right eye, the left eye, and total by visit. Frequency and percentage of present anterior uveitis responses will be summarized for new and recurrent onset by visit. New uveitis presence is defined as any new occurrence in either eye for subjects without previous uveitis or a new onset in a different eye in those with previous uveitis.

10.1.13 36-Item Short Form (SF-36v2) Questionnaire

The 36-Item Short Form, Version 2 (SF-36v2) Questionnaire with 4 week recall will be completed by the subject at the designated study visits listed in Table 15 of Appendix D. The SF-36v2 health survey consists of 36 general health questions. It has 2 components: physical and mental. For each component, a transformed summary score is calculated using 8 sub-domains: physical functioning, role-physical, bodily pain, general health, vitality, social functioning, role-emotional, and mental health. The SF-36v2 coding and scoring methods are described in Appendix A.

10.1.14 EuroQol EQ-5D

The EuroQol (EQ-5D) measures quality of life (QOL) and is collected at the designated study visits listed in Table 15 of Appendix D. The prevailing preference-based scoring function for the EuroQol Group's EQ-5D was derived from the general population of the United Kingdom (UK) in the early 1990s. The questionnaire can be found in Appendix T of the Study M13-375 Protocol (Amendment 2).

10.1.15 Work Productivity and Activity Impairment Questionnaire: Axial Spondyloarthritis, Version 2.0 (WPAI-Axial Spondyloarthritis)

The WPAI-Axial SpA questionnaire was developed to measure the effect of overall health and specific symptoms on productivity at work and outside of work. It consists of 6 questions. A lower WPAI-Axial SpA score indicates an improvement. The WPAI

Axial SpA is collected at the designated study visits listed in Table 15 of Appendix D. The The WPAI-Axial SpA coding and scoring methods are described in Appendix C.

10.1.16 Joint Evaluation

Anatomical joints are evaluated for swelling and tenderness at the designated study visits listed in Table 15 of Appendix D. The 34 anatomical joints in Table 10 are assessed in both the left and right side of the body for tenderness. The joints to be examined for swelling are the same as those examined for tenderness, with the exception of the hip joints.

Table 10. Anatomical Joints Assessed

Temporomandibular	Sternoclavicular	Acromio-Clavicular	Shoulder
Elbow	Wrist	Metacarpophalangeal I	Metacarpophalangeal II
Metacarpophalangeal III	Metacarpophalangeal IV	Metacarpophalangeal V	Thumb Interphalangeal
Proximal Interphalangeal II	Proximal Interphalangeal III	Proximal Interphalangeal IV	Proximal Interphalangeal V
Distal Interphalangeal II	Distal Interphalangeal III	Distal Interphalangeal IV	Distal Interphalangeal V
Hip[a]	Knee	Ankle	Tarsus
Metatarsophalangeal I	Metatarsophalangeal II	Metatarsophalangeal III	Metatarsophalangeal IV
Metatarsophalangeal V	Great Toe/Hallux	Interphalangeal II	Interphalangeal III
Interphalangeal IV	Interphalangeal V		

a. Hip joints are not assessed for swelling.

At each study visit, a joint evaluator will assess whether a particular joint is "tender or painful," where presence of tenderness is scored as "1" and the absence of tenderness is scored as "0," provided the joint was not replaced ("9") or could not be assessed ("NA") due to other reasons (e.g., post-corticosteroid joint injection). The total tender joint count (TJC68) will be derived as the sum of all "1s" thus collected and no penalty will be considered for the joints not assessed or those which have been replaced. A similar method will be followed for the derivation of total swollen joint count (SJC66). Thus, the range for TJC68 will be 0 to 68 and 0 to 66 for SJC66.

10.1.17 Other Efficacy Variables

- Patient's Global Assessment of Disease Activity NRS
  - The range is 0 to 10 with no activity being indicated by 0 and severe activity by 10.
- Patient's Assessment of Total Back Pain NRS
  - The range is 0 to 10 with no pain being indicated by 0 and most severe pain by 10.
- Physician's Global Assessment of Disease Activity NRS
  - The range is 0 to 10 with no activity being indicated by 0 and severe activity by 10.
- Patient's Assessment of Nocturnal Back Pain NRS
  - The range is 0 to 10 with no pain being indicated by 0 and worst possible pain by 10.
- Patient's Global Assessment of Pain NRS
  - The range is 0 to 10 with no pain being indicated by 0 and severe pain by 10.

10.1.18 MRI of the Spine and Sacroiliac Joints (SPARCC Method)

All subjects will have an MRI evaluation of the cervical, thoracic and lumbar regions of the spine as well as the sacroiliac joints at screening visit and study completion visit. Images will be sent to the central imaging laboratory designated by the Sponsor. Procedures for performing MRIs and scoring methods are provided in the "AbbVie Protocol M13-375 (19539) Imaging Charter" Version 3.0 from BioClinica Inc. (released on 27 December 2013). The SpondyloArthritis Research Consortium of Canada (SPARCC) scores for spine and sacroiliac (SI) joints are calculated by adding up the dichotomous outcomes from evaluations of the presence, depth, and intensity of bone marrow edema lesions of the spine and SI joints, respectively. The maximum SPARCC score for spine is 108 and that for SI joints is 72.

10.2 Visit Windows for Variables

Each variable (safety or efficacy) will be summarized and analyzed according to the visit windows as described in Section 6.0.

10.3 Efficacy Analysis for OL Period 1

The efficacy results for OL Period 1 will be presented up to and including Week 28.

In addition, the following variables will be assessed at each scheduled visit and including Week 28:

- ASDAS (continuous score)
- ASDAS Inactive Disease (ASDAS < 1.300)
- ASDAS Moderate Disease (1.3 ≤ ASDAS < 2.1)
- ASDAS High Disease (2.1 ≤ ASDAS ≤ 3.5)
- ASDAS Very High Disease (ASDAS > 3.5)
- ASDAS Major Improvement (a change from baseline ≤ –2.000)
- ASDAS Clinically Important Improvement (a change from baseline ≤ –1.100)
- ASAS20, ASAS40, ASAS 5/6, and ASAS Partial Remission
- Bath AS Disease Activity Index (BASDAI)
- BASDAI50
- Inflammation/Morning Stiffness (mean of BASDAI Questions 5 and 6)
- Physician's Global Assessment of Disease Activity
- Patient's Global Assessment of Disease Activity
- Patient's Assessment of Total back pain
- Patient's Assessment of Nocturnal Back Pain
- Patient's Global Assessment of Pain
- Bath AS Functional Index (BASFI)
- hs-CRP (mg/L)
- hs-CRP (normal vs. above upper limit of normal)
- Linear Bath AS Metrology Index ($BASMI_{lin}$)

- Swollen Joint Count (66 joints)
- Tender Joint Count (68 joints)
- MASES
- Plantar fascia enthesitis
- Dactylitis
- Anterior Uveitis
- HAQ-S
- Short Form-36v2 Health Survey questionnaire
- EQ-5D
- Work Productivity and Activity Impairment – Axial Spondylorarthritis (WPAI-Axial SpA)

In addition to the above variables, each of the variables mentioned in Section 10.1 will also be summarized at each time point of the analysis window as shown in Table 1 and Table 2 as applicable.

Discrete variables will be summarized using count and percentages. Continuous efficacy variables will be summarized by summary statistics (number of subjects, mean, 95% confidence interval, standard deviation, first quartile, median, third quartile, minimum, maximum) and change from Baseline. This will be done for the observed values.

10.4 Efficacy Analysis for DB Period 2

10.4.1 Primary Efficacy Variable

10.4.1.1 Analysis of Primary Efficacy Endpoint

The primary efficacy endpoint is the proportion of subjects who do not experience a flare in Period 2 by Week 68 of the study, where a flare is defined as having any 2 consecutive study visits with ASDAS $\geq$ 2.100. Subjects who discontinue during the DB Period 2 or who are missing flare data at Week 68 will be treated as non-responders (flare) according to the "non responder imputation" (NRI) criterion. The proportion of subjects who do not experience flare will be compared and summarized by randomized

treatment group using the mITT population (excluding post rescue data). The null hypothesis for this comparison is that the response rates between the adalimumab and placebo groups are not different; the alternative hypothesis is that the response rates are different. The treatment comparison will be performed using a two-sided Pearson's chi-square test or Fisher's exact test (if ≥ 25% of the cells have expected counts less than 5) with a significance level $\alpha = 0.05$.

10.4.1.2 Sensitivity Analysis of Primary Efficacy Endpoint

Sensitivity analyses of the primary efficacy endpoint will be conducted using an observed case analysis for the proportion of subjects not experiencing a flare up to Week 68 for the mITT (exclude post rescue data). In addition, the primary efficacy endpoint will also be summarized and analyzed using NRI for PPP (excluding post rescue data). The treatment comparison will be performed using a two-sided Pearson's chi-square test or Fisher's exact test (if ≥ 25% of the cells have expected counts less than 5) with a significance level $\alpha = 0.05$.

10.4.2 Analyses of Secondary Efficacy Endpoints

Unless otherwise stated, all statistical comparisons of secondary efficacy variables will be conducted using the two-sided test with significance level $\alpha = 0.05$.

Time to event analysis will be performed for following variables:

- Time to flare defined as ASDAS ≥ 2.100 at 2 consecutive visits
- Time to partial flare defined as ASDAS ≥ 1.300 but < 2.100 at 2 consecutive visits

Time to flare/partial flare will be analyzed using a Cox proportional hazards model with treatment as a factor and baseline ASDAS as a covariate. Kaplan-Meier curves will also be provided for flare/partial flare of two treatment groups. The comparison will be analyzed by the log-rank test.

At each time point of the visit window as described in Table 3, Table 4, Table 5, Table 6 and Table 7 variables. The efficacy measurements will be summarized and analyzed, treatment comparison will be performed for the following variables and other variables mentioned in Section 10.1.

- ASDAS Inactive Disease
- ASDAS Major Improvement
- ASDAS Clinically Important Improvement
- ASAS20, ASAS40, ASAS 5/6, and ASAS Partial Remission
- BASDAI50
- HAQ-S

For the Rescued Population, the following variables will be summarized descriptively at 12 weeks after initiation of Rescue Therapy, as well as for other post-rescue visits where appropriate.

- ASDAS Inactive Disease (ASDAS < 1.300)
- ASDAS Major Improvement (a change from baseline ≤ –2.000)
- ASDAS Clinically Important Improvement (a change from baseline ≤ –1.100)
- ASAS20, ASAS40, ASAS 5/6, and ASAS Partial Remission
- Bath AS Disease Activity Index 50 (BASDAI50)
- HAQ-S

Discrete efficacy variables will be summarized using count and percentages and will be compared between adalimumab and placebo groups during the double-blind period using Pearson's chi-square test or Fisher's exact test (if ≥ 25% of the cells have expected counts less than 5).

Continuous efficacy variables will be summarized by summary statistics (number of subjects, mean, 95% confidence interval, standard deviation, first quartile, median, third quartile, minimum, maximum). A mixed-effect repeated measures model (MMRM)

including the fixed effects of treatment, visit, baseline score and treatment by visit interaction as covariate will be used for observed values. The least square mean and 95% confidence interval obtained from the model will be presented.

10.5 Handling of Multiplicity

No multiplicity adjustment will be applied to the efficacy analyses.

10.6 Efficacy Subgroup Analysis

Analyses for the primary efficacy endpoint based on the subgroups of gender (male, female), age (< 40, 40 – 65, > 65 years), race (white, non-white), baseline hs-CRP (normal vs. elevated), baseline HLA-B27 status (positive, negative), symptom duration (< 5, ≥ 5 years) will be performed for DB Period 2. Subgroups may be combined in case of small number of subjects. Analyses of treatment by each subgroup interaction for the primary efficacy analysis will be performed using logistic regression. Exploratory multiple logistic regression analysis may be performed for the aforementioned list of subgroup variables using stepwise selection. The analysis will be performed for the mITT population (excluding post rescue data); missing data will be imputed using NRI.

11.0 Safety Analysis

11.1 General Considerations

Safety analyses will include reporting of adverse events, laboratory, and vital sign measurements.

There will be 4 sets of safety analyses to be performed.

1. All safety analyses will be performed using the FAS for OL Period 1.
2. All safety analyses will be performed using the DB Safety Population (excluding OL rescue therapy data) for DB Period 2.
3. All safety analyses will be performed using the Rescued Population during the rescue period.

4. The following safety analyses will be performed for the Any ADA population starting from the first dose of the OL Period 1 to the last dose of DB/OL rescue ADA + 70 days. Data from the DB placebo exposure will be excluded from this analysis.
    a. All AE analyses
    b. Potentially clinically significant values/abnormalities in laboratory data and vital signs

Adverse event analyses will include all treatment-emergent AEs up to 70 days after adalimumab dose. The following summary statistics for laboratory, and vital signs measurements will be presented for subjects who have both Baseline and post-baseline values for laboratory tests and vital signs: the mean value at Baseline and at each respective protocol specified visit, and the mean, standard deviation, and median for changes from Baseline. Categorical data will be summarized using frequencies and percentages. The number of non-missing values will be given.

11.2 Analysis of Adverse Events

Treatment-emergent adverse events (TEAEs) will be tabulated. In addition, serious adverse events (SAEs) collected between the signing of the informed consent and the first dose will be reported separately.

11.2.1 Treatment-Emergent Adverse Events

Treatment-emergent AEs are defined as AEs that begin either on or after the first dose of the study medication, and up to within 70 days after the last dose of the study medication, except for those subjects that initiate adalimumab therapy not supplied in the context of the clinical trial after the end of study participation. Events where the onset date is the same as the study drug start date are assumed to be treatment-emergent. If an incomplete onset date was collected for an adverse event, the event will be assumed to be treatment emergent unless there is other evidence that confirms the event was not treatment-emergent (e.g., the event end date was prior to the study drug start date).

Adverse event data will be summarized and presented using primary MedDRA system organ classes (SOCs) and preferred terms (PTs) according to Version 19.1 or higher of the MedDRA coding dictionary. The system organ classes will be presented in alphabetical order and the preferred terms will be presented in alphabetical order within each system organ class.

11.2.1.1 Adverse Event Overview

The number and percentage of subjects experiencing TEAEs will be summarized for the following AE categories:

- Any adverse event (AE)
- Any serious AE
- Any AE leading to discontinuation
- Any severe event
- Any AE related to study drug
- Any AE leading to death

AEs of Special Interest (AESI)

- Any infection
- Any serious infection
- Any legionella infection
- Any diverticulitis
- Any opportunistic infection excluding oral candidiasis and TB
- Any oral candidiasis
- Any tuberculosis
- Any active tuberculosis
- Any latent tuberculosis
- Any parasitic infection
- Any reactivation of hepatitis B
- Any progressive multifocal leukoencephalopathy (PML)

- Any malignancy
- Any lymphoma
- Any hepatosplenic T-cell lymphoma (HSTCL)
- Any non-melanoma skin cancer (NMSC)
- Any melanoma
- Any leukaemia
- Any malignancy other than lymphoma, HSTCL, leukaemia, NMSC or melanoma
- Any allergic reaction including angioedema/anaphylaxis
- Any lupus-like reactions and systemic lupus erythematosus
- Any vasculitis
- Any cutaneous vasculitis
- Any non-cutaneous vasculitis
- Any sarcoidosis
- Any autoimmune hepatitis
- Any myocardial infarction
- Any cerebrovascular accident
- Any CHF
- Any pulmonary embolism
- Any interstitial lung disease
- Any intestinal perforation
- Any pancreatitis
- Any Stevens-Johnson Syndrome
- Any erythema multiforme
- Any worsening/new onset of psoriasis
- Any demyelinating disorder
- Any amyotrophic lateral sclerosis
- Any reversible posterior leukoencephalopathy syndrome (RPLS)
- Any hematologic disorders including pancytopenia

- Any liver failure and other liver event
- Any adalimumab administration related medication error
- Any injection site reaction

Additional AEs may be considered for tabulation/summary based on recommendations from Clinical and Safety as deemed appropriate. Any event with an unknown severity will be considered severe and any AE with an unknown relationship will be considered drug related.

11.2.1.2 Adverse Events by System Organ Class and Preferred Term

Select TEAEs will be summarized and presented using primary Medical Dictionary for Regulatory Activities (MedDRA) system organ classes (SOCs) and preferred terms (PTs) according to Version 19.1 or later of the MedDRA coding dictionary. The SOCs will be presented in alphabetical order, and the PTs will be presented in alphabetical order within each SOC.

A subject who reports more than 1 AE in different SOCs will be counted only once in the overall total. A subject who reports 2 or more different PTs, which are in the same SOC, will be counted only once in the SOC total. A subject who reports more than 1 AE with the same PT will be counted only once for that PT using the most extreme incident (i.e., most severe for the severity tables and most related for the relationship tables).

11.2.1.3 Adverse Events by Maximum Severity

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

11.2.1.4 Adverse Events by Maximum Relationship

TEAEs will also be summarized by maximum relationship to study drug, as assessed by the investigator. If a subject has an AE with an unknown relationship, then the subject will be counted in the relationship category of "unknown," even if the subject has another occurrence of the same event with a relationship present. The only exception is if the subject has another occurrence of the same AE with a relationship assessment of probably related or possibly related. In this case, the subject will be counted under the probably related or the possibly related category, respectively.

11.2.1.5 Adverse Events by "At Least Possibly Related" Relationship

TEAEs will also be summarized by relationship defined by "at least possibly related" to study drug, as assessed by the investigator. If a subject has an AE with an unknown relationship, then the subject will be counted in as 'related.'

11.2.1.6 Serious Adverse Events (Including Deaths) and Adverse Events Leading to Study Drug Discontinuation

All serious adverse events (SAEs), deaths, and adverse events leading to discontinuation of study drug will be listed. The number and percentage of subjects experiencing SAEs (including deaths) and adverse events leading to discontinuation of study drug will be tabulated according to the primary MedDRA SOC and MedDRA PT for each treatment group.

11.2.1.7 Frequent (≥ 5%) Adverse Events by System Organ Class and Preferred Term

TEAEs occurring for more than ≥ 5% (based on at least 1 of the treatment arms) of subjects will be summarized and presented using primary MedDRA preferred terms (PTs) by decreasing frequency.

11.2.1.8 Adverse Events of Special Interests

The special TEAEs listed in Section 11.2.1.1 will be summarized and presented using primary MedDRA system organ classes (SOCs) and preferred terms (PTs):

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

11.2.1.9 Adverse Events by 100 Patient Years

TEAEs will be summarized by event rate per 100 patient years, defined as

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

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

11.3 Analysis of Laboratory Data

11.3.1 Variables and Units

All laboratory parameters to be collected in the study are listed in the table below. Laboratory parameters will be reported using the standard international (SI) units. The laboratory test short names and units of all laboratory variables for the study are as follows:

	Units
Hematology	
White Blood Cell (WBC) Count	$\times 10^9$/L
Red Blood Cell (RBC) Count	$\times 10^{12}$/L
Hemoglobin	g/L
Hematocrit	Ratio
Platelet count	$\times 10^9$/L
Neutrophils	$\times 10^9$/L
Basophils	$\times 10^9$/L
Eosinophils	$\times 10^9$/L
Lymphocytes	$\times 10^9$/L
Monocytes	$\times 10^9$/L
Bands	$\times 10^9$/L
Chemistry	
Total Bilirubin	µmol/L
Alkaline Phosphatase (ALP)	U/L
Serum glutamic oxaloacetic transaminase (SGOT/AST)	U/L
Serum glutamic pyruvic transaminase (SGPT/ALT)	U/L
Total Protein	g/L
Albumin	g/L
Glucose	mmol/L
Cholesterol	mmol/L
Triglycerides	mmol/L
Blood Urea Nitrogen (BUN)	mg/dL
Creatinine	µmol/L
Uric acid	µmol/L
Sodium	mmol/L
Potassium	mmol/L
Calcium	mmol/L
Inorganic Phosphorus	mmol/L

	Units
Urinalysis	
Specific Gravity	-
pH	-
Protein	g/L
Glucose	mmol/L
Ketones	mmol/L
Blood	
Nitrite	-
Serology	
hs-CRP	mg/L
Pregnancy test (Human Chorionic Gonadatropin)	-
Anti-Nuclear Antibody (ANA)	
anti-dsDNA reflex only if ANA is positive	-

11.3.2 Statistical Methods

11.3.2.1 Analysis for Continuous Laboratory Data for OL Period 1

Mean changes from Baseline to post-baseline visits of OL Period 1 for FAS will be summarized with the baseline mean, the visit mean, the mean change from baseline, standard deviation, and median. No statistical comparison will be taken for OL Period 1.

11.3.2.2 Shift Table Analyses for OL Period 1

The laboratory data will be categorized as low, normal, or high based on the normal ranges of the laboratory used in this study. The shift tables will tabulate the number and percentage of subjects with Baseline values below/within/above the normal range versus minimum/maximum/final values below/within/above the normal range.

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

- < 1.5 × ULN,
- ≥ 1.5 × ULN TO < 3 × ULN,
- ≥ 3 × ULN TO < 5 × ULN,
- ≥ 5 × ULN TO < 8 × ULN, and
- ≥ 8 × ULN,

where ULN is the upper normal limit.

Shift tables showing shift from Baseline to maximum values will be presented using these 5 categories.

11.3.2.3 Analysis for Potentially Clinically Significant Laboratory Values for OL Period 1

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

- ALT ≥ 2.5 × ULN, or
- AST ≥ 2.5 × ULN, or
- Alkaline phosphatase ≥ 2.5 × ULN, or
- Total bilirubin ≥ 1.5 × ULN.

For selected laboratory parameters, a listing of all subjects with any laboratory determination during each Period meeting Common Toxicity Criteria (CTC) (Version 3) of Grade 3 or higher will be provided. For each of these subjects, the whole course of the respective parameter will be listed.

11.3.2.4 Analysis for Continuous Laboratory Data for DB Period 2

Mean changes from Baseline to post-baseline visits of Double blind safety population will be summarized with the baseline mean, the visit mean, the mean change from baseline, standard deviation, and median.

11.3.2.5 Shift Table Analyses for DB Period 2

The laboratory data will be categorized as low, normal, or high based on the normal ranges of the laboratory used in this study. The shift tables will tabulate the number and percentage of subjects with Baseline values below/within/above the normal range versus minimum/maximum/final values below/within/above the normal range.

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

- $< 1.5 \times$ ULN,
- $\geq 1.5 \times$ ULN TO $< 3 \times$ ULN,
- $\geq 3 \times$ ULN TO $< 5 \times$ ULN,
- $\geq 5 \times$ ULN TO $< 8 \times$ ULN, and
- $\geq 8 \times$ ULN.

where ULN is the upper limit of normal.

Shift tables showing shift from Baseline to minimum/maximum/final values will be presented using the above 5 categories.

11.3.2.6 Analysis for Potentially Clinically Significant Laboratory Values for DB Period 2

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

- ALT $\geq 2.5 \times$ ULN, or
- AST $\geq 2.5 \times$ ULN, or
- Alkaline phosphatase $\geq 2.5 \times$ ULN, or
- Total bilirubin $\geq 1.5 \times$ ULN.

For selected laboratory parameters, a listing of all subjects with any laboratory determination during each Period meeting Common Toxicity Criteria (CTC) (Version 3) of Grade 3 or higher will be provided. For each of these subjects, the whole course of the respective parameter will be listed.

11.4 Analysis of Vital Signs and Weight

11.4.1 Variables and Criteria Defining Abnormality

Vital sign variables include sitting systolic blood pressure, sitting diastolic blood pressure, pulse, body temperature, and weight. The criteria for potentially clinically significant vital sign findings are presented in Table 11. The number and percentage of subjects meeting the criteria for potentially clinically significant vital sign values will be summarized by the actual treatment groups.

Table 11. 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 bpm from Baseline
	High	Value ≥ 120 bpm and/or increase ≥ 15 bpm from Baseline

11.4.2 Statistical Methods for Vital Sign and Weight

11.4.2.1 Statistical Methods for Vital Sign and Weight of OL Period 1

Mean changes from Baseline to post-baseline visits will be summarized with the baseline mean, the visit mean, mean change from baseline, standard deviation, and median. The number and percentage of subjects meeting the criteria for potentially clinically significant vital sign values will be summarized. No statistical comparison will be taken

for OL Period 1. Vital sign results satisfying the criteria for potentially clinically significant findings will be identified in a listing.

11.4.2.2 Statistical Methods for Vital Sign and Weight of DB Period 2

Mean changes from Baseline to post baseline visits will be summarized with the baseline mean, the visit mean, mean change from baseline, standard deviation, and median. The number and percentage of subjects meeting the criteria for potentially clinically significant vital sign values will be summarized. Vital sign results satisfying the criteria for potentially clinically significant findings will be identified in a listing.

11.5 Analysis of ECG Parameters

No post-baseline ECGs were obtained for this study.

11.6 Analysis for Other Safety Variables

Not applicable for the current planned analysis.

11.7 Safety Subgroup Analysis

Not applicable for the current planned analysis.

12.0 Special Statistical Topics

Not applicable for the current planned analysis.

13.0 Appendix

Appendix A	Scoring the SF-36v2
Appendix B	EQ-5D Scoring Algorithm
Appendix C	Scoring the WPAI-Axial SpA
Appendix D	Visit Schedule per Protocol

Appendix A. Scoring the SF-36v2

The SF-36v2™ consists of 36 questions covering 8 health domains: physical functioning, bodily pain, role limitations due to physical problems, role limitations due to emotional problems, general health perceptions, mental health, social function, and vitality. The SF 36v2™ will be scored for the 8 sub-domains according to the standard SF-36v2™ scoring algorithms (0 to 100 scale) explained in the SF-36v2™ Manual and Interpretation Guide (Ware et al 1993). The physical component summary (PCS) and mental component summary (MCS) will be scored according to the standard SF-36v2™ scoring algorithm (0 to 100 scale) explained in the SF-36v2™ Physical and Mental Health Summary Scales Manual (Ware et al 1994). A higher score indicates a better health state.

Table 12. Coding and Scoring for the 8 SF-36 (Version 2) Scales

	Item	Coding	Range of Raw Score	Normalization*
Physical functioning	*Items 3a to 3j*	Yes, limited a lot = 1 Yes, limited a little = 2 No, not limited = 3	10 – 30	(S – 10)/20 × 100
Role-physical	*Items 4a to 4d*	All of the time = 1 Most of the time = 2 Some of the time = 3 A little of the time = 4 None of the time = 5	4 – 20	(S – 4)/16 × 100
Bodily pain	*Item 7*	None = 6 Very mild = 5.4 Mild = 4.2 Moderate = 3.1 Severe = 2.2 Very severe = 1	2 – 12	(S – 2)/10 × 100
	Item 8 if both Items 7 and 8 are answered	Not at all and Item 7 equals 'None' = 6 Not at all and Item 7 not equal 'None' = 5 A little bit = 4 Moderately = 3 Quite a bit = 2 Extremely = 1		
	Item 8 if Item 7 is not answered	Not at all = 6 A little bit = 4.75 Moderately = 3.5 Quite a bit = 2.25 Extremely = 1		

	Item	Coding	Range of Raw Score	Normalization*
General health	*Item 1*	Excellent = 5 Very good = 4.4 Good = 3.4 Fair = 2 Poor = 1	5 – 25	(S – 5)/20 × 100
	Items 11a and 11c	Definitely true = 1 Mostly true = 2 Don't know = 3 Mostly false = 4 Definitely false = 5		
	Items 11b and 11d	Definitely true = 5 Mostly true = 4 Don't know = 3 Mostly false = 2 Definitely false = 1		
Vitality	*Items 9a and 9e*	All of the time = 5 Most of the time = 4 Some of the time = 3 A little of the time = 2 None of the time = 1	4 – 20	(S – 4)/16 × 100
	Items 9g and 9i	All of the time = 1 Most of the time = 2 Some of the time = 3 A little of the time = 4 None of the time = 5		
Social functioning	*Item 6*	Not at all = 5 Slightly = 4 Moderately = 3 Quite a bit = 2 Extremely = 1	2 – 10	(S – 2)/8 × 100
	Item 10	All of the time = 1 Most of the time = 2 Some of the time = 3 A little of the time = 4 None of the time = 5		

	Item	Coding	Range of Raw Score	Normalization*
Role-emotional	*Items 5a to 5c*	All of the time = 1 Most of the time = 2 Some of the time = 3 A little of the time = 4 None of the time = 5	3 – 15	(S – 3)/12 × 100
Mental health	*Items 9b, 9c and 9f*	All of the time = 1 Most of the time = 2 Some of the time = 3 A little of the time = 4 None of the time = 5	5 – 25	(S – 5)/20 × 100
	Items 9d and 9h	All of the time = 5 Most of the time = 4 Some of the time = 3 A little of the time = 2 None of the time = 1		

* S = raw score = sum of item scores after coding.

The standardization of each scale score is shown in Table 13.

Table 13. Standardization of the SF-36 (Version 2) Scale Scores

Item	Standardization (Z-Score)
Physical functioning	PF_Z = (PF-83.29094)/23.75883
Role-physical	RP_Z = (RP-82.50964)/25.52028
Bodily pain	BP_Z = (BP-71.32527)/23.66224
General health	GH_Z = (GH-70.84570/20.97821
Vitality	VT_Z = (VT-58.31411)/20.01923
Social functioning	SF_Z = (SF-84.30250)/22.91921
Role-emotional	RE_Z = (RE-87.39733)/21.43778
Mental heath	MH_Z = (MH-74.98685)/17.75604

Note: Based on the 1998 general United States population.

SF-36 physical and mental component transformed scores are calculated using the conventions in Table 14.

Table 14. SF-36 Aggregate Component and Transformed Scores

Score	Calculation
Aggregate Standardized Score	
Physical component	AGG_PHYS = (PF_Z * 0.42402) + (RP_Z * 0.35119) + (BP_Z * 0.31754) + (GH_Z * 0.24954) + (VT_Z * 0.02877) + (SF_Z * –0.00753) + (RE_Z * –0.19206) + (MH_Z * –0.22069)
Mental component	AGG_MENT = (PF_Z * –0.22999) + (RP_Z * –0.12329) + (BP_Z * -0.09731) + (GH_Z * –0.01571) + (VT_Z * 0.23534) + (SF_Z * 0.26876) + (RE_Z * 0.43407) + (MH_Z * 0.48581)
Transformed Summary	
Transformed physical	PCS = 50 + (AGG_PHYS * 10)
Transformed mental	MCS = 50 + (AGG_MENT * 10)

Missing Data in an Individual Questionnaire

In the event that data is missing for an individual item from the sub-domains of the SF-36v2™ Health Status Survey, the average value of the completed items in the corresponding sub-domain will be used as an estimate of the missing item. If more than 50 percent of the items from a sub-domain are missing for an individual questionnaire, the corresponding deficient sub-domain(s) will be excluded from analyses. For example, if only 3 of the 5 general health questions are answered, the mean of those 3 answers will be used to fill in the responses of the remaining 2 general health questions. However, if only 1 or 2 general health questions are answered, the general health score would be set to missing. In addition, if the respondent is missing any 1 of the 8 scale scores, then the physical and the mental component scores will be set to missing.

Each of the 8 scales and the 2 components (physical and mental) will be reported using norm-based scoring as per the standard guidelines adopted by AbbVie.

Appendix B. EQ-5D Scoring Algorithm

The EQ-5D self-reported questionnaire (EQ-5D) consists of 2 modules comprising the EQ-5D descriptive system and the EQ VAS. EQ-5D is a standardized instrument for use as a measure of general health related quality of life. It provides a descriptive profile and a single index value for health status.

Descriptive System

The EQ-5D descriptive system comprises 5 dimensions of health (mobility, self-care, usual activities, pain/discomfort anxiety/depression) to describe patient's current health state. Each dimension comprises 5 levels (no problems, slight problems, moderate problems, severe problems, unable to perform activity) with corresponding numeric scores 1, 2, 3, 4 and 5. Only a single response is required for each dimension. A unique EQ-5D health state is defined by combining 1 level from each of the 5 dimensions.

A total of 3125 possible health states are defined in this way. Each state is referred to as a 5-digit code. For example, State 11111 indicates no problem with any of the fine dimensions, while State 11335 indicates no problems with mobility or self-care, some problems with performing usual activities, moderate pain or discomfort, and extreme anxiety or depression.

UK Version

If a score of other than 1 is chosen for an item, then the weights below should be subtracted from the constant. The weighted average of slope (a fixed coefficient 0.9675) will be multiplied by the sum of the five decrements when calculating values for all health states.

	Item Weights				
Dimensions	**1**	**2**	**3**	**4**	**5**
Mobility	0	0.051	0.063	0.212	0.275
Self-Care	0	0.057	0.076	0.181	0.217
Usual Activity	0	0.051	0.067	0.174	0.190
Pain/Discomfort	0	0.060	0.075	0.276	0.341
Anxiety/Depression	0	0.079	0.104	0.296	0.301
Constant	1				

Based on this scoring system, the predicted value for State 23245 is

1-0.9675*(0.051 + 0.076 + 0.051 + 0.276 + 0.301) = 0.270

EQ VAS

The EQ VAS records the respondent's self-rated health status on a vertical graduated (0 to 100) visual analogue scale with 0 being the "Worst Imaginable Health State" and 100 being "Best Imaginable Health State." It generates a self-rating of current health related quality of life. It should be used with the 5-digit health state classification to build a composite picture of the respondent's health status.

Reference:

Feng y, Devlin N, Shah k, et al. New methods for modelling EQ-5D-5L value sets: an application to English data, January 2016, OHE research papers.

Devlin N, Shah K, Feng Y, et al. Valuing health-related quality of life: an EQ-5D-5L value set for England, January 2016, OHE research papers.

Appendix C. Scoring the WPAI-Axial SpA

The Work Productivity and Activity Impairment Questionnaire: Axial Spondylorarthritis, V2.0 (WPAI-Axial SpA) was developed to measure the effect of overall health and specific symptoms on productivity at work and outside of work. A description of the six questions asked in the WPAI-Axial SpA is as follows:

Questions:

Q1. Currently employed? (Yes/No).

Q2. Hours missed from work due to Axial Spondyloarthritis.

Q3. Hours missed due to other reasons.

Q4. Hours actually worked.

Q5. Degree Axial Spondyloarthritis affected productivity while working.

Q6. Degree Axial Spondyloarthritis affected regular activities other than job.

The following 6 measures will be derived based on the responses from the 6 questions. The 4 main impairment scores (S1 to S4) are expressed as *percent impairment* based on the above questions.

Scores:

S0. Employment: defined below

S1. Absenteeism: Percent work time missed due to SpA:

$$100 \times \left[\frac{Q2}{Q2 + Q4} \right]$$

S2. Presenteeism: Percent impairment while working due to SpA:

$$100\times\left[\frac{Q5}{10}\right]$$

S3. Percent overall work impairment due to SpA:

$$100\times\left[\frac{Q2}{Q2+Q4}+\left\{1-\frac{Q2}{Q2+Q4}\right\}\times\frac{Q5}{10}\right]$$

S4. Percent activity impairment due to SpA:

$$100\times\left[\frac{Q6}{10}\right]$$

S5. Did subject miss work (defined below). This is needed to derive the proportion of subjects who missed work.

Missing Data Handling Conventions

When calculating the WPAI: Axial SpA scores, the following computational notes should be followed.

- Define Employment as a binary YES or NO variable where YES corresponds to "Employed" and NO corresponds to "Not Employed."
  - A subject will be considered "employed" at a given visit if Q1 = YES or Q2 > 0 or Q4 > 0.
  - A subject will be considered "unemployed" at a given visit if Q1 = NO and no positive hours recorded under Q2 and Q4 (i.e., if Q1 = NO AND Q2 ≤ 0 AND Q4 ≤ 0, then UNEMPLOYED).

  - Employment status for a subject will be considered "missing" at a given visit if Q1 = missing and no positive hours recorded under Q2 and Q4.
- If a subject is "unemployed" or employment status is "missing," then S1, S2, and S3 will be set to "missing."
- If Q2 = 0 and Q4 = 0 or missing then Q2/(Q2 + Q4) = missing (i.e., S1 = missing).
- If Q2 = 0 and Q4 = 0, then set S3 to missing.
- If Q2 is missing or Q4 is missing, then set S1 and S3 to missing.
- If Q4 = missing, then DO NOT set Q5 = missing.
- If Q5 is missing, then apply the following rules:
  - If Q2 > 0, Q4 = 0, and Q5 = missing, then S3 = 100%.
  - If Q2 = 0, Q4 > 0, and Q5 = missing, then S3 is missing.
  - If Q2 > 0, Q4 > 0, and Q5 = missing, then S3 is missing.
- Determine if a subject missed work (based on Q2) in order to analyze the proportion of subjects who missed work:
  - Create a binary (yes or no) "missed work" variable.
  - A subject will be considered as yes to missed work if Q2 is greater than 0.
  - If Q2 = missing, then MISSED WORK = missing.
  - If Q2 > 0, then MISSED WORK = "yes."
  - If Q2 = 0, then MISSED WORK = "no."
  - Therefore, the proportion of subjects who missed work will be counted based on the number of subjects with MISSED WORK = YES.

Appendix D. Visit Schedule per Protocol

Table 15. Study Activities

Activity	SCR (≤ 42 D)	BL[a] (D1)	Open-Label – Period 1 Week 2	4	8	12	16	20	24	28	Double-Blind – Period 2 Week 32	36	40	44	48	52	56	60	64	68	12 Weeks PF[b]	RT for Those that Flare at Week 60 – 68[c] 72	76	80	Disc. Visit[d]	70-Day F/U Call
Informed Consent	X																									
Inclusion/ Exclusion Criteria	X	X[e]																								
Medical/ Surgical History	X	X[e]																								
Vital Signs/ Weight/ Height[f]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	
Anterior Uveitis	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	
Physical Exam	X	X				X		X[i]	X[i]	X		X				X				X		X[i]	X[i]	X	X	
12-Lead ECG[g]	X[h]																									
Chest X-Ray[g]	X[h]															X[j]										
AP Pelvis																										
X-Ray	X[h]																									

Activity	SCR (≤ 42 D)	BL[a] (D1)	Open-Label – Period 1 Week 2	4	8	12	16	20	24	28	Double-Blind – Period 2 Week 32	36	40	44	48	52	56	60	64	68	12 Weeks PF[b]	RT for Those that Flare at Week 60 – 68[c] 72	76	80	Disc. Visit[d]	70-Day F/U Call
TB Screening (PPD Skin Test OR IGRA)	X[h]															X[f]										
Pregnancy Test	X[k]	X[l]						X[l,i]	X[l,i]	X[l,i]										X[l]		X[l,i]	X[l,i]	X[l]	X[l]	
Chemistry and Hematology	X	X[m]		X	X	X	X	X	X	X		X		X		X				X		X[v]	X[v]	X	X	
Urinalysis[n]	X	X[m]				X		X[i]	X[i]	X		X				X				X		X[i]	X[i]	X	X	
HLA-B27	X[h]																									
Antinuclear Antibody (ANA)/reflex Anti dsDNA antibody	X[h]																									
HBV Screening	X																									
hs-CRP	X	X[m]		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	
HIV	X[o]																									
Biomarkers[p]		X						X[i]	X[i]	X		X[p]	X[p]	X[p]	X[p]	X[p]	X[p]	X[p]	X[p]	X	X[p]	X[i]	X[i]	X	X	
Serum ADA Concentration[p]		X				X		X[i]	X[i]	X		X	X[p]	X[p]	X[p]	X	X[p]	X[p]	X[p]	X	X[p]	X[i]	X[i]	X	X	

Activity	SCR (≤ 42 D)	BL[a] (D1)	Open-Label – Period 1								Double-Blind – Period 2										12 Weeks PF[b]	RT for Those that Flare at Week 60 – 68[c]			Disc. Visit[d]	70-Day F/U Call
			Week								Week															
			2	4	8	12	16	20	24	28	32	36	40	44	48	52	56	60	64	68		72	76	80		
Anti-ADA antibody (AAA)[r]		X				X		X[i]	X[i]	X		X	X[p]	X[p]	X[p]	X	X[p]	X[p]	X[p]	X	X[p]	X[i]	X[i]	X	X	
PG Sample		X[q]																								
MRI of the Spine and SI Joints	X[h]							X[r]	X[r]	X[r]																
TJC/SJC		X				X		X[i]	X[i]	X		X				X				X		X[i]	X[i]	X	X	
BASDAI	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	
$BASMI_{lin}$		X				X		X[i]	X[i]	X		X				X				X		X[i]	X[i]	X	X	
MASES		X				X		X[i]	X[i]	X		X				X				X		X[i]	X[i]	X	X	
Plantar Fascia Enthesitis		X				X		X[i]	X[i]	X		X				X				X		X[i]	X[i]	X	X	
Dactylitis		X				X		X[i]	X[i]	X		X				X				X		X[i]	X[i]	X	X	
BASFI		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	
Physician's Global Assessment of Disease Activity NRS		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	

Activity	SCR (≤ 42 D)	BL[a] (D1)	Open-Label – Period 1								Double-Blind – Period 2										12 Weeks PF[b]	RT for Those that Flare at Week 60 – 68[c]			Disc. Visit[d]	70-Day F/U Call
			Week								Week															
			2	4	8	12	16	20	24	28	32	36	40	44	48	52	56	60	64	68		72	76	80		
Patient's Assessment of Nocturnal Back Pain NRS		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	
Patient's Assessment of Total Back Pain NRS	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	
Patient's Global Assessment of Disease Activity NRS	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	
Patient's Global Assessment of Pain NRS		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	
HAQ-S		X				X		X[i]	X[i]	X		X				X				X		X[i]	X[i]	X	X	
SF-36V2 Health Survey		X				X		X[i]	X[i]	X		X				X				X		X[i]	X[i]	X	X	
WPAI:Axial SpA		X				X		X[i]	X[i]	X		X				X				X		X[i]	X[i]	X	X	
EQ-5D		X				X		X[i]	X[i]	X		X				X				X		X[i]	X[i]	X	X	

Activity	SCR (≤ 42 D)	BL[a] (D1)	Open-Label – Period 1								Double-Blind – Period 2										12 Weeks PF[b]	RT for Those that Flare at Week 60 – 68[c]			Disc. Visit[d]	70-Day F/U Call
			Week								Week															
			2	4	8	12	16	20	24	28	32	36	40	44	48	52	56	60	64	68		72	76	80		
Prior and Concomitant Therapy Assessment	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	
Monitor AE	X[s]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X[t]
ASDAS Calculation	X	X						X[u]	X[u]	X[u]		X[v]	X[v]	X[v]	X[v]	X[v]	X[v]	X[v]	X[v]	X[v]						
Enrollment		X																								
Randomization										X[u]																
Dispense Study Drug		X		X	X	X	X	X	X	X[u]	X	X	X	X	X	X	X	X	X	X[w]		X[w]	X[w]			
Admin of Study Drug[x]		X	X	X	X	X			X																	
Optional Blood Sample Collection Sub-Study		X						X[y]	X[y]	X		X[z]	X[z]	X[z]	X[z]	X[z]	X[z]	X[z]	X[z]	X[z]						

Ada = Adalimumab; BL = Baseline; D = Day(s); Disc Visit = Discontinuation Visit; SCR = Screening; RT = Rescue Therapy; PF = Post Flare; F/U = Follow-up

a. Baseline Visit is defined as Day 1, the day of first study drug administration.

b. If subject meets flare criteria, these assessments are required 12 weeks post flare in addition to assessments for the scheduled visit when the flare was identified. If any of these procedures are already collected at the regularly scheduled visit, there is no need to obtain these procedures twice.

c. For subjects who meet the flare criteria at Weeks 60, 64 or 68, rescue therapy with OL adalimumab will be provided for 12 weeks and final study visit will be at Weeks 72, 76 or 80, respectively.

d. Subjects who prematurely discontinue the study should complete the procedures outlined for the Discontinuation Visit as soon as possible after the last dose of study drug and preferably prior to the administration of new therapies. No Discontinuation Visit procedures will be required for those subjects that complete the Week 68 visit (or Weeks 72, 76 or 80 for those subjects who flare at Weeks 60, 64 or 68, respectively).

e. Medical History and Inclusion/Exclusion Criteria must be confirmed by the Investigator prior to enrollment to verify subject eligibility.

f. Height will be measured at Screening Visit only.

g. Subject can have a repeat CXR and/or ECG at any time during the study only if in the opinion of the investigator, clinically significant AEs develop that warrant a repeat exam.

h. These procedures will not be required if the subject had a previous normal CXR, normal ECG or negative PPD within 90 days of screening provided all protocol required documentation is available at the site and nothing has changed in the subject's health status since the time of the test that warrants a repeat test. MRI of the SI joint and spine or AP Pelvis x-ray will also not be required if the subject had a previous test within 30 days and 1 year (365 days) of Screening, respectively. Any documentation of past positive PPD results may be acceptable; however, negative PPD test done more than 90 days prior to the Screening visit will need to be repeated. If the subject has a positive PPD test or IGRA test, has had a past ulcerative reaction to PPD placement and/or a CXR consistent with prior TB exposure, the subject will be required to initiate and complete at least the first 2 weeks (or per local guidelines, whichever is longer) of the prescribed TB prophylaxis, or have documented completion of a full course of TB prophylaxis prior to Baseline. There is no need to redraw the HLA-B27 or ANA once the results are available from the central laboratory chosen for this study.

i. These procedures should only be performed if the visit is the subject's last visit due to not meeting ASDAS remission criteria (Weeks 20, 24 or 28) or due to a flare and has received 12 weeks of rescue therapy (Weeks 72 or 76).

j. An annual PPD/IGRA test will be required for those subjects with a negative TB test at Screening. If the annual PPD/IGRA test is positive, a CXR must be performed, the investigator must contact the study designated physician to review the individual data and agree on a decision whether or not additional evaluation is needed.

k. All females of childbearing potential will have a serum pregnancy test at Screening that is performed at the central laboratory.

l. All females of childbearing potential will have a urine pregnancy test performed at Baseline prior to study enrollment and at study discontinuation/completion. The urine pregnancy kit supplies will be provided to the sites via the central laboratory (where allowed by local guidelines). Monthly pregnancy tests will be performed throughout the study if required by country regulatory authorities. Any subject with a positive urine pregnancy test must have a negative serum test performed at the central laboratory prior to enrollment or continuation in the study.

m. Laboratory assessments (chemistry, hematology, hs-CRP and urinalysis) will 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.

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

o. If required by country regulatory authorities to confirm eligibility, subjects will be tested for antibodies to HIV, and document that the test has been performed. The 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.

p. Flare is defined as 2 consecutive study visits with ASDAS ≥ 2.100. Subjects that flare during the double-blind period (Period 2) will need the following assessments during the visit that flare was determined as well as 12 weeks following the determination of the flare: biomarkers (refer to Protocol Amendment 2, Section 5.3.1.1 for further information), serum adalimumab and AAA (refer to Protocol Amendment 2, Section 5.3.2 for further information). These assessments are in addition to the regularly scheduled assessments outlined in the table above.

q. The pharmacogenic (PG) sample is optional. A separate PG consent form must be signed prior to sample collection. The PG sample is ideally collected at Baseline, but may be collected at any time throughout the study.

r. Site staff should schedule the Week 28 MRI during the Baseline visit, if possible. For subjects that are discontinued due to not fulfilling the ASDAS remission criteria at Weeks 20 or 24, the site should attempt to reschedule the MRI within 2 weeks (14 days) of the discontinuation visit. If the MRI cannot be rescheduled, the Week 28 MRI will be considered missed.

s. SAEs will be collected starting from time of signing informed consent; non-serious AEs will be collected starting from the 1st study drug dose. Any AE that occurs between Screening and Baseline should be captured as medical history.

t. Site personnel will contact all subjects by telephone approximately 70 days after the last dose of study drug to determine the occurrence of AEs or SAEs. Subjects that initiate adalimumab therapy not supplied in the context of the clinical trial after the end of study participation need not be contacted by phone as new AEs or SAEs should be reported through the mechanism used for all post marketing adverse experiences.

u. Week 20, calculation of the Week 16 ASDAS will be done using the Week 16 visit parameters including the Week 16 hs-CRP. If the subject does not meet the criteria for remission (ASDAS < 1.300), the subject will be discontinued. At Week 24, calculation of the Week 20 ASDAS will be done using the Week 20 visit parameters including the Week 20 hs-CRP. If the subject does not meet the criteria for remission (ASDAS < 1.300), the subject will be discontinued. At Week 28, calculation of the Week 24 ASDAS will be done using the Week 24 parameters including the Week 24 hs-CRP; the Week 28 ASDAS will be done using Week 28 parameters and the Week 24 hs-CRP (as the Week 28 results will not be available). If any of these time points do not meet the criteria for remission (ASDAS < 1.300), the subject will be discontinued. For hs-CRP elevations secondary to AEs or using a previous hs-CRP, please refer to Section 5.1 or ASDAS Calculation in Protocol Amendment 2, Section 5.3.1.1.

v. Beginning at Week 36, ASDAS will be calculated via IVRS/IWRS for the current and previous study visit to determine fulfillment of flare criteria. Assessments from the current and previous study visit will be used for the calculation except for the hs-CRP which will be based on the results from the previous study visit's blood draw for both calculations. For subjects who meet flare criteria during Period 2 and start rescue therapy with OL adalimumab, further calculation of ASDAS is not required. For any subject missing 2 consecutive visits within the double-blind period (Period 2) the site must contact the AbbVie Medical Monitor for discussion regarding the subject's continued participation in the study.

w. Subjects would only receive study drug at Weeks 68, 72 or 76, if they flared at Week 56, 60 or 64 respectively.

x. Study drug will be administered in the Investigator's office at Baseline, Week 2, 4, 8, 12, and Week 24 to ensure proper technique.

y. Optional Blood Sample Collection Sub-Study only: use a separate sub-study central lab kit. Samples will be collected at Week 28 or at Weeks 20 or 24 only if the visit is the subject's last visit. (Optional Blood Sample Collection Sub-Study samples are in addition to the regularly scheduled assessments outlined in the table above.)

z. Optional Blood Sample Collection Sub-Study only: use a separate sub-study central lab kit. In subjects who meet flare criteria during Weeks 36 – 68, samples will be collected only once at time of flare. Flare is defined as 2 consecutive study visits with ASDAS ≥ 2.100. (Optional Blood Sample Collection Sub-Study samples are in addition to the regularly scheduled assessments outlined in the Table above.)

Document Approval

Study M13375 - Statistical Analysis Plan Version 3 - 02Mar2017 (E3 16.1.9)

Version: 1.0 **Date:** 02-Mar-2017 11:08:37 PM **Company ID:** 03022017-00F9F68148ADC4-00001-en

Signed by:	Date:	Meaning Of Signature:
[redacted]	02-Mar-2017 04:51:35 PM	Approver
	02-Mar-2017 04:57:23 PM	Approver
	02-Mar-2017 11:08:36 PM	Approver