1.0 Title Page

Statistical Analysis Plan

Study M13-538

Phase 2 Study, Multicenter, Open-Label Extension (OLE) Study in Rheumatoid Arthritis Subjects Who Have Completed a Preceding Phase 2 Randomized Controlled Trial (RCT) with ABT-494

Date: 03 May 2021

Version 3.0

2.0 Table of Contents

List of Tables

List of Figures

List of Appendices

3.0 Introduction

This statistical analysis plan (SAP) describes the statistical analyses to be completed by the Data and Statistical Science Department for upadacitinib Study M13-538. It provides details to further elaborate statistical methods as outlined in the protocol and describes analysis conventions to guide the statistical programming work.

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

4.0 Study Objectives, Design and Procedures

4.1 Study Objectives

The primary objective is to evaluate the long-term safety, tolerability, and efficacy of upadacitinib in RA subjects who have completed Study M13-550 or Study M13-537 Phase 2 RCT with upadacitinib.

4.2 Overall Study Design and Plan

This is a Phase 2, multicenter, open-label extension (OLE) study to assess the long-term safety, tolerability, and efficacy of upadacitinib in RA subjects who have completed Study M13-550 or Study M13-537 Phase 2 RCT with upadacitinib. This study has enrolled 494 subjects. Subjects will be treated for approximately 312 weeks.

Only those subjects who have met all of the specified inclusion criteria and none of the exclusion criteria will have an option to enter into the OLE study to receive upadacitinib, as long as the subject is willing and the Investigator believes that continuing therapy with upadacitinib is appropriate.

All eligible subjects will be assigned to upadacitinib 6 mg BID immediately after or up to 30 days following the Last Visit of Study M13-550 or Study M13-537 Phase 2 RCT. Subjects who are unable to tolerate 6 mg BID will be discontinued from the study. At Week 6, if a subject fails to achieve at least 20% improvement from RCT Baseline in

Tender Joint Count (TJC) and Swollen Joint Count (SJC), upadacitinib dose should be increased to 12 mg BID as long as the Investigator has no safety concerns. After 6 weeks of treatment with upadacitinib 12 mg BID, the improvement in TJC and SJC will be re-assessed at next scheduled visit (Week 12). If the subject on 12 mg BID fails to achieve at least 20% improvement in TJC and SJC from RCT Baseline the subject will be discontinued.

At Week 12, the same process will be followed. If a subject still on 6 mg BID fails to achieve at least 20% improvement from RCT Baseline in TJC and SJC ABT-494 dose should be increased to 12 mg BID. The improvement in TJC and SJC will be re-assessed after 6 weeks of treatment with 12 mg BID at Week 18 (an optional study visit). Subjects who fail to achieve at least 20% improvement from RCT Baseline in TJC and SJC with 12 mg BID at Week 18 will be discontinued.

After Week 12, if a subject fails to show at least 20% improvement from RCT Baseline in TJC and SJC at 2 consecutive scheduled study visits then the subject will be discontinued.

Starting at Week 6 and during any scheduled visits thereafter, upadacitinib dose may be increased from 6 mg BID to 12 mg BID (or 30 mg QD from January 2017) if a subject fails to achieve the Low Disease Activity (LDA) status (CDAI > 10) and has no safety concerns per Investigator's judgment.

At any visit, upadacitinib dose may be decreased back to 6 mg BID per Investigator's judgment due to either an adverse event or reaching one of the protocol specific toxicity management thresholds. Dose increase back to 12 mg BID (or 30 mg QD from January 2017) is not allowed.

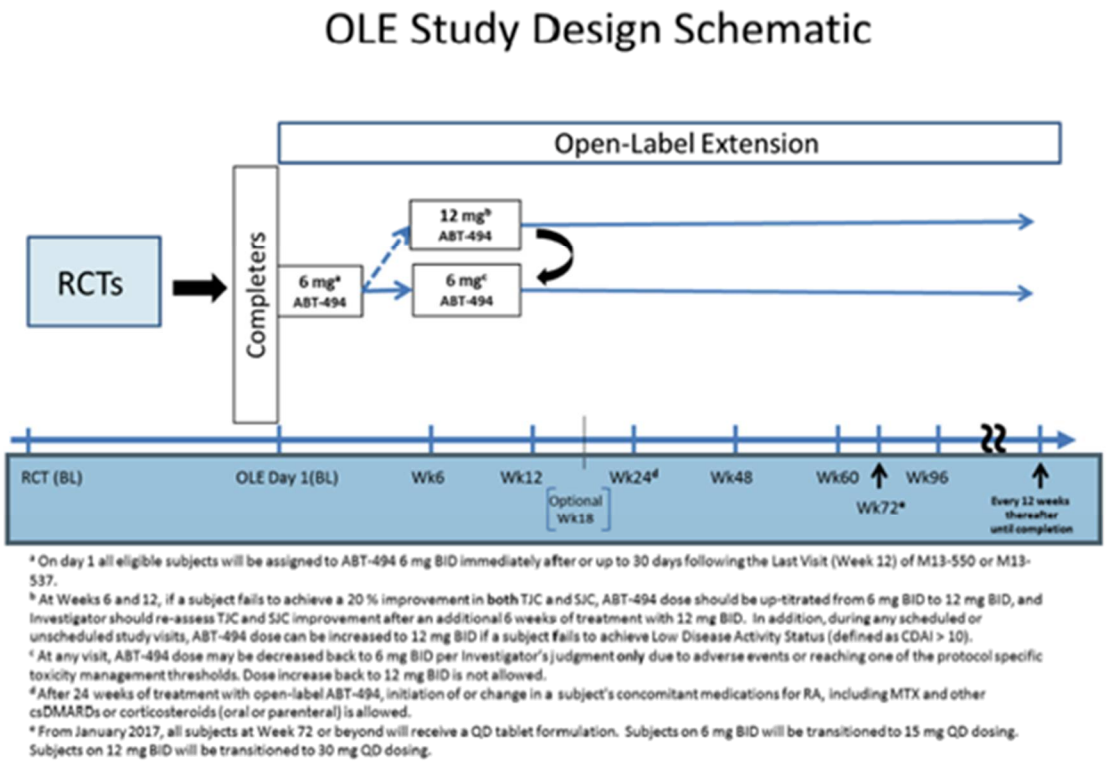
From January 2017, all subjects who are at Week 72 or beyond will receive a once–daily tablet formulation. Subjects who are on 6 mg BID capsule dosing will be transitioned to 15 mg QD tablet dosing. Subjects who are on 12 mg BID capsule dosing will be transitioned to 30 mg QD tablet dosing. The capsule formulation will not be available to subjects once the transition to tablet formulation has occurred. The change to once daily

tablet formulation in this study is based on extrapolation of pre-clinical efficacy models and analyses of PK, pharmacodynamic, safety, and efficacy data from the Phase 1 studies in healthy volunteers (single and multiple ascending dose Studies M13-401 and M13-845, respectively) and completed Phase 2b randomized, controlled studies (RCTs) in RA subjects (Studies M13-537 and M13-550). The doses selected, 15 mg QD and 30 mg QD, are expected to approximate the exposures achieved with the 6 mg BID and 12 BID immediate release formulations, respectively, which should be efficacious with an acceptable safety profile. Preliminary results from an ongoing relative bioavailability study demonstrated that single 15 mg and 30 mg doses of the once-daily formulation provide comparable AUC and significantly lower C_{max} than single doses of 12 mg immediate release (IR) formulation (AUC equivalent to 6 mg BID dose tested in Phase 2) and 24 mg IR formulation (AUC equivalent to 12 mg BID dose tested in Phase 2), respectively, under fasting conditions. At steady-state, the once-daily dose of 15 mg QD is predicted to achieve comparable daily AUC and C_{max}, and non-inferior C_{trough} to 6 mg BID IR formulation. In the Phase 2 RCTs, the 6 mg BID dose was shown to achieve the near maximum efficacy. The once-daily dose of 30 mg QD is predicted to achieve comparable daily AUC and C_{max}, and non-inferior C_{trough} to 12 mg BID IR formulation. In the Phase 2 RCTs, the 12 mg BID dose was clearly shown to achieve the plateau of efficacy.

Starting with Amendment 5, subjects who are receiving 30 mg QD open-label upadacitinib will have the option to decrease the dose to 15 mg QD based on investigator's discretion.

A schematic of the overall study design is shown in Figure 1.

Figure 1. Study Design Schematic



[a] On day 1 all eligible subjects will be assigned to ABT-494 6 mg BID immediately after or up to 30 days following the Last Visit (Week 12) of M13-550 or M13-537.
[b] At Weeks 6 and 12, if a subject fails to achieve a 20 % improvement in **both** TJC and SJC, ABT-494 dose should be up-titrated from 6 mg BID to 12 mg BID, and Investigator should re-assess TJC and SJC improvement after an additional 6 weeks of treatment with 12 mg BID. In addition, during any scheduled or unscheduled study visits, ABT-494 dose can be increased to 12 mg BID if a subject fails to achieve Low Disease Activity Status (defined as CDAI > 10).
[c] At any visit, ABT-494 dose may be decreased back to 6 mg BID per Investigator's judgment **only** due to adverse events or reaching one of the protocol specific toxicity management thresholds. Dose increase back to 12 mg BID is not allowed.
[d] After 24 weeks of treatment with open-label ABT-494, initiation of or change in a subject's concomitant medications for RA, including MTX and other csDMARDs or corticosteroids (oral or parenteral) is allowed.
[e] From January 2017, all subjects at Week 72 or beyond will receive a QD tablet formulation. Subjects on 6 mg BID will be transitioned to 15 mg QD dosing. Subjects on 12 mg BID will be transitioned to 30 mg QD dosing.

Treatment Period

The treatment period for OLE can begin immediately after or up to 30 days following the Last Visit of Study M13-550 or Study M13-537 RCT and will continue for approximately 312 weeks in OLE. At Last Visit (Week 12) or up to 30 days following the Last Visit (Week 12) of Study M13-550 or Study M13-537, subjects who meet all the inclusion criteria and none of the exclusion criteria described in protocol Section 5.2.1 and Section 5.2.2 will be enrolled into the study. Subjects will visit the study site at Weeks 6, 12, 18 (Optional), 24, 36, 48, 60, 72, 84, 96 and every 12 weeks thereafter until the end of the study or if they terminate early from the study. A ± 7-day window is permitted around

scheduled study visits. The last dose of study drug is taken the evening prior to the final visit.

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

Starting with Amendment 5, subjects who are receiving 30 mg QD open-label upadacitinib will have the option to decrease to the 15 mg QD dose based on investigator's discretion.

Follow-Up Period

All subjects will have a follow-up visit approximately 30 days after the last administration of study drug to obtain information on any new or ongoing adverse events (AEs), and to collect vital signs and clinical laboratory tests. The 30 day follow up visit may be a telephone call if a site visit is not possible. Vital signs and laboratory tests may not be required.

Protocol Modifications due to State of Emergency or Pandemic like COVID-19

Study visits may be impacted by a pandemic situation (including the COVID-19 pandemic) or any state of emergency. If visits cannot be conducted onsite due to travel restrictions or other pandemic-related reasons, virtual visits, visits at alternative locations, or changes in the visit frequency and timing of study procedures, among others may be performed. Additional details are provided in the subsequent sections. Every effort should be made to ensure the safety of subjects and site staff, while maintaining the integrity of the study.

4.3 Interim Analysis and Data Base Lock

Interim Analyses may be performed, as appropriate, to generate long term safety and efficacy data and to help plan future studies.

Prior to interim analyses, interim data cuts will be taken, where database versions will be preserved and discrepant data will be clarified. The analyses will be conducted by the clinical statistics department at AbbVie.

4.4 Data Monitoring Committee (DMC) Activities

An independent external Data Monitoring Committee (DMC) is used to review unblinded safety data at regular intervals during the conduct of the study. The DMC will provide recommendation to an AbbVie Point of Contact on whether to continue, modify, or terminate studies after each review. When needed, high-level unblinded efficacy data may also be requested by the DMC and be reviewed so that the DMC can assess benefit:risk of any emerging safety differences.

5.0 Analysis Populations and Analysis Windows

5.1 Analysis Populations

Open-Label Treated Population

The open-label treated population consists of all subjects who have received at least one dose of open-label study medication. It will be used for all long-term efficacy, safety, and baseline analysis.

5.2 Definition of Baseline

The baseline value will be determined by the last non-missing measurement recorded on or before the date of the first dose of study drug in RCT. This applies to both efficacy and safety analysis.

5.3 Analysis Windows

Definition of OL-Rx Days (Days Relative to the First Dose of Open-Label Study Drug)

For the purpose of defining analysis windows for the open-label study, OL-Rx Days are calculated for each time point relative to the first dose of open-label study drug. They are defined as the number of days between the day of the first dose of open-label study drug and the specific time point. OL-Rx days are negative values when the time point of interest is prior to the first dose day of the open-label study drug, while OL-Rx days are positive values when the time point of interest is on or after the first dose day of open-label study drug. The day of the first dose of open-label study drug is defined as OL-Rx Day 1, while the day immediately prior to the first dose day of open-label study drug is defined as OL-Rx Day –1 (there is no OL-Rx Day 0). OL-Rx days are used to map actual study visits to the protocol specified study visits.

Definition of Analysis Windows

The following rules will be applied to assign actual subject visits to protocol-specified visits. For each protocol-specified study visit, a target study day will be identified 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 collection date does not fall into multiple visit windows. If a subject has two or more actual visits in one visit window, the visit closest to the target day will be used for analysis. If two visits are equidistant from the target day, then the later visit will be used for analysis.

The visit window and the target OL-Rx day for each protocol-specified visit are displayed in Table 1.

Table 1. Analysis Windows for Efficacy, Laboratory Assessments and Vital Signs

Protocol Specified Visit Week	Lower Bound	Target Day[a]	Upper Bound
Baseline	Last non-missing measurement recorded on or before the date of the first dose of study drug in RCT		
6	2	43	64
12	65	85	127
24	128	169	211
36	212	253	295
48	296	337	379
60	380	421	463
72	464	505	547
84	548	589	631
96	632	673	715
108	716	757	799
120	800	841	883
132	884	925	967
144	968	1009	1051
156	1052	1093	1135
168	1136	1177	1219
180	1220	1261	1303
192	1304	1345	1387
204	1388	1429	1471
216	1472	1513	1555
228	1556	1597	1639
240	1640	1681	1723
252	1724	1765	1807
264	1808	1849	1891
276	1892	1933	1975
288	1976	2017	2059
300	2060	2101	2143
312	2144	2185	2227

a. Relative to the day of first dose of open-label study drug.

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

6.1 Demographic and Baseline Characteristics

The demographic data and baseline characteristics will be reported for the Baseline visit of the preceding RCT Studies M13-537 and M13-550. Demographic and baseline characteristics will be summarized for the Open-Label Treated Population. 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 frequencies and percentages. The summary statistics will be computed for each treatment group, and the combined upadacitinib group.

Main Demographic and Baseline Characteristics

- Sex (male/female)
- Age (years)
- Age Categories (< 45, [45,65), ≥ 65 years)
- Race (White, Black, American Indian/Alaska Native, Native Hawaiian or Other Pacific Islander, Asian, Other)
- Geographic region (Western Europe, Eastern Europe, North America, South/Central America, Other)
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino)
- Weight (kg)
- Weight Categories (< 60 kg, ≥ 60 kg)
- Height (cm)
- Body Mass Index (BMI) (kg/m^2)
- Body Mass Index (BMI) Category (kg/m^2) (< 18.5, [18.5,25), [25,30), ≥ 30)

RA Medical History and Characteristics

- Duration of RA in years
- Number of prior DMARDs used (for Study M13-537 subjects)

- MTX as first prior DMARD (Yes/No) (for Study M13-537 subjects)
- Number of prior biologic DMARD(s) used (for Study M13-550 subjects)
- Oral steroid use at baseline (yes, no)
- Oral steroid dose (prednisone equivalent) at baseline

Baseline RA Disease Characteristics

- Individual scores of each ACR component (e.g., hsCRP, SJC, TJC, etc.)
- DAS28 (CRP)
- DAS28 categories:
  - DAS28 > 5.1 (High Disease Activity)
  - DAS28 ≤ 5.1
- Clinical Disease Activity Index (CDAI) and categories:
  - CDAI > 22 (High Disease Activity)
  - CDAI ≤ 22
- Simplified Disease Activity Index (SDAI) and categories:
  - SDAI > 26 (High Disease Activity)
  - SDAI ≤ 26
- Anti-cyclic citrullinated peptide (Anti-CCP) (units)
- Anti-CCP status (Positive or Negative)
- Rheumatoid Factor (RF) (KU/L)
- RF status (Positive or Negative)

Patient Report Outcomes at Baseline

- Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-F)
- Work Instability Scale for Rheumatoid Arthritis (RA-WIS)
- EQ-5D-5L

Clinical Tests at Screening

- Chest x-ray
- ECG
- Tuberculin PPD skin test, QuantiFERON TB Gold test
- Serum pregnancy test

Immunization History

- BCG immunization
- Herpes Zoster immunization

Tobacco/Nicotine and Alcohol Use

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

6.2 Medical History

Medical history data (defined relative to the preceding RCT studies) will be summarized and presented for Open-Label Treated Population using body systems and conditions/diagnoses as captured on the CRF. 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 by treatment group sequence (defined in Section 9.1) as well total as for Open-Label Treated Population. Subjects reporting more than one condition/diagnosis within a body system will be counted only once for that body system. No statistical comparison will be performed for medical history reporting.

6.3 Previous Treatment and Concomitant Medications

Prior and concomitant medications will be summarized by treatment group sequence (defined in Section 9.1) as well as total for Open-Label Treated Population. A prior

medication is defined as any medication taken prior to the first dose of study drug in the preceding RCT study. This includes medications with a start date before the first study drug administration date, regardless of the end date of these medications. Medications taken on the day of the first dose of study drug are not counted as prior medications. A concomitant medication is defined as any medication, other than study drug, that started prior to the first dose of open-label study drug and continued to be taken after the first dose of open-label study drug or any medication that started after the first dose of open-label study drug and within 1 day of the last dose of study drug.

The number and percentage of subjects who received a prior or concomitant medication will be tabulated separately by the generic name assigned by the most current version of the World Health Organization (WHO) Drug Dictionary.

6.4 Protocol Deviations

Protocol deviations based on ICH deviation criteria are categorized as follows:

1. Those who entered the study even though they did not satisfy the entry criteria
2. Those who developed withdrawal criteria during the study and were not withdrawn
3. Those who received the wrong treatment or incorrect dose, and
4. Those who received an excluded or prohibited concomitant medication.

The protocol deviations listed above will be summarized and listed by treatment group sequence (defined in Section 9.1) as well as overall for the Open-Label Treated Population.

7.0 Patient Disposition

The number of subjects treated, completed or discontinued in the open-label treatment period will be summarized by treatment group sequence (defined in Section 9.1) as well as overall for the Open-Label Treated Population. The number of subjects discontinuing

from the open-label treatment period will also be summarized with the reasons for discontinuation collected from CRF by the following categories:

- Adverse event
- Withdrew consent
- Lost to follow-up
- Lack of efficacy
- Subject non-compliance
- Requires alternative (or prohibited) therapy
- COVID-19 infection
- COVID-19 logistical restrictions
- Other

8.0 Study Drug Exposure and Compliance

8.1 Study Drug Exposure

The duration of exposure to study drug in M13-538 will be summarized for the Open-Label Treated Population by the following groups:

1. Upadacitinib 6 mg BID/15 mg QD

   This includes upadacitinib 6 mg BID or 15 mg QD exposure from subjects who never titrated, as well as the upadacitinib 6 mg BID or 15 mg QD exposure prior to dose titration for subjects who titrated (censored at the time of up titration to 12 mg BID/30 mg QD).

2. Upadacitinib 12 mg BID/30 mg QD

   This includes upadacitinib 12 mg BID or 30 mg QD exposure for subjects who titrated up (censored at the time of down titration to 6 mg BID/15 mg QD for subjects who titrated up and back down).

3. Upadacitinib 6 mg BID/15 mg QD post down titration

   This includes upadacitinib 6 mg BID/15 mg QD exposure after down titration for subjects who titrated up and back down.

4. Any upadacitinib

   This includes any upadacitinib exposure in Study M13-538.

The first three exposure groups are mutually exclusive in terms of exposure, and the three groups combined is represented by the fourth group.

The duration of exposure to study drug will be summarized for each exposure group as specified above, with the number of subjects, mean, standard deviation, median, minimum and maximum values. In addition, the number and percentage of subjects exposed to study drug will be summarized for the following cumulative duration intervals.

Open-label exposure:

- ≥ 6 months
- ≥ 12 months
- ≥ 18 months
- ≥ 2 years
- ≥ 2.5 years
- ≥ 3 years
- ≥ 4 years
- ≥ 5 years

8.2 Compliance

Study drug compliance (TC) will summarized for the Open-Label Treated Population. The treatment compliance is defined as the number of capsules/tablets taken (i.e., the difference between the number of capsules/tablets dispensed and the number of

capsules/tablets returned) during the subject's participation in the open-label treatment period divided by the number of capsules/tablets planned to be taken by the subject during the open-label treatment period of the study. In Study M13-538, subjects will switch from twice-daily formulation (capsules) to once-daily formulation (tablets). The impact of missing a once-daily tablet should be weighted twice as much as that for a twice-daily capsule. Specifically, TC will be calculated using the following formula:

$$TC = \frac{\text{number of twice daily capsules taken} + \text{number of once daily tablets taken} \times 2}{(\text{Number of days from first to last dose of open label study drug}) \times 2} \cdot 100\%$$

9.0 Efficacy Analysis

9.1 General Considerations

Long-term efficacy analysis will be performed on the Open-Label Treated Population by the treatment group sequences as defined below:

1. Upadacitinib never titrated (6 mg BID/15 mg QD),
2. Upadacitinib titrated up and not down (6 mg BID/15 mg QD to 12 mg BID/30 mg QD), and
3. Upadacitinib titrated up and down (6 mg BID/15 mg QD to 12 mg BID/30 QD and then back down to 6 mg BID/15 mg QD).

No statistical tests will be conducted; only descriptive statistics and confidence intervals will be provided. All data will be summarized based on As Observed data (defined in Section 9.1.1).

9.1.1 Definition of Missing Data Imputation

As Observed (AO)

The AO analysis will not impute values for missing evaluations. All observed data will be used in the analysis. The AO analysis will be applied to long-term efficacy analysis.

Mixed Model Repeated Measures (MMRM) and Generalized Linear Mixed Model (GLMM)

The repeated measure analysis will be conducted using mixed model including As Observed measurements at all visits. MMRM will be used for continuous endpoints and GLMM will be used for binary endpoints. The mixed models will include the categorical fixed effects of treatment, visit and treatment-by-visit interaction. Prior RCT study will also be included as a fixed factor in the model. For the MMRM analysis of change from baseline in continuous endpoints, the baseline measurement will be included as a continuous fixed covariate. The categorical fixed effect of subject's discontinuation status may also be included in the model as appropriate. Unstructured, Toeplitz, compound symmetry, or other covariance structures may be considered.

9.2 Long-Term Efficacy Analysis

Assessments to evaluate long-term efficacy will be analyzed for the following measures at Weeks 6, 12, 24, 36, 48, 60, 72, 84, 96 and every 12 weeks until Week 312.

- ACR20/50/70 response rates;
- Change from baseline in individual ACR components;
- Change from baseline in DAS28 (CRP);
- Change from baseline in CDAI and SDAI;
- Proportion of subjects achieving LDA and proportion of subjects achieving CR based on DAS28 (CRP), CDAI and SDAI criteria;
- Change from baseline in Patient Reported Outcomes, including FACIT-Fatigue Scale, RA-WIS, and EQ-5D;

Analyses will be based on As Observed (AO) data. Descriptive statistics will be provided for each treatment group sequence in the Open-Label Treated Population as defined in Section 9.1. These include the number of observations, mean, standard deviation, and 95% CI for continuous endpoints; and frequencies and percentages with 95% CI using normal approximation for binary endpoints. In addition, longitudinal analysis will be

performed using MMRM or GLMM as described in Section 9.1.1 for all endpoints for treatment sequences 1 and 2. Point estimates and 95% CI from the model will be provided. Plots over time will be provided.

Table 1 below provides the overview of the long-term efficacy analysis to be performed on different endpoints.

Table 2. Summary of Long-Term Efficacy Analysis

Measure of Interest	Analysis Method
Binary endpoints: • ACR20 at each visit • ACR50 at each visit • ACR70 at each visit • LDA based on DAS28 (CRP) at each visit • LDA based on CDAI and SDAI at each visit • CR based on DAS28 (CRP) at each visit • CR based on CDAI and SDAI at each visit	• Point estimate and 95% CI of the response rate for each treatment sequence • Point estimate and 95% CI of the response rate using GLMM model for treatment sequences 1 and 2 • Plots over time (for ACR responses and LDA/CR based on DAS28 and CDAI) • Imputation: AO • Population: open-label treated population
Continuous endpoints: • Change from baseline for DAS28 (CRP) at each visit • Change from baseline for CDAI and SDAI at each visit • Change from baseline for each ACR individual component (i.e., TJC, SJC, Physician's Global Assessment of Disease Activity, Patient's Global Assessment of Disease Activity, Patient's Assessment of Pain, HAQ-DI, hsCRP) at each visit • Change from baseline for other PROs (FACIT, RA-WIS, EQ-5D) at each visit	• Point estimate and 95% CI of mean change from baseline for each treatment sequence together with SD • Point estimate and 95% CI of change from baseline using MMRM model for treatment sequences 1 and 2 • Plots over time (for DAS28 (CRP), CDAI and SDAI) • Imputation: AO • Population: open-label treated population

9.3 Handling of Multiplicity

There is no statistical testing, and therefore multiplicity adjustment is not applicable.

9.4 Efficacy Variables Definitions and Conventions

9.4.1 ACR Criteria

ACR criteria are a commonly used standard criteria set mentioned in the guidance of American College of Rheumatology and US FDA to evaluate the effectiveness of investigational drug in RA clinical trials. It is a composite measurement calculated based on the improvement over a set of core measurements.

ACR20 is defined as at least 20% improvement (compared to baseline values) in tender and swollen joint counts and at least 20% improvement in 3 of the remaining 5 core set measures (subject global assessment of pain, subject global assessment of disease activity, physician global assessment of disease activity, subject assessment of physical function and acute phase reactant hsCRP).

ACR50 and ACR70 are similarly defined with at least 50% and 70% improvement, respectively.

A subject will be classified as an ACR20 (ACR50, ACR70) responder, if the following conditions are met:

1. ≥ 20 % (50%, 70%) improvement from baseline in tender joint count (TJC68) and
2. ≥ 20 % (50%, 70%) improvement from baseline in swollen joint count (SJC66) and
3. ≥ 20% (50%, 70%) improvement from baseline in at least 3 of the following 5:
   - Patient's assessment of pain
   - Patient's global assessment of disease activity (PGA)
   - physician's global assessment of disease activity (PhGA)
   - patient's self-assessment of physical function (i.e., measured by Health Assessment Questionnaire (HAQ-DI score)
   - Acute-phase reactant value hsCRP/ESR

The following table illustrates examples for As Observed ACR calculations.

Example	TJC 68	SJC 66	Component 1	Component 2	Component 3	Component 4	Component 5	ACR20-Response?
A	1	1	1	1	1	.	.	Yes
B	1	0	1	1	1	1	1	No
C	.	0	.	.	.	.	.	No
D	1	.	1	1	1	1	1	.
E	1	1	0	0	0	1	1	No
F	.	.	0	0	0	.	.	No
G	1	1	1	1	0	0	.	.

Legend: 1 = ≥ 20% improved compared to baseline; 0 = < 20% improved compared to baseline; "." Missing

Windowing Rule for ACR Response Calculation:

- ACR component values will first be determined at each date within a visit window.
- ACR component values at each date will be combined to determine the observed ACR composite score at each date in each window.
- After this calculation, if multiple non-missing ACR composite scores are available within a given visit window, the non-missing ACR composite score closest to the target day will be used. If two composite scores have the same distance from the target day, the later one will be used. The corresponding date will be used as the observed ACR response date in the derived efficacy dataset.
- If a non-missing ACR composite score is not available for any day within a given visit window, the windowed component values for that visit will be used to calculate the ACR composite score for that visit window (component value windowing follow the same rules as in steps described above). The date of observed ACR composite score will be determined by the first available ACR component date, in the order of TJC, SJC, Pain, PGA, PhGA, HAQ-DI, CRP/ESR, in the derived efficacy dataset.

9.4.2 Joint Evaluation

Anatomical joints are evaluated for swelling and tenderness at every study visit. The 34 anatomical joints in Table 3 are assessed in this study for both the left and right side of the body.

Table 3. Anatomical Joints Assessed for Calculation of Tender and Swollen Joint Counts (TJC68 and SJC66)

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 assessed whether a particular joint was "tender or painful" where presence of tenderness was scored as "1" and the absence of tenderness was 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), which is based on 68 joints, 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), which is based on 66 joints as the hip joints are excluded. Thus, the range for TJC68 will be 0 to 68 and 0 to 66 for SJC66.

9.4.3 Patient's Global Assessment of Disease Activity Visual Analog Scale (VAS)

The subject will assess his/her disease activity for the past 24 hours using a Patient's Global Assessment of Disease VAS. The range is 0 to 100 mm with no activity being indicated by 0 and severe activity by 100.

9.4.4 Physician's Global Assessment of Disease Activity Visual Analog Scale (VAS)

The physician will assess Patient's disease activity at the time of visit using a Physician's Global Assessment of Disease VAS. The range is 0 to 100 mm with no activity being indicated by 0 and severe activity by 100.

9.4.5 Patient's Global Assessment of Pain

The subject will assess his/her pain in the previous week using a Patient's Global Assessment Pain VAS. The range is 0 to 100 mm with no pain being indicated by 0 and severe pain by 100.

9.4.6 Disease Activity Score Based on CRP (DAS28 [CRP])

DAS28 (CRP) is a composite index to assess disease activity in RA patients using hsCRP measurement. The DAS provides a score between 0 and 10, indicating how active the rheumatoid arthritis is at the time of measurement.

It can be calculated based on Tender Joint Count, Swollen Joint Count, Patient's Global Assessment of Disease Activity (PtGA) (in mm), and hsCRP (in mg/L).

$$\text{DAS28 (CRP)} = 0.56 \times \sqrt{(\text{TJC28*})} + 0.28 \times \sqrt{(\text{SJC28**})} + 0.36 \times \ln(\text{hsCRP}^{\&} + 1) + 0.014 \times \text{PtGA}^{\text{»}} + 0.96$$

* TJC28 refers to the Subject's total Tender Joint Count out of the provided 28 evaluated joints.
** SJC28 refers to the Subject's total Swollen Joint Count out of the provided 28 evaluated joints.
& hsCRP refers to the high-sensitivity c-reactive protein lab value. hsCRP unit in the DAS28 (CRP) equation is expressed as mg/L.
# ESR refers to the Erythrocyte sedimentation rate. ESR unit in the DAS28 (ESR) equation is expressed as mm/hr.
» PtGA refers to the Patient's Global Assessment of Disease Activity.
where $\sqrt{}$ is square root and ln is natural log.

TJC28 and SJC28 represent total tender joint count and total swollen joint count, respectively, based on the 28 joints (including the left and right side of the body) listed in the table below.

Table 4. Anatomical Joints for DAS28 (CRP) Calculation

Shoulder	Elbow	Wrist	Thumb Interphalangeal
Metacarpophalangeal I	Metacarpophalangeal II	Metacarpophalangeal III	Metacarpophalangeal IV
Metacarpophalangeal V	Proximal Interphalangeal II hand	Proximal Interphalangeal III hand	Proximal Interphalangeal IV hand
Proximal Interphalangeal V hand	Knee		

To calculate observed DAS28 scores, the observed component value will be calculated first. Then the components will be included in the calculation per the DAS28 formula. If any observed component is missing in a window, then the observed DAS28 score will be missing.

9.4.7 Clinical Disease Activity Index (CDAI)

CDAI is a composite continuous index to assess disease activity without using hsCRP measurement.

It can be calculated based on TJC28, SJC28, Patient's Global Assessment of Disease Activity (PtGA) (in cm, 0 – 10) and Physician's Global Assessment of Disease Activity (PhGA) (in cm, 0 – 10).

CDAI = TJC28 + SJC28 + PtGA (cm) + PhGA (cm).

To calculate observed CDAI scores, the observed component value will be calculated first. Then the components will be included in the calculation per the CDAI formula selected. If any observed component is missing in a window, then the observed CDAI score will be missing.

9.4.8 Simplified Disease Activity Index (SDAI)

SDAI is a composite continuous index to assess disease activity based on TJC28, SJC28, Patient's Global Assessment of Disease Activity (PtGA) (in cm, 0 – 10), Physician's

Global Assessment of Disease Activity (PhGA) (in cm, 0 – 10) and hsCRP (mg/dL). It can be derived as follows:

SDAI = TJC28 + SJC28 + PtGA (cm) + PhGA (cm) + hsCRP (mg/dL).

To calculate observed SDAI scores, the observed component value will be calculated first. Then the components will be included in the calculation per the SDAI formula selected. If any observed component is missing in a window, then the observed SDAI score will be missing.

9.4.9 Clinical Remission (CR) and Low Disease Activity (LDA)

Clinical remission (CR) and low disease activity (LDA) based on DAS28 (CRP) and CDAI are defined as follows:

Disease Status	DAS28 [CRP]	CDAI
LDA	≤ 3.2	≤ 10
CR	< 2.6	≤ 2.8

9.4.10 Disability Index of Health Assessment Questionnaire (HAQ-DI)

HAQ-DI is a self-reported patient outcome measurement. It is calculated as the mean of the scores from 8 following categories with a range 0 – 3: Dressing and Grooming, Rising, Eating, Walking, Hygiene, Reach, Grip, and Activities. The higher the score, the more likely to associate with morbidity and mortality for the RA patient.

The maximum score for all the questions in each category is considered as the score for the category. The Standard disability index (HAQ DI) takes into account the subject's use of aids or devices or assistance in the scoring algorithm for a disability category. For each of the eight disability categories these is an AIDS OR DEVICES companion variable(s) that is used to record the type of assistance, if any, a subject uses for his/her usual activities. If aids or devices and/or assistance from another person are checked for a disability category, the score for this category is set to 2 (much difficulty), if the original

score is 0 (no difficulty) or 1 (some difficulty). The HAQ DI is then calculated by summing the adjusted categories scores and dividing by the number of categories answered. The HAQ DI cannot be calculated if the patient does not have scores for at least 6 categories.

Minimum clinically important difference in HAQ-DI is defined as change from Baseline ≤ –0.22.

9.4.11 Functional Assessment of Chronic Illness Therapy (FACIT) – Fatigue Scale

Fatigue is one of the most frequent complaints of individuals with RA and is strongly associated with loss of independence, decreased physical activity and functional decline. One validated tool to measure fatigue is the FACIT Fatigue scale v4. The FACIT Fatigue Scale is a short, 13-item, easy to administer tool that measures an individual's level of fatigue during their usual daily activities over the past week. The level of fatigue is measured on a four point Likert scale (4 = not at all fatigued to 0 = very much fatigued). The Fatigue scale is ranged from 0 to 52 and the higher the score, the better the quality of life.

Item score for each item is calculated by either subtracting from 4 or adding 0 depending on whether it is a reversal item or not. FACIT Fatigue Scale is then calculated by adding up all item scores, multiplying by 13 and dividing by the number of items answered. It is essentially a prorated subscale if there are missing values for some items. If less than or equal to 50% of the items are answered (e.g., 6 out of 13), the proration is not acceptable and the scale will not be computed. Detailed calculation and missing value handling are documented in Appendix A.

9.4.12 Work Instability Scale for RA (RA-WIS)

The 23-item RA-WIS is a simple validated tool to evaluate Work Instability (the consequence of a mis-match between an individual's functional ability and their work tasks). It can be self-administered by the patients. To calculate the RA-WIS score,

one can simply add up the number of "true" responses. The RA-WIS can be scored in three bands indicating low (< 10), medium (10 – 17), and high (> 17) risk of work disability. The state of 'mismatch' between functional incapacity and work demands can threaten continuing employment if not resolved.

9.4.13 EuroQoL-5D (EQ-5D)

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

10.0 Safety Analysis

10.1 General Considerations

Safety analyses will include reporting of adverse events, laboratory, and vital signs measurements in open-label treatment period only. Safety analyses will be carried out using the Open-Label Treated Population, which includes all subjects who received at least one dose of open-label study medication.

Long-term safety analyses include reporting of AE rate adjusted by cumulative exposure, mean change from baseline in laboratory parameters and vital sign variables, and frequency of potentially clinically significant laboratory and vital signs values.

The treatment-emergent adverse event (TEAE) rate per 100 patient-years of exposure will be presented by the exposure groups defined in Section 8.1, namely

1. Upadacitinib 6 mg BID/15 mg QD
2. Upadacitinib 12 mg BID/30 mg QD
3. Upadacitinib 6 mg BID/15 mg QD post down titration
4. Any upadacitinib

Listing of subjects with TEAEs by SOC and PT will be provided. Frequency tables and listings of subjects meeting criteria for potentially clinically significant vital sign values and for potentially clinically significant laboratory values will be provided by the exposure groups defined in Section 8.1.

Mean changes from baseline in all continuous laboratory parameters and vital signs variables at each visit will be summarized by the treatment group sequences defined in Section 9.1, namely

1. Upadacitinib never titrated (6 mg BID/15 mg QD),
2. Upadacitinib titrated up and not down (6 mg BID/15 mg QD to 12 mg BID/30 mg QD), and
3. Upadacitinib titrated up and down (6 mg BID/15 mg QD to 12 mg BID/30 QD and then back down to 6 mg BID/15 mg QD).

10.2 Analysis of Adverse Events

A treatment-emergent Adverse Event (TEAE) is defined as an adverse event with an onset date that is on or after the first dose of open-label study drug, and no more than 30 days after the last dose of open-label study drug; or an adverse event with onset date before the first dose of open-label study drug, but increased in severity on or after the first dose of

open-label study drug, and no more than 30 days after the last dose of open-label study drug.

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 that the event was not treatment-emergent (e.g., the event end date was prior to the open-label study drug start date).

Adverse event data will be presented by system organ classes (SOCs) and preferred terms (PTs) using the Medical Dictionary for Drug Regulatory Activities (MedDRA) version 19.0 or most up to date version. All adverse event tables will be sorted in alphabetical order by SOC and PT and descending percentages in the Any upadacitinib group.

10.2.1 Overview of Adverse Events Rates per 100 Patient-Years of Study Drug Exposure

The number and percentage of subjects experiencing TEAEs will be summarized for each exposure group in the open-label treated population for the following AE categories:

- All TEAEs
- Treatment-emergent serious adverse events (SAEs)
- Treatment-emergent severe adverse events
- TEAEs reasonably possibly related to study drug
- TEAEs of special interest
- TEAEs leading to discontinuation of study drug
- TEAE leading to death
- All death

For this calculation, 1 year will be considered to be 365.25 days. For each exposure group, the numerator of the overall rate will be the total number of TEAEs reported for

the event; that is, a subject can contribute more than one event to the numerator. For each treatment group, the denominator of the rates will be the total number of days exposed to study drug summed across all treated subjects divided by 365.25. Please refer to Section 8.1 for the calculation of study drug exposure. The AE rate per 100 patient-years of exposure will be calculated as ([numerator/denominator])*100. The number of AEs reported (numerator), the total number of years of study drug exposure (denominator), and the AE rate per 100 patient-years will be presented for each exposure group.

In the AE overview summary, any event with an unknown severity will be considered as severe and any AE with an unknown relationship will be considered as having "reasonable possibility" of being related to study drug.

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

10.2.2 Adverse Events Rates per 100 Patient-Years of Study Drug Exposure by System Organ Class and Preferred Term

the TEAE rate per 100 patient-years of exposure will be tabulated by SOC and MedDRA PT for each exposure group in the open-label treated population. The SOCs will be presented in alphabetical order, and the PTs will be presented in alphabetical order within each SOC.

The following summaries of adverse events will be generated:

- All TEAEs
- Treatment-emergent serious adverse events (SAEs)
- Treatment-emergent severe adverse events
- TEAEs reasonably possibly related to study drug
- Frequent TEAEs (reported at ≥ 2 E/100PYin any exposure group)
- TEAEs leading to discontinuation of study drug
- TEAE leading to death

10.2.3 Serious Adverse Events (Including Deaths) and Adverse Events Leading to Study Drug Discontinuation Rates per 100 Patient-Years of Study Drug Exposure

All serious adverse events (SAEs), deaths, and adverse events leading to discontinuation of study drug will be summarized by SOC and PT or each exposure group.

10.2.4 Frequent Adverse Events by System Organ Class and Preferred Term

TEAEs occurring at ≥ 2 E/100PY in any of the exposure groups will be summarized by MedDRA PT in decreasing rate.

10.2.5 Adverse Events of Special Interests Rates per 100 Patient-Years of Study Drug Exposure

The Adverse Events of Special Interests (AESI) categories will be summarized and presented for each exposure group using SOC and MedDRA PT. The AESI categories will be identified per Standard MedDRA Queries (SMQs)/Company MedDRA Queries (CMQs) in Table 5 below. Adjudicated cardiovascular events will be summarized and presented by treatment groups using CAC adjudicated categories.

Table 5. AESI for Upadacitinib with SMQs/CMQs/PTs Searches

AESI	Type of MedDRA Query	Broad or Narrow Search	SMQ/CMQ Search Criteria
Serious Infections	CMQ		"Infections" – Subset for SAEs
Opportunistic Infection excluding Tuberculosis and Herpes Zoster	CMQ		"Opportunistic Infection excluding Tuberculosis and Herpes Zoster"
Possible Malignancy	SMQ	Narrow	"Malignancies"
Malignancy	SMQ		"Malignant tumours"
Non-Melanoma Skin Cancer (NMSC)	SMQ	Narrow	Skin Malignant tumours (Narrow SMQ) removing Melanoma CMQ
Malignancy excluding NMSC			"Malignant tumours" SMQ removing NMSC output
Lymphoma	SMQ		"Malignant Lymphomas"
Hepatic Disorder	SMQ	Narrow	"Drug Related Hepatic Disorders"
Adjudicated Gastrointestinal Perforations	Output from adjudication		
Anemia	CMQ		"Non-Hemolytic and Non-Aplastic Anemias"
Neutropenia	CMQ		"Hematological Toxicity – Neutropenia"
Lymphopenia	CMQ		"Hematological Toxicity – Lymphopenia"
Herpes Zoster	CMQ		"Herpes Zoster"
Creatine Phosphokinase (CPK) Elevation	PT		Search only for the PT of "Blood creatine phosphokinase increased"
Renal Dysfunction	SMQ	Narrow	"Acute Renal, Failure"
Active Tuberculosis	CMQ		"Active Tuberculosis"
Adjudicated Cardiovascular Events	Output from CAC		
MACE*			
Cardiovascular Death			
Non-fatal Myocardial Infarction			
Non-fatal Stroke			

Table 5. AESI for Upadacitinib with SMQs/CMQs/PTs Searches (Continued)

AESI	Type of MedDRA Query	Broad or Narrow Search	SMQ/CMQ Search Criteria
Undetermined/Unknown Cause of Deaths			
Other Cardiovascular events			
Adjudicated Thrombotic Events	Output From CAC		
Venous Thromboembolic Events**			
Deep Vein Thrombosis			
Pulmonary Embolism			
Other Venous Thrombosis			
Arterial Thromboembolic Events (non-cardiac, non-neurologic)			

CAC = Cardiovascular Adjudication Committee; CMQ = company MedDRA query; PT = preferred term; SMQ = standard MedDRA query

* MACE; Major Adverse Cardiovascular Events, defined as cardiovascular death, non-fatal myocardial infarction and non-fatal stroke.

** Venous thromboembolic events (VTE) include deep vein thrombosis (DVT) and pulmonary embolism (PE) (fatal and non-fatal).

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

10.2.6 Listing of 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.

10.3 Analysis of Laboratory Data

10.3.1 Variables and Units

All laboratory parameters to be collected in this study are listed below. Laboratory parameters will be reported using the standard international (SI) units.

Table 6. List of Laboratory Variables

Laboratory Variables
Hematology
White Blood Cell (WBC) Count
Red Blood Cell (RBC) Count
Hemoglobin
Hematocrit
Platelets count
Neutrophils
Basophils
Eosinophils
Lymphocytes
Monocytes
Reticulocytes
Bands
Chemistry
Total Bilirubin
Alkaline Phosphatase (ALP)
Serum glutamic oxaloacetic transaminase (SGOT/AST)
Serum glutamic pyruvic transaminase (SGPT/ALT)
Total Protein
Albumin
Glucose
Triglycerides
Blood Urea Nitrogen (BUN)
Creatinine
Uric acid

Table 6. List of Laboratory Variables (Continued)

Laboratory Variables
Chemistry (Continued)
Sodium
Potassium
Calcium
Inorganic Phosphorus
Creatine Phosphokinase (CPK)
Chloride
Bicarbonate
Cholesterol
LDL cholesterol
HDL cholesterol
Urinalysis
Specific Gravity
pH
Protein
Glucose
Ketones
Blood
Microscopic Examination
Immunologic Laboratory Assessments
Rheumatoid Factor (RF)
Anti-CCP antibody
Advanced Lipid Panel
Apo A1
Apo B
Other
hSCRP
HBs Ab
HBc Ab
Hbs Ag

Table 6. List of Laboratory Variables (Continued)

Laboratory Variables
Other (Continued)
HBV DNA PCR reflex only
HCV Ab
HCV RNA reflex only
QuantiFERON-TB Gold

10.3.2 Assessment of Mean Change from Baseline in Clinical Laboratory Values

Analyses of mean change from baseline in continuous hematology, chemistry, and urinalysis variables which are measured longitudinally will be performed by visits and by treatment group. For each change from baseline analysis, the following summary statistics will be presented for each treatment group sequence: sample size, baseline mean, visit mean, and the mean, standard deviation, and median of the changes from baseline.

10.3.3 Assessment of Shift from Baseline in Clinical Laboratory Variables

The baseline and post-baseline laboratory observations will be categorized as Grade 1, Grade 2, Grade 3, and Grade 4 according to OMERACT criteria (Rheumatology Common Toxicity Criteria v.2.0). For creatine phosphokinase and creatinine NCI CTC criteria will be used.

For each laboratory variable, shift tables will be generated that cross tabulate the subjects as deemed appropriate by treatment sequence:

- Category of the baseline value versus category of the final value.
- Category of the baseline value versus maximum category.
- Category of the baseline value versus minimum category.

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

No statistical tests will be performed for this analysis.

10.3.4 Potentially Clinically Significant Laboratory Values

The criteria for potentially clinically significant laboratory values will be determined by OMERACT criteria of Grade 3 or 4. For creatine phosphokinase, creatinine, and parameters that are not covered in the OMERACT criteria, NCI CTC criteria will instead be used. For laboratory parameters where values available, a listing of all subjects with any laboratory determination during each period meeting OMERACT criteria of Grade 3 or 4 will be provided by Grade. For each of these subjects, the whole course of the respective parameter will be listed.

10.3.5 Analyses for Liver Elevations

According to FDA's Guidance for Industry "Drug-Induced Liver Injury: Premarketing clinical evaluation (July 2009), when aminotransferase (AT) abnormalities indicating hepatocellular injury are accompanied by evidence of impaired hepatic function (bilirubin elevation > 2 × ULN), in the absence of evidence of biliary obstruction (i.e., significant elevation of ALP) or some other explanation of the injury (e.g., viral hepatitis, alcohol hepatitis), the combined finding (i.e., Hy's Law cases) represents a signal of a potential for the drug to cause severe DILI.

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

- ALT ≥ 3 × ULN
- ALT ≥ 5 × ULN
- ALT ≥ 10 × ULN

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

A listing of potentially clinically significant liver elevations based on criteria specified above will be provided. For each of these subjects, the whole course of the respective parameter will be listed.

10.4 Analysis of Vital Signs and Weight

10.4.1 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 7.

Table 7. 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 decrease ≥ 20 mmHg from Baseline
	High	Value ≥ 160 mmHg and increase ≥ 20 mmHg from Baseline
Diastolic blood pressure	Low	Value ≤ 50 mmHg and decrease ≥ 15 mmHg from Baseline
	High	Value ≥ 105 mmHg and increase ≥ 15 mmHg from Baseline
Weight	High	> 7% increase from baseline
	Low	> 7% decrease from baseline

10.4.2 Analysis of Long-Term Vital Signs

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

A listing of all subjects with any vital sign values meeting the criteria for potentially clinically significant vital signs will be provided. For each of these subjects, the whole course of the respective parameter will be listed.

11.0 Summary of Changes

Version Number	Date	Description of Revision
1.0	15 July 2015	Original Version
2.0	25 October 2016	This version of SAP includes the changes summarized below: Section 4.2 – Design Diagram • Update the length of study from 96 week to 264 week and the actual number of subjects participating in the study. • Update the information on switching to once-daily tablet formulation. Section 4.3 (in version 1.0) – Sample Size • Delete the section because it does not provide useful information. Section 4.3 – Interim Analysis • Remove the part of the description of efficacy/safety analysis that is no longer applicable to our updated analysis scheme. • Change the title to Interim Analysis because the section only contains information on interim analysis. Section 5.1 – Open-Label Treated Population • Change the wording for three treatment groups in open-label treated population to make them it clear. • Include the information on the use of as randomized open-label treated population for supportive efficacy analysis up to Week 24. Section 6.1 – Demographic and Baseline Characteristics • Include information to clarify that baseline is calculated at the baseline visit of the preceding RCT studies. • Include geographic region for demographic characteristics. • Include duration of prior DMARDs used (for Study M13-537 subjects), number of mechanisms of action for prior biologic DMARD(s) used (for Study M13-550 subjects). Section 6.3 – Previous Treatment and Concomitant Medications • Change the definition of prior medication to be any medication taken prior to the first dose of open-label study drug.

Version Number	Date	Description of Revision
		Section 8.1 – Study Drug exposure • Change the duration intervals from exclusive to cumulative. • Delete the part for the calculation of study drug exposure based on double-blind and open-label periods combined. Section 8.2 – Compliance • Change the definition of treatment compliance to account for the fact of switching to once-daily formulation. Section 9.1 – General Considerations • Update the analysis scheme to include a long term efficacy analysis as the main analysis, which is based on the open-label treated population. • Update the analysis scheme to include a supportive efficacy analysis up to Week 24, which is based on the as randomized open-label treated population and a subset of most important endpoints. Section 9.2 (in Version 1.0) – Exploratory Efficacy Variables • Delete the section because it is redundant. Section 9.2 – Handling of Multiplicity • Add information to clarify why handling of multiplicity is not needed. Section 9.4 – Efficacy Variables Definitions and Conventions • Include the suggestion of the use of the corresponding programming specification for details on the computation of composite endpoints. • Add a subsection for Joint Evaluation. • Add a subsection for Patient's Global Assessment of Disease Activity Visual Analog Scale (VAS). • Add a subsection for Physician's Global Assessment of Disease Activity Visual Analog Scale (VAS). • Add a subsection for Patient's Global Assessment of Pain. • Add a subsection for Simplified Disease Activity Index (SDAI). • Delete the subsection for SF-36. Section 9.4.1 – ACR Criteria • Remove the detailed instructions for the computation of ACR and relocate them into the programming specification. Section 9.4.9 – Low Disease Activity • Change the Criterion for LDA base on DAS28 (CRP) to be $2.6 \leq$ DAS28 (CRP) ≤ 3.2.

Version Number	Date	Description of Revision
		Section 10 – Safety Analysis • Remove the language involving statistical testing (i.e., ANCOVA, Fisher's exact test) because it is not meaningful to compare treatment groups in open-label treated population. • Update MedDRA version to 19.0. • Add the phrase "in open-label treated population" in multiple places to clarify the population on which analysis will be performed. Section 10.2 – Analysis of Adverse Events • Remove the part for summarizing SAEs collected between the signing of the informed consent and the first dose. Section 10.2.1.1 – Adverse Event Overview • Delete the description regarding how to handle unknown severity. Section 10.2.1.2 – Adverse Events by System Organ Class and Preferred Term • Update the language according to the current Immunology convention. Section 10.2.1.3 – Adverse Events by Maximum Severity • Change "Select TEAEs" to "TEAEs." Section 10.2.1.4 – Adverse Events by Maximum Relationship • Update the language regarding TEAE summarization to be "TEAEs will be summarized by maximum relationship to ABT-494 only." • Change "probably related or possibly related" to "Reasonable Possibility." Section 10.2.1.5 (in version 1.0) – Adverse Events by "Reasonably Possibly Related" Relationship • Delete the section due to redundancy. Section 10.2.1.7 – Adverse Events of Special Interests • Update the list for summarizing AESI. Section 10.2.1.8 – Adverse Events by 100 Patient Years • Include information on summarizing AESI. Section 10.3.2.4 – Potentially Clinically Significant Laboratory Values • Add NCI CTC criteria for parameters not covered in the OMERACT criteria.

Version Number	Date	Description of Revision
		Section 10.4.1 – Variables and Criteria Defining Abnormality • Update the table "Criteria for Potentially Clinically Significant Vital Sign Findings" according to the current Immunology convention.
3.0	03 May 2021	This version of SAP includes the changes summarized below: Section 3.0 – Introduction • Remove disease background to align with phase 3 standards. Section 4.2 – Overall Study Design and Plan • Update the length of study from 264 weeks to 312 weeks. • Include details of Amendment 5 providing the option to decrease dose to 15 mg QD. • Add clarification that follow-up visit is available after completion of OLE last visit or after PD visit. Section 4.3 – Data Monitoring Committee (DMC) Activities • Include to align with phase 3 standards. Section 5.1 – Analysis Populations • Remove treatment group sequences and relocate to efficacy analysis. • Remove As Randomized Open-Label Treated Population. Section 5.2 – Definition of Baseline • Add clarification that RCT baseline will be used for both efficacy and safety analysis. Section 5.3 – Analysis Windows • Include to align with phase 3 standards. Section 5.3 (in Version 2.0) – Variables Used for Stratification of Randomization • Remove since it is not needed for this study. Section 5.4 (in Version 2.0) – Subgroups Definition • Remove since it is not needed for this study. Section 6.1 – Demographics and Baseline Characteristics • Update to align with phase 3 standards where applicable. Section 6.3 – Prior medication • Update to define prior medication as any medication prior to the first dose of preceding RCT study. • Update language to align with phase 3 standards. Section 6.4 – Protocol Deviations • Include to align with phase 3 standards. Section 7.0 – Patient Disposition

Version Number	Date	Description of Revision
		• Add Other category in reasons for discontinuation. Section 8.1 – Study Drug Exposure • Add exposure groups for the Open-Label Treated Population. • Update duration of exposure to utilize exposure group rather than treatment sequence group. Section 9.1 – Efficacy Analysis General Considerations • Remove Supportive Efficacy Analysis Up to Week 24. • Add treatment group sequences relocated from analysis population. Section 9.1.1 – Definition of Missing Data Imputation • Add to describe imputation for MMRM and GLMM models. Section 9.2 – Long-Term Efficacy Analysis • Add to elaborate on long-term efficacy analysis. • Update assessments after Week 96 to occur every 12 weeks to be consistent with data collection. • Add proportion of subjects achieving SDAI criteria as reported measurement. • Add long-term longitudinal analysis performed using MMRM or GLMM performed only for treatment sequences 1 and 2. • Update Table 2 according to above changes. Section 9.3 (in Version 2.0) – Efficacy Subgroup Analysis • Remove since it is not needed for this study. Section 9.3.1 – ACR Criteria • Add definition of ACR20, ACR50, and ACR70 along with illustrated example • Add windowing rule for ACR response calculation. Section 9.3.6 – Disease Activity Score Based on CRP (DAS28[CRP]) • Update formula to provide descriptive footnotes. • Add description of Observed DAS28 score calculation. Section 9.3.7 – Clinical Disease Activity Index (CDAI) • Add description of Observed CDAI score calculation. Section 9.3.8 – Simplified Disease Activity Index (SDAI) • Add description of Observed SDAI score calculation. Section 9.3.9 – Clinical Remission (CR) and Low Disease Activity (LDA) • Replace Section 9.4.9 (in Version 2.0) and Section 9.4.10 (in Version 2.0) with table of definitions. Section 10.1 – Safety Analysis General Considerations

Version Number	Date	Description of Revision
		• Change the AE summary to be presented by the added exposure groups. • Add clarification that the laboratory value analysis will be presented by treatment sequence groups. Section 10.2.1 – Adverse Event Rates per 100 Patient-Years • Change the TEAE summary to be presented by the added exposure groups. • Remove Frequent TEAE categories. Section 10.2.1 – Adverse Event Rates per 100 Patient-Years • Change the TEAE by SOC and MedDRA PT summary to be presented by the added exposure groups. • Redefine Frequent TEAE category to greater-than or equal-to 2 events per 100 patient-years. • Remove Frequent TEAE reasonably possibly related to study drug category. Section 10.2.3 (in Version 2.0) – TEAEs by Maximum Severity • Delete since this is not reported for long-term. Section 10.2.4 (in Version 2.0) – TEAEs by Maximum Relationship • Delete since this is not reported for long-term. Section 10.2.3 – Serious Adverse Events and Adverse Events Leading to Study Drug Discontinuation • Change the SAE, deaths, and adverse events leading to study drug discontinuation summary to be presented by the added exposure groups. Section 10.2.4 – Frequent AE by SOC and MedDRA PT • Redefine Frequent TEAEs to greater-than or equal-to 2 events per 100 patient-years. Section 10.2.5 – Safety Analysis General Considerations • Change the AESI summary to be presented by exposure groups. Section 10.3.3 – Assessment of Shift from Baseline in Clinical Laboratory Variables • Update shift from baseline analysis to align with phase 3 standards where applicable. Section 10.5 (in Version 2.0) – Analysis of ECG Parameters • Delete due to redundancy with baseline characteristics. Appendix D – Statistical Analysis to Account for Impact of COVID-19 Pandemic

Version Number	Date	Description of Revision
		• Add detail of additional patient disposition summaries and long-term data handling required to account for impact of COVID-19

12.0 Appendix

Appendix A. FACIT-Fatigue Scale Calculation

FACIT-Fatigue Subscale Scoring Guidelines (Version 4)

Instructions:*

1. Record answers in "item response" column. If missing, mark with an X.
2. Perform reversals as indicated, and sum individual items to obtain a score.
3. Multiply the sum of the item scores by the number of items in the subscale, then divide by the number of items answered. This produces the subscale score.
4. **The higher the score, the better the QOL.**

Subscale	Item Code	Reverse Item?		Item Response		Item Score
FATIGUE	HI7	4	-	_____	=	_____
SUBSCALE	HI12	4	-	_____	=	_____
	An1	4	-	_____	=	_____
	An2	4	-	_____	=	_____
	An3	4	-	_____	=	_____
	An4	4	-	_____	=	_____
Score range: 0 – 52	An5	0	+	_____	=	_____
	An7	0	+	_____	=	_____
	An8	4	-	_____	=	_____
	An12	4	-	_____	=	_____
	An14	4	-	_____	=	_____
	An15	4	-	_____	=	_____
	An16	4	-	_____	=	_____
					Sum individual item scores:	_____
					Multiply by 13:	_____
					Divide by number of items answered:	_____ **= Fatigue**

Subscale Score

Handling missing items. If there are missing items, subscale scores can be prorated. This is done by multiplying the sum of the subscale by the number of items in the subscale, then dividing by the number of items actually answered. This can be done on the scoring guide or by using the formula below:

Prorated subscale score = [Sum of item scores] × [N of items in subscale] ÷ [N of items answered]

Appendix B. EQ5D-5L Manual

Please refer

http://www.euroqol.org/fileadmin/user_upload/Documenten/PDF/Folders_Flyers/UserGuide_EQ-5D-5L_v2.0_October_2013.pdf.

Appendix C. OMERACT Criteria

Rheumatology Common Toxicity Criteria v.2.0 **Based on Woodworth TG, et al. Standardizing assessment of adverse effects in rheumatology clinical trials II. Status of OMERACT Drug Safety Working Group May 2006: OMERACT 8. Standardizing Assessment and Reporting of Adverse Effects in Rheumatology Clinical Trials: Enabling Description of Comparative Safety Profiles for Antirheumatic Therapies**				
	1 – Mild Asymptomatic, or transient Short duration (< 1 week) No change in life style No medication or OTC	**2 – Moderate** Symptomatic Duration (1 – 2 weeks) Alter lifestyle occasionally Meds relieve (may be prescription) Study drug continued	**3 – Severe** Prolonged symptoms, reversible, major functional impairment Prescription meds/partial relief May be hospitalized < 24 hr Temporary study drug discontinuation, or/and dose reduced	**4 – Includes Life Threatening** At risk of death Substantial disability, especially if permanent Multiple meds Hospitalised > 24 hr Study drug discontinued
A. Allergic/Immunologic				
A1. Allergic reaction/hypersensitivity (includes drug fever)	Transient rash: drug fever < 38°C: transient, asymptomatic bronchospasm	Generalised urticaria responsive to meds; or drug fever > 38°C, or reversible bronchospasm	Symptomatic bronchospasm requiring meds; symptomatic urticaria persisting with meds, allergy related oedema/angioedema	Anaphylaxis, laryngeal/pharyngeal oedema, requiring resuscitation
A2. Autoimmune reaction	Seriologic or other evidence of autoimmune reaction, but patient asymptomatic: all organ function normal and no treatment is required (e.g., vitiligo)	Evidence of autoimmune reaction involving a non-essential organ or functions, requiring treatment other than immunosuppressive drugs (e.g., hypothyroidism)	Reversible autoimmune reaction involving function of a major organ or toxicity requiring short term immunosuppressive treatment (e.g., transient colitis or anaemia)	Causes major organ dysfunction, or progressive, not reversible, or requires long term administration of high dose immunosuppressive therapy
A3. Rhinitis (includes sneezing, nasal stuffiness, post-nasal discharge)	Transient, non-prescription meds relieve	Prescription med. required, slow	Corticosteroids or other prescription med. with persistent disabling symptoms such as impaired exercise tolerance	NA

Rheumatology Common Toxicity Criteria v.2.0 **Based on Woodworth TG, et al. Standardizing assessment of adverse effects in rheumatology clinical trials II. Status of OMERACT Drug Safety Working Group May 2006: OMERACT 8. Standardizing Assessment and Reporting of Adverse Effects in Rheumatology Clinical Trials: Enabling Description of Comparative Safety Profiles for Antirheumatic Therapies**				
A. Allergic/Immunologic (Continued)				
A4. Serum sickness	Transient, non-prescription meds relieve	Symptomatic, slow response to meds (e.g., oral corticosteroids)	Prolonged; symptoms only partially relieved by meds; parenteral corticosteroids required	Major organ dysfunction, requires long-term high-dose immunosuppressive therapy
A5. Vasculitis	Localised, not requiring treatment; or rapid response to meds; cutaneous	Symptomatic, slow response to meds (e.g., oral corticosteroids)	Generalised, parenteral corticosteroids required or/and short duration hospitalisation	Prolonged, hospitalisation, ischemic changes, amputation
B. Cardiac				
B1. Arrhythmia	Transient, asymptomatic	Transient, but symptomatic or recurrent, responds to meds	Recurrent/persistent; maintenance prescription	Unstable, hospitalisation required, parenteral meds
B2. Cardiac function decreased	Asymptomatic decline in resting ejection fraction by > 10%, but < 20% of baseline value	Asymptomatic decline of resting ejection fraction ≥ 20% of baseline value	CHF responsive to treatment	Severe or refractory CHF
B3. Edema	Asymptomatic (e.g., 1 + feet/calves), self-limited, no therapy required	Symptomatic (e.g., 2 + feet/calves), requires therapy	Symptoms limiting function (e.g., 3 + feet/calves, 2 + thighs), partial relief with treatment prolonged	Anasarca; no response to treatment
B4. Hypertension (new onset or worsening)	Asymptomatic, transient increase by > 20 mmHg (diastolic) or to > 150/100 if previously normal, no therapy required	Recurrent or persistent increase > 150/100 or by > 10 mmHg (diastolic), requiring and responding readily to treatment	Symptomatic increase > 150/100, > 20 mmHg, persistent, requiring multi agency therapy, difficult to control	Hypertensive crisis
B5. Hypotension (without underlying diagnosis)	Transient, interOpen-label treatedent, asymptomatic, orthostatic decrease in blood pressure > 20 mmHg	Symptomatic, without interference with function, recurrent or persistent > 20 mmHg decrease, responds to treatment	Syncope or symptomatic, interferes with function, requiring therapy and sustained medical attention, dose adjustment or drug discontinuation	Shock

Rheumatology Common Toxicity Criteria v.2.0 **Based on Woodworth TG, et al. Standardizing assessment of adverse effects in rheumatology clinical trials II. Status of OMERACT Drug Safety Working Group May 2006: OMERACT 8. Standardizing Assessment and Reporting of Adverse Effects in Rheumatology Clinical Trials: Enabling Description of Comparative Safety Profiles for Antirheumatic Therapies**				
B. Cardiac (Continued)				
B6. Myocardial ischaemia	Transient chest pain/ECG changes; rapid relief with nitro	Recurring chest pain, transient ECG ST-T changes; treatment relieves	Angina with infarction, no or minimal functional compromise, reduce dose or discontinue study drug	Acute myocardial infarction, arrthymia or/and CHF
B7. Pericarditis/ pericardial effusion	Rub heard, asymptomatic	Detectable effusion by echocardiogram, symptomatic NSAID required	Detectable on chest x-ray, dyspnoea; or pericardiocentesis; requires corticosteroids	Pulsus alternates with low cardiac output; requires surgery
B8. Phlebitis/thrombosis/ Embolism (excludes injection sites)	Asymptomatic, superficial, transient, local, or no treatment required	Symptomatic, recurrent, deep vein thrombosis, no anticoagulant therapy required	Deep vein thrombosis requiring anticoagulant therapy	Pulmonary embolism
C. General (Constitutional)				
C1. Fatigue/malaise (asthenia)	Increase over baseline; most usual daily functions maintained, short term	Limits daily function interOpen-label treatedently over time	Interferes with basic ADL, persistent	Unable to care for self, bed or wheelchair bound > 50% of day debilitating, hospitalisation
C2. Fever (pyrexia) (note: fever due to drug allergy should be coded as allergy)	Transient, few symptoms 37.7 – 38.5°C	Symptomatic, recurrent 38.6 – 39.9°C. Relieved by meds	≥ 40°C; ≤ 24 h, persistent symptoms; partial response to meds	≥ 40°C, debilitating, > 24 h, hospitalisation; no relief with meds
C3. Headache	Transient or interOpen-label treatedent, no meds or relieved with OTC	Persistent, recurring, non-narcotic analgesics relieve	Prolonged with limited response to narcotic medicine	Intractable, debilitating, requires parenteral meds
C4. Insomnia	Difficulty sleeping, short term, no interfering with function	Difficulty sleeping interfering with function, use of prescription med	Prolonged symptoms, with limited response to narcotic meds	Debilitating, hospitalisation; no relief with meds

Rheumatology Common Toxicity Criteria v.2.0 **Based on Woodworth TG, et al. Standardizing assessment of adverse effects in rheumatology clinical trials II. Status of OMERACT Drug Safety Working Group May 2006: OMERACT 8. Standardizing Assessment and Reporting of Adverse Effects in Rheumatology Clinical Trials: Enabling Description of Comparative Safety Profiles for Antirheumatic Therapies**				
C. General (Constitutional) (Continued)				
C5. Rigors, chills	Asymptomatic, transient, no meds, or non-narcotic meds relieve	Symptomatic, narcotic meds relieve	Prolonged symptoms, with limited response to narcotic meds	Debilitating, hospitalisation; no relief with meds
C6. Sweating (diaphoresis)	Episodic, transient	Frequent, short term	Frequent, drenching, disabling	Dehydration, requiring IV fluids/hospitalization > 24 hrs
C7. Weight gain	5% – 9.9%	10% – 19.9%	20% – 30%	NA
C8. Weight loss	5% – 9.9%	10% – 19.9%	20% – 30%	NA
D. Dermatologic				
D1. Alopecia	Subjective, transient	Objective, fully reversible	Patchy, wig used, partly reversible	Complete, or irreversible even if patchy
D2. Bullous eruption	Localised, asymptomatic	Localised, symptomatic, requiring treatment	Generalised, responsive to treatment; reversible	Prolonged, generalised, or requiring hospitalisation for treatment
D3. Dry skin	Asymptomatic, controlled with emollients	Symptoms eventually (1 – 2 wks) controlled with emollients	Generalised, interfering with ADL > 2 wks, persistent pruritis, partially responsive to treatment	Disabling for extended period, unresponsive to ancillary therapy and requiring study drug discontinuation for relief
D4. Injection site reaction	Local erythema, pain, pruritis, < few days	Erythema, pain, oedema, may include superficial phlebitis, 1 – 2 wks	Prolonged induration, superficial ulceration; includes thrombosis	Major ulceration necrosis requiring surgery
D5. Petechiae (without vasculitis)	Few, transient asymptomatic	Dependent areas, persistent up to 2 wks	Generalised, responsive to treatment; reversible	Prolonged, irreversible, disabling

Rheumatology Common Toxicity Criteria v.2.0 **Based on Woodworth TG, et al. Standardizing assessment of adverse effects in rheumatology clinical trials II. Status of OMERACT Drug Safety Working Group May 2006: OMERACT 8. Standardizing Assessment and Reporting of Adverse Effects in Rheumatology Clinical Trials: Enabling Description of Comparative Safety Profiles for Antirheumatic Therapies**				
D. Dermatologic (Continued)				
D6. Photosensitivity	Transient erythema	Painful erythema and oedema requiring topical treatment	Blistering or desquamation, requires systematic corticosteroids	Generalised exfoliation or hospitalisation
D7. Pruritis	Localised, asymptomatic, transient, local treatment	Intense, or generalised, relieved by systematic medication	Intense or generalised; poorly controlled despite treatment	Disabling, irreversible
D8. Rash (not bullous)	Erythema, scattered macular/popular eruption; pruritis transient; TOC or no meds	Diffuse macular/popular eruption or erythema with pruritus; dry desquamation; treatment required	Generalised, moist desquamation, requires systemic corticosteroids; responsive to treatment; reversible	Exfoliative or ulcerating; or requires hospitalisation; or parenteral corticosteroids
D9. Induration/fibrosis/ Thickening (not sclerodermal)	Localized, high density on palpation, reversible, no effect on ADL and not disfiguring	Local areas < 50% body surface, not disfiguring, transient interference with ADL, reversible	Generalized, disfiguring, interferes with ADL, reversible	Disabling, irreversible, systemic symptoms
E. Ear/Nose/Throat				
E1. Hearing loss	Transient, interOpen-label treatedent, no interference with function	Symptomatic, treatment required, reversible	Interferes with function; incomplete response to treatment	Irreversible deafness
E2. Sense of smell	Slightly altered	Markedly altered	Complete loss, reversible	Complete loss, without recovery
E3. Stomatitis	Asymptomatic	Painful, multiple, can eat	Interferes with nutrition, slowly reversible	Requires enteral support; residual dysfunction
E4. Taste disturbance (dysgeusia)	Transiently altered; metallic	Persistently altered; limited effect on eating	Disabling, effect on nutrition	NA
E5. Tinnitus	InterOpen-label treatedent, transient, no interference with function	Requires treatment, reversible	Disabling, or associated with hearing loss	Irreversible deafness

Rheumatology Common Toxicity Criteria v.2.0 **Based on Woodworth TG, et al. Standardizing assessment of adverse effects in rheumatology clinical trials II. Status of OMERACT Drug Safety Working Group May 2006: OMERACT 8. Standardizing Assessment and Reporting of Adverse Effects in Rheumatology Clinical Trials: Enabling Description of Comparative Safety Profiles for Antirheumatic Therapies**				
E. Ear/Nose/Throat (Continued)				
E6. Voice changes (includes hoarseness, loss of voice, laryngitis)	InterOpen-label treatedent hoarseness, able to vocalise	Persistent hoarseness, able to vocalise	Whispered speech, slow return of ability to vocalise	Unable to vocalize for extended
E7. Xerostomia (dry mouth)	Transient dryness	Relief with meds	Interferes with nutrition, slowly reversible	Extended duration interference with nutrition, requires parenteral nutrition
F. Eye/Ophthalmologic				
F1. Cataract	Asymptomatic, no change in vision, non-progressive	Symptomatic, partial visual loss, progressive	Symptoms impairing function, vision loss requiring treatment, including surgery	NA
F2. Conjunctivitis	Asymptomatic, transient, rapid response to treatment	Symptomatic, responds to treatment, changes not interfering with function	Symptoms prolonged, partial response to treatment, interferes with function	NA
F3. Lacrimation increased (tearing, watery eyes)	Symptoms not requiring treatment, transient	Symptomatic, treatment required, reversible	Unresponsive to treatment with major effect on function	NA
F4. Retinopathy	Asymptomatic, non-progressive, no treatment	Reversible change in vision; readily responsive to treatment	Disabling change in vision ophthalmological findings reversible, sight improves over time	Loss of sight
F5. Vision changes (e.g., blurred, photophobia, night blindness, vitreous floaters)	Asymptomatic, transient, no treatment required	Symptomatic, vision changes not interfering with function, reversible	Symptomatic, vision changes interfering with function	Loss of sight

Rheumatology Common Toxicity Criteria v.2.0 **Based on Woodworth TG, et al. Standardizing assessment of adverse effects in rheumatology clinical trials II. Status of OMERACT Drug Safety Working Group May 2006: OMERACT 8. Standardizing Assessment and Reporting of Adverse Effects in Rheumatology Clinical Trials: Enabling Description of Comparative Safety Profiles for Antirheumatic Therapies**				
F. Eye/Ophthalmologic (Continued)				
F6. Xerophtalmia (dry eyes)	Mild scratchiness	Symptomatic without interfering with function, requires artificial tears	Interferes with vision/function, corneal ulceration	Loss of sight
G. Gastrointestinal				
G1. Anorexia	Adequate food intake, minimal weight loss	Symptoms requiring oral nutritional supplementation	Prolonged, requiring iv support	Requires hospitalization for nutritional support
G2. Constipation	Asymptomatic, transient, responds to stool softener, OTC laxatives	Symptomatic, requiring prescription laxatives, reversible	Obstipation requiring medical intervention	Bowel obstruction. Surgery required
G3. Diarrhea	Transient, increase of 2 – 3 stools/day over pre-treatment (no blood or mucus), OTC agents relieve	Symptomatic, increase 4 – 6 stools/day, nocturnal stools, cramping, requires treatment with prescription meds	Increase > 6 stools/day, associated with disabling symptoms, e.g., incontinence, severe cramping, partial response to treatment	Prolonged, dehydration, unresponsive to treatment, requires hospitalization
G4. Dyspepsia (heartburn)	Transient, interOpen-label treatedent, responds to OTC antacids, H-2 blockers	Prolonged, recurrent, requires prescription meds, relieved by meds	Persistent despite treatment, interferes with function, associated with GI bleeding	NA
G5. GI bleed (gastritis, gastric or duodenal ulcer diagnosed-define aetiology)	Asymptomatic, endoscopic finding, haemocult + stools, no transfusion, responds rapidly to treatment	Symptomatic, transfusion ≤ 2 units needed; responds to treatment	Haematemesis, transfusion 3 – 4 units, prolonged interference with function	Recurrent, transfusion > 4 units, perforation, requiring surgery, hospitalisation
G6. Haematochezia (rectal bleeding)	Haemorrhodial, asymptomatic, no transfusion	Symptomatic, transfusion ≤ 2 units, reversible	Recurrent, transfusion > 3 – 4 units	> 4 units, hypotension, requiring hospitalization

Rheumatology Common Toxicity Criteria v.2.0 **Based on Woodworth TG, et al. Standardizing assessment of adverse effects in rheumatology clinical trials II. Status of OMERACT Drug Safety Working Group May 2006: OMERACT 8. Standardizing Assessment and Reporting of Adverse Effects in Rheumatology Clinical Trials: Enabling Description of Comparative Safety Profiles for Antirheumatic Therapies**				
G. Gastrointestinal (Continued)				
G7. Hepatitis	Laboratory abnormalities, asymptomatic, reversible	Symptomatic laboratory abnormalities, not interfering with function, slowly reversible	Laboratory abnormalities persistent > 2 wks, symptoms interfere with function	Progressive, hepato-renal, anasarca, pre-coma or coma
G8. Nausea, or nausea/vomiting (use diagnostic term)	Transient, interOpen-label treatedent, minimal interference with intake, rapid response to meds	Persistent, recurrent, requires prescription meds, intake maintained	Prolonged, interferes with daily function and nutritional intake, periodic iv fluids	Hypotensive, hospitalization, parenteral nutrition, unresponsive to out-patient management
G9. Pancreatitis	Anylase elevation, interOpen-label treatedent nausea/vomiting, transient, responds rapidly to treatment	Amylase elevation with abdominal pain, nausea, occasional vomiting, responsive to treatment	Severe, persistent abdominal pain with pancreatitic enzyme elevation, incomplete or slow response to treatment	Complicated by shock, haemorrhage (acute circulatory failure)
G10. Proctitis	Perianal pruritus, haemorrhoids (new onset), transient, or interOpen-label treatedent, relieved by OTC meds	Tenesmus or ulcerations, anal fissure, responsive to treatment, minimal interference with function	Unresponsive to treatment, marked interference with function	Mucosal necrosis with haemorrhage, infection, surgery required
H. Musculoskeletal				
H1. Avascular necrosis	Asymptomatic MRI changes, non-progressive	MRI changes and symptoms responsive to rest and analgesia	MRI changes, symptoms requiring surgical intervention	Wheelchair bound; surgical repair not possible
H2. Arthralgia	InterOpen-label treatedent transient symptoms, no meds or relieved by OTC meds	Persistent or recurrent symptoms, resolve with meds, little effect on function	Severe symptoms despite meds impairs function	Debilitating, hospitalisation required for treatment
H3. Leg cramps	Transient, interOpen-label treatedent, does not interfere with function	Recurrent symptoms, minimally interferes with function or sleep, responds to meds	Persistent, prolonged interference with function or sleep, partial or no response to meds	NA

Rheumatology Common Toxicity Criteria v.2.0 **Based on Woodworth TG, et al. Standardizing assessment of adverse effects in rheumatology clinical trials II. Status of OMERACT Drug Safety Working Group May 2006: OMERACT 8. Standardizing Assessment and Reporting of Adverse Effects in Rheumatology Clinical Trials: Enabling Description of Comparative Safety Profiles for Antirheumatic Therapies**				
H. Musculoskeletal (Continued)				
H4. Myalgia	Occasional; does not interfere with function	Frequent, requires meds (non-narcotic); minor effects on function	Major change in function/lifestyle, narcotic pain meds	Debilitating, profound weakness, requires wheelchair, unresponsive to meds
I. Neuropsychiatric				
I1. Anxiety or Depression (mood alteration)	Symptomatic, does not interfere with function; no meds	Frequent symptoms, responds to meds; interferes with ADL at times	Persistent, prolonged symptoms, partial or no response to meds, limits daily function	Suicidal ideation or danger to self
I2. Cerebrovascular ischaemia	NA	Single transient ischaemic event, responsive to treatment	Recurrent transient ischaemic events	Cerebrovascular vascular accident with permanent disability
I3. Cognitive disturbance	Subjective symptoms, transient, interOpen-label treatedent, not interfering with function	Objective symptoms, persisting, interferes with daily function occasionally	Persistent, or worsening objective symptoms; interferes with routine daily routine	Debilitating/disabling and permanent; toxic psychosis
I4. Depressed consciousness (somnolence)	Observed, transient, interOpen-label treatedent, not interfering with function	Somnolence or sedation, interfering with function	Persistent, progressive, obundation, stupor	Coma
I5. Inability to concentrate	Subjective symptoms, does not interfere with function	Objective findings, interferes with function	Persistent, prolonged objective findings or organic cause	NA
I6. Insomnia (in absence of pain)	Occasional difficulty sleeping, transient interOpen-label treatedent, not interfering with function	Recurrent difficulty sleeping; requires meds for relief; occasional interference with function	Persistent or worsening difficulty sleeping; severely interferes with routine daily function	NA
I7. Libido decreased	Decrease in interest	Loss of interest; influences relationship	Persistent, prolonged interfering with relationship	NA

Rheumatology Common Toxicity Criteria v.2.0 **Based on Woodworth TG, et al. Standardizing assessment of adverse effects in rheumatology clinical trials II. Status of OMERACT Drug Safety Working Group May 2006: OMERACT 8. Standardizing Assessment and Reporting of Adverse Effects in Rheumatology Clinical Trials: Enabling Description of Comparative Safety Profiles for Antirheumatic Therapies**				
I. Neuropsychiatric (Continued)				
I8. Peripheral motor neuropathy	Subjective or transient loss of deep tendon reflexes; function maintained	Objective weakness, persistent, no significant impairment of daily function	Objective weakness with substantial impairment of function	Paralysis
I9. Peripheral sensory neuropathy (sensory disturbance)	Subjective symptoms without objective findings, transient, not interfering with function	Objective sensory loss, persistent, not interfering with function	Prolonged sensory loss or paraethesias interfering with function	NA
I10. Seizure	NA	Recurrence of old seizures, controlled with adjustment of medication	Recurrence/exacerbation with partial response to medication	Recurrence not controlled, requiring hospitalization; new seizures
I11. Vertigo (dizziness)	Subjective symptoms, transient, interOpen-label treatedent, no treatment	Objective findings, recurrent, meds relieve, occasionally interfering with function	Persistent, prolonged, interfering with daily function; partial response to medication	Debilitating without response to medication, hospitalization
J. Pulmonary				
J1. Asthma	Occasional wheeze, no interference with activities	Wheezing, requires oral meds, occasional interference with function	Debilitating, requires nasal O_2	Requires ventilator assistance
J2. Cough	Transient, interOpen-label treatedent, occasional OTC meds relieve	Persistent, requires narcotic or other prescription meds for relief	Recurrent, persistent coughing spasms without consistent relief by meds, interferes with function	Interferes with oxygenation; debilitating
J3. Dyspnea	Subjective, transient, no interference with function	Symptomatic, interOpen-label treatedent or recurring, interferes with exertional activities	Symptomatic during daily routine activities, interferes with function, treatment with interOpen-label treatedent nasal O2 relieves	Symptomatic at rest, debilitating, requires constant nasal O_2

Rheumatology Common Toxicity Criteria v.2.0 **Based on Woodworth TG, et al. Standardizing assessment of adverse effects in rheumatology clinical trials II. Status of OMERACT Drug Safety Working Group May 2006: OMERACT 8. Standardizing Assessment and Reporting of Adverse Effects in Rheumatology Clinical Trials: Enabling Description of Comparative Safety Profiles for Antirheumatic Therapies**				
J. Pulmonary (Continued)				
J4. Pleuritic pain (pleurisy)	Transient, interOpen-label treatedent symptoms, no treatment or OTC meds relieve	Persistent symptoms, requires prescription meds for relief	Prolonged symptoms, interferes with function, requires frequent narcotic pain relief	Debilitation, requiring hospitalisation
J5. Pneumonitis (pulmonary infiltrates)	Asymptomatic radiographic changes, transient, no treatment required	Symptomatic, persistent, requiring corticosteroids	Symptomatic, requiring treatment including O_2	Debilitating, not reversible; or requiring assisted ventilation
J6. Pulmonary function decreased (FVC or carbon monoxide diffusion capacity – DLCO)	76% – 90% of pre-treatment value	51% – 75% of pre-treatment value	26% – 50% of pre-treatment value	≤ 25% of pre-treatment value
Laboratory Data				
K. Haematology				
K1. Hgb (g/dl) decrease from pre-treatment	1.0 – 1.4	1.5 – 2.0	2.1 – 2.9, or Hgb < 8.0, > 7.0	≥ 3.0; or Hgb < 7.0
K2. Leukopenia (total WBC) × 1000	3.0 – 3.9	2.0 – 2.9	1.0 – 1.9	< 1.0
K3. Neutropenia (× 1000)	1.5 – 1.9	1.0 – 1.4	0.5 – 0.9	< 0.5
K4. Lymphopenia (× 1000)	1.5 – 1.9	1.0 –1.4	0.5 – 0.9	< 0.5
K5. Platelets (× 1000)	75 – LLN	50 – 74.9	20 – 49.9; platelet transfusion required	< 20; recurrent platelet transfusions

Rheumatology Common Toxicity Criteria v.2.0 **Based on Woodworth TG, et al. Standardizing assessment of adverse effects in rheumatology clinical trials II. Status of OMERACT Drug Safety Working Group May 2006: OMERACT 8. Standardizing Assessment and Reporting of Adverse Effects in Rheumatology Clinical Trials: Enabling Description of Comparative Safety Profiles for Antirheumatic Therapies**				
L. Chemistry				
L1. Hypercalcaemia (mg/dl)	1.1 × ULN – 11.5	11.6 – 12.5	12.6 – 13.5; or symptoms present	> 13.5; or associated coma
L2. Hyperglycemia (mg/dl) Fasting	140 – 160	161 – 250	251 – 500	> 500, or associated with ketoacidosis
L3. Hyperkalaemia (mmol/l)***	5.5 – 5.9	6.0 – 6.4	6.5 – 7.0 or any ECG change	> 7.0 or any arrhythmia
L5. Hypocalcaemia (mg/dl)	0.9 × LLN – 7.8	7.7 – 7.0	6.9 – 6.5; or associated with symptoms	< 6.5 or occurrence of tetany
L6. Hypoglycemia (mg/dl)	55 – 64 (no symptoms)	40 – 54 (or symptoms present)	30 – 39 (symptoms impair function)	< 30 or coma
L7. Hyponatraemia (mmol/l)***	-	125 – 129	120 – 124	< 120
L8. Hypokalaemia (mg/dl)	-	3.0 – 3.4	2.5 – 2.9	< 2.5
L9. CPK (also if polymyositis-disease	1.2 – 1.9 × ULN	2.0 – 4.0 × ULN	4.0 × ULN with weakness but without life-threatening signs or symptoms	> 4.0 × ULN with signs or symptoms of rhabdomyolysis or life-threatening
L10. Serum uric acid	1.2 – 1.6 × ULN	1.7 – 2.9 × ULN	3.0 – 5.0 × ULN or gout	NA
L11. Creatinine (mg/dl)	1.1 – 1.3 × ULN	1.4* – 1.8 × ULN	1.9 – 3.0 × ULN	> 3.0 × ULN
L12. SGOT (AST)	1.2 – 1.5 × ULN	1.6 – 3.0 × ULN	3.1 – 8.0 × ULN	> 8.0 × ULN
L13. SGPT (ALT)	1.2 – 1.5 × ULN	1.6 – 3.0 × ULN	3.0 – 8.0 × ULN	> 8.0 × ULN
L14. Alkaline phosphatase	1.1 – 1.5** × ULN	1.6 – 3.0 × ULN	3.0 – 5.0 × ULN	> 5.0 × ULN

Rheumatology Common Toxicity Criteria v.2.0 **Based on Woodworth TG, et al. Standardizing assessment of adverse effects in rheumatology clinical trials II. Status of OMERACT Drug Safety Working Group May 2006: OMERACT 8. Standardizing Assessment and Reporting of Adverse Effects in Rheumatology Clinical Trials: Enabling Description of Comparative Safety Profiles for Antirheumatic Therapies**				
Laboratory Data (Continued)				
L. Chemistry (Continued)				
L15. T. bilirubin	1.1 – 1.4 × ULN	1.5 – 1.9 × ULN	2.0 – 3.0 × ULN	> 3.0 × ULN
L16. LDH	1.3 – 2.4 × ULN	2.5 – 5.0 × ULN	5.1 – 10 × ULN	> 10 × ULN
M. Urinalysis				
M1. Haematuria	Micro only	Gross, no clots	Clots, transfusion < 2 units	Transfusion required
M2. Proteinuria (per 24 h)	300 – 500 mg (tr/1+)	501 – 1999 mg (2+)	2 – 5.0 g (3+) nephrotic syndrome	5.0 g (4+) anasarca
M3. WBC in urine	NA	NA	Indicating acute interstitial nephritis	Associated with acute renal failure
M4. Uric acid crystals	Present without symptoms	NA	With stones or symptoms of stones (e.g., renal colic)	Causing renal outflow obstruction and hospitalization

* In L11, 1.5 – 1.8 × ULN is changed to 1.4 – 1.8 × ULN.

** In L14, 1.1 – 2.0 × ULN is changed to 1.1 – 1.5 × ULN.

*** In L3, L7 and L8, mg/dl is changed to mmol/l.

**** For CPK and Creatinine NCI CTC grading will be used. For CPK therefore the following gradings apply: Grade 1: > ULN – 2.5 × ULN; Grade 2: > 2.5 – 5.0 × ULN; Grade 3: > 5.0 – 10.0 × ULN; Grade 4: > 10.0 × ULN; For Creatinine the following gradings apply: Grade 1: > 1 – 1.5 × Baseline; > ULN – 1.5 × ULN; Grade 2: > 1.5 – 3.0 × Baseline; > 1.5 – 3.0 × ULN; Grade 3: > 3.0 baseline; > 3.0 – 6.0 × ULN; Grade 4: > 6.0 × ULN.

Appendix D. Statistical Analysis to Account for Impact of COVID-19 Pandemic

1.0 Overview

The COVID-19 pandemic is interfering with the conduct of many ongoing trials, with potential impacts on treatment duration and collection, analysis and the interpretation of clinical trial data. Some protocol-specified visits in the clinical trials may be impacted due to COVID-19 infection or logistical restrictions during the pandemic. For example, some scheduled visits may be missed due to self-quarantine or local government restrictions on travel; some visits may also be delayed or canceled due to healthcare resource constraints during the pandemic. Impacted visits due to COVID-19 will be recorded in the database.

This appendix describes the additional analyses and updates to existing analyses due to COVID-19 impact.

2.0 Patient Disposition

Patient disposition and reason for discontinuation will be summarized as described in Section 7.0. For subjects who prematurely discontinued study drug, the reasons for discontinuation will be summarized with two additional categories capturing discontinuation due to COVID-19 infection or logistical restrictions related to the COVID-19 pandemic, as described in Section 7.0.

In addition, the number and percentage of subjects with scheduled study visits affected by COVID-19 pandemic will be summarized by treatment group sequences. The impact on study visits will be summarized by the following categories (as collected in CRF) by visit:

- Missed visit
- Virtual visit
- In person, partial assessments done

3.0 Long Term Efficacy Analysis

Missing data could occur due to various reasons, including missing visits/assessments, early withdrawal from the study, or missing due to COVID-19 infection or logistic restriction. The probability of having missed visits and missing data due to COVID-19 infection or logistical restrictions related to the COVID-19 pandemic can be reasonably assumed to be unrelated to the unobserved values. Therefore, for the purpose of statistical analysis, it is reasonable to assume that these missing data are missing at random (MAR) and the statistical models that require MAR assumption are appropriate.

As described in Section 9.2, longitudinal analysis models MMRM and GLMM will be used for long term efficacy analysis, and will be maintained for long term efficacy analysis in the presence of missing data due to COVID-19.

4.0 Safety Analysis

In listings of adverse events and deaths, a flag indicating whether the event or death was related to COVID-19 infection will be presented. A listing of COVID-19 related adverse events may be provided.

Statistical Analysis Plan for Study M13-538

Phase 2 Study, Multicenter, Open-Label Extension (OLE) Study in Rheumatoid Arthritis Subjects Who Have Completed a Preceding Phase 2 Randomized Controlled Trial (RCT) with Upadacitinib (ABT-494)

Prevnar 13® Vaccine Sub-Study

Date: 06 March 2020

Version 2.0

Table of Contents

List of Tables

List of Figures

1.0 Introduction

This Statistical Analysis Plan (SAP) describes the statistical analyses for the Prevnar 13® vaccine sub-study of upadacitinib Study M13-538 titled "Phase 2 Study, Multicenter, Open-Label Extension (OLE) Study in Rheumatoid Arthritis Subjects who have Completed a Preceding Phase 2 Randomized Controlled Trial (RCT) with Upadacitinib (ABT-494)."

The SAP will not be updated in case of administrative changes or amendments to the protocol unless the changes impact the analysis.

Unless noted otherwise, all analyses will be performed using SAS Version 9.4 (SAS Institute Inc., Cary, NC 27513) or later under the UNIX operating system.

2.0 Study Design and Objectives

2.1 Objectives and Hypotheses

The objective of this sub-study is to assess the impact of upadacitinib treatment (15 mg QD and 30 mg QD) with background MTX on immunological responses to Prevnar 13® in RA patients.

2.2 Study Design Overview

This sub-study will evaluate humoral immune response following Prevnar 13® administration in RA subjects enrolled in the main Study M13-538 treated with either upadacitinib 15 mg QD or 30 mg QD on a stable dose of background MTX. Immune response will be measured at three timepoints (pre-vaccination, 4 weeks post-vaccination, and 12 weeks post-vaccination). Enrollment will continue until approximately 150 subjects have been recruited to participate.

Subjects must be on a stable dose of:

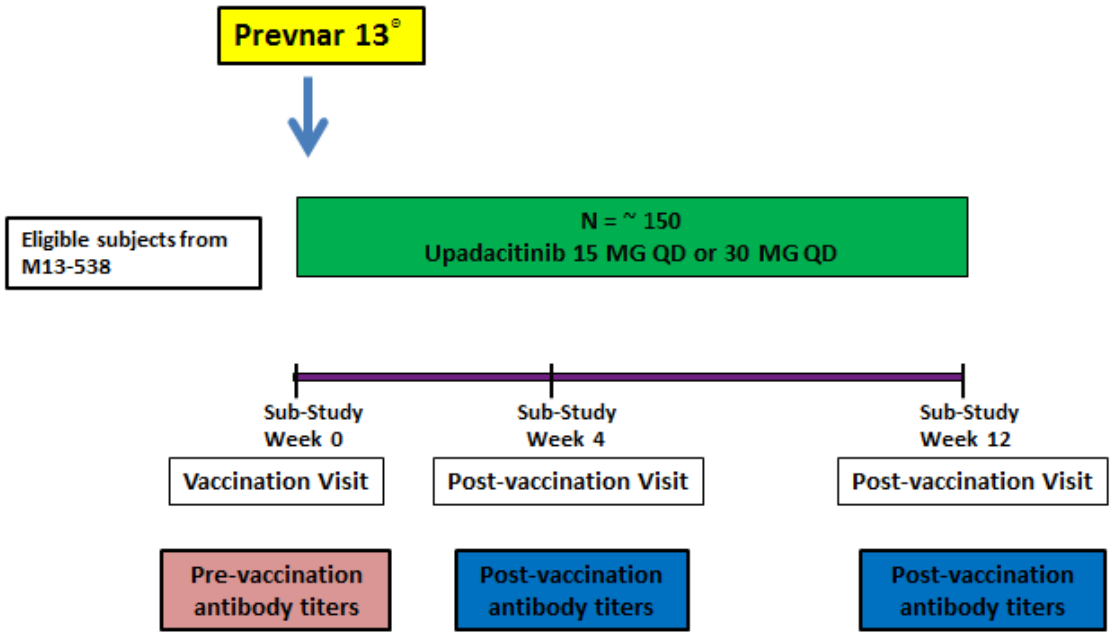
- upadacitinib (15 mg QD or 30 mg QD as assigned in the main study) for a minimum of 4 weeks prior to the Vaccination visit and for at least 4 weeks afterwards.
- background methotrexate for a minimum of 4 weeks prior to the Vaccination visit and for at least 4 weeks afterwards.

If subject is on corticosteroids, they must remain on a stable dose of ≤ 10 mg/day of prednisone or equivalent corticosteroid therapy for at least 4 weeks after the Vaccination visit.

At the Vaccination visit, subjects who have provided a written informed consent for the sub-study will be assessed for eligibility. If eligible, blood samples for the measurement of pre-vaccination antibody titers will be collected, followed by administration of Prevnar 13® as per the local label prescribing requirements. The criteria and procedures of the main study must still be adhered to in addition to those for the sub-study.

The schematic of the sub-study is shown in Figure 1.

Figure 1. Study Schematic



2.3 Treatment Assignment and Blinding

This is a sub-study of the open-label extension Study M13-538. All eligible subjects will receive Prevnar 13®.

2.4 Sample Size Determination

A total sample size of approximately 150 subjects is planned for the sub-study, where approximately 100 subjects are expected to be on the upadacitinib 15 mg QD dose. Assuming a satisfactory humoral response rate of 60%, the planned sample size of approximately 100 subjects for upadacitinib 15 mg QD group will provide a margin of error of approximately 10% for this dose group when using a 95% confidence interval.

3.0 Endpoints

3.1 Primary Endpoint(s)

The primary endpoint is:

- Proportion of subjects with satisfactory humoral response to Prevnar 13® at the Week 4 visit of the sub-study (4 weeks post-vaccination).

Satisfactory humoral response to Prevnar 13® is defined as greater than or equal to (≥) 2 fold increase in antibody concentration from the vaccination baseline in at least 6 out of the 12 pneumococcal antigens (1, 3, 4, 5, 6B, 7F, 9V, 14, 18C, 19A, 19F and 23F).

3.2 Secondary Endpoint(s)

The secondary endpoints are:

- Proportion of subjects with satisfactory humoral response to Prevnar 13® at the Week 12 visit of the sub-study (12 weeks post-vaccination).
- Geometric Mean Fold Rise (GMFR) of anti-pneumococcal antibody levels to each of the 12 pneumococcal antigens above vaccination baseline values at Week 4 and Week 12 of the sub-study.

3.3 Other Efficacy Endpoint(s)

Other endpoints:

- Proportion of subjects with at least 2-fold increase in antibody concentration from the vaccination baseline in each of the 12 pneumococcal antigens at the Week 4 and Week 12 visit of the sub-study.

3.4 Safety Endpoint(s)

The safety endpoints will include Adverse events (AEs).

4.0 Analysis Populations

The Full Analysis Set (FAS) includes all subjects enrolled in the sub-study who received Prevnar 13® and received at least one dose of study drug (upadacitinib) after vaccination during the sub-study. The FAS will be used for all baseline analyses and analyses of humoral response.

The Safety Analysis Set includes all subjects who received Prevnar 13® in the sub-study.

5.0 Subject Disposition

A summary of subject accountability will be provided where the number of subjects in each of the following categories will be summarized (by upadacitinib treatment and overall):

- Subjects who received Prevnar 13® vaccination (i.e., Safety Analysis Set)
- Subjects in the FAS
- Subjects who completed the vaccine sub-study (defined as subjects who completed Week 12 visit of the sub-study)

For subjects who discontinued upadacitinib during the first 4 weeks of the vaccine sub-study, the reasons for discontinuation will be summarized.

6.0 Study Drug Duration

Not applicable.

7.0 Analysis Conventions

Definition of Baseline

In this sub-study, a subject's baseline is defined as the last non-missing observation on or before the date of receiving Prevnar 13® vaccination.

Definition of Rx Days (Days Relative to the Date of Vaccination)

Rx Days are defined as the number of days between the day of vaccination and the specific time point. Rx days are negative values when the time point of interest is prior to the vaccination. Rx days are positive values when the time point of interest is on or after the date of vaccination. The day of vaccination is defined as Rx Day 1, while the day prior to vaccination is defined as Rx Day –1 (there is no Rx Day 0). Rx days are used to map actual study visits to the protocol-specified study visits.

Definition of Analysis Windows

Table 1 displays the visit window and the target/nominal day for each study visit.

Table 1. Visit Windows for Satisfactory Humoral Response, and anti-pneumococcal antibody levels

Window Label	Nominal Day	Interval
Baseline	1	≤ 1
Week 4	29	[15, 57]
Week 12	85	[58, 112]

8.0 Demographics, Baseline Characteristics, Medical History, and Concomitant Medications

Demographics and baseline disease characteristics will be summarized for FAS of the sub-study overall and by upadacitinib dose groups. Demographics and baseline disease characteristics will be defined relative to the baseline of the vaccine sub-study.

Categorical variables will be summarized with the number and percentage of subjects; percentages will be calculated based on the number of non-missing observations. Continuous variables will be summarized with descriptive statistics (number of non-missing observations, mean and standard deviation, median, minimum and maximum).

8.1 Demographics and Baseline Characteristics

Continuous demographic variables include age, weight and body mass index (BMI). Categorical demographic variables include sex, race, ethnicity, age (< 40, ≥ 40 – < 65 years, ≥ 65), BMI (< 18.5, ≥ 18.5 - < 25, ≥ 25 - < 30, ≥ 30).

Disease characteristics include the following:

- Duration of any upadacitinib treatment up to the sub-study baseline
- Duration of RA in years
- Concomitant csDMARD at baseline (MTX alone, MTX and other csDMARD, csDMARD other than MTX, none)
- MTX dose at baseline
- Oral steroid use at baseline
- Oral steroid dose at baseline
- CDAI
- CDAI (≤ 2.8, > 2.8 – ≤ 10, > 10)
- DAS28 (CRP)

8.2 Medical History

Not Applicable. This is a short-term sub-study, and any medical history will be summarized for the main study.

8.3 Concomitant Medications

Not Applicable. This is a short-term sub-study, and any concomitant medications will be summarized for the main study.

9.0 Efficacy Analyses

9.1 General Considerations

All efficacy analyses will be conducted in the FAS Population. Subjects will be classified into upadacitinib 15 mg QD dose group or upadacitinib 30 mg QD dose group based on their upadacitinib dose at the vaccination visit.

9.2 Handling of Missing Data

All efficacy analyses will be based on observed cases, which will not impute values for missing evaluations, and thus a subject who does not have an evaluation on a scheduled visit will be excluded from the analysis for that visit.

9.3 Primary Efficacy Endpoint(s) and Analyses

9.3.1 Primary Efficacy Endpoint(s)

The primary endpoint is:

- Proportion of subjects with satisfactory humoral response to Prevnar 13® at the Week 4 visit of the sub-study (4 weeks post-vaccination).

Satisfactory humoral response to Prevnar 13® is defined as greater than or equal to (≥) 2-fold increase in antibody concentration from the vaccination baseline in at least 6 out of the 12 pneumococcal antigens (1, 3, 4, 5, 6B, 7F, 9V, 14, 18C, 19A, 19F and 23F).

9.3.2 Handling of Missing Data for the Primary Efficacy Endpoint(s)

All efficacy analyses will be based on observed cases.

9.3.3 Primary Efficacy Analysis

The proportion of subjects achieving Satisfactory Humoral Response, defined as greater than or equal to (≥) 2-fold increase in antibody concentration from the vaccination

baseline in at least 6 out of the 12 pneumococcal antigens (1, 3, 4, 5, 6B, 7F, 9V, 14, 18C, 19A, 19F and 23F), will be summarized by count and percentages for each upadacitinib dose group at 4 weeks post-vaccination and 12 weeks post-vaccination. 95% confidence intervals for each upadacitinib dose group will be provided.

9.3.4 Additional Analyses of the Primary Efficacy Endpoint(s)

None.

9.4 Secondary Efficacy Analyses

9.4.1 Key Secondary Efficacy Analyses

At 4 weeks post-vaccination and 12 weeks post-vaccination, the geometric mean fold rise (GMFR) and corresponding 95% confidence intervals of anti-pneumococcal antibodies as compared with pre-vaccination will be summarized for each upadacitinib dose group. Calculations will be done using logarithmically transformed assay results, and then back-transformed to the original scale. Lower limit of quantification of 0.005 will be assigned to the zero concentration values.

9.4.2 Supportive Secondary Efficacy Analyses

None.

9.5 Additional Efficacy Analyses

The proportion of subjects achieving at least 2-fold increase in antibody concentration from the vaccination baseline in each of the 12 pneumococcal antigens will be summarized by count and percentages for each upadacitinib dose group at 4 weeks post-vaccination and 12 weeks post-vaccination. 95% confidence intervals for each upadacitinib dose group will be provided.

9.6 Efficacy Subgroup Analyses

To evaluate the consistency of the efficacy over different groups, the primary efficacy endpoint will be analyzed in the following subgroups:

- Age group (< 40, ≥ 40 – < 65 years, ≥ 65)
- Oral steroid use at baseline (Yes/No)
- Concomitant csDMARD at baseline (MTX alone, MTX and other csDMARD, csDMARD other than MTX, none)
- CDAI (≤ 2.8, > 2.8 – ≤ 10, > 10)

10.0 Safety Analyses

10.1 General Considerations

Safety data will be summarized for the Safety Analysis Set. Safety summaries will be presented by treatment group, in which treatment group is based on subjects' upadacitinib dose at the vaccination visit, including a total group for all subjects.

10.2 Adverse Events

Adverse events (AEs) will be summarized and presented using primary MedDRA System Organ Classes (SOCs) and preferred terms (PTs) according to the version of the MedDRA coding dictionary used for the study at the time of database lock. The actual version of the MedDRA coding dictionary used will be noted in the AE tables and in the clinical study report. Any AEs with onset after the vaccination and no more than 30 days post-vaccination will be summarized, overall and by upadacitinib dose groups. The SOCs will be presented in alphabetical order, and the PTs will be presented in alphabetical order within each SOC. Specific adverse events will be counted once for each subject for calculating percentages, unless stated otherwise.

10.2.1 Treatment-Emergent Adverse Events

No treatment-emergent AEs are defined in this sub-study. All AEs will be summarized.

10.2.2 Adverse Event Overview

An overview of AEs will be presented consisting of the number and percentage of subjects experiencing at least one event for each of the following AE categories:

- Any AE
- Any serious AE
- Any AE leading to discontinuation of upadacitinib
- Any severe AE
- Any AE with reasonable possibility of being related to upadacitinib
- Any AE leading to death
- All deaths

In the AE overview summary, any event with an unknown severity will be considered as severe and any AE with an unknown relationship will be considered as having "reasonable possibility" of being related to upadacitinib.

10.2.3 Adverse Events by SOC and/or PT

Adverse events will be summarized by SOC and PT; by maximum relationship to upadacitinib as assessed by the investigator (e.g., reasonable possibility or no reasonable possibility) and SOC and PT; by maximum severity and SOC and PT; and by subject number and SOC and PT. Severe AEs will be summarized by SOC and PT. Specific adverse events will be counted once for each subject for calculating percentages, unless stated otherwise.

In the summary of AEs by maximum severity and by maximum relationship, if the same adverse event occurs multiple times within a subject, the highest severity and level of relationship to upadacitinib will be reported. 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. 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 "reasonable possibility." In this case, the subject will be counted under the "reasonable possibility" category.

Adverse events will also be summarized by PT and sorted by decreasing frequency for the total group.

In addition, to assess potentially vaccine-related adverse events, events with onset within 7 days post-vaccination will be summarized by SOC and PT and by maximum severity and SOC and PT.

10.2.4 Adverse Events per 100 Patient-Years of Exposure

Not applicable. This is a short-term study with 12-week duration. No adverse events per 100 Patient-Years of exposure will be summarized.

10.2.5 SAEs (Including Deaths) and Adverse Events Leading to Study Drug Discontinuation

SAEs (including deaths), AEs leading to study drug discontinuation and AEs leading to death will be summarized by SOC and PT and in listing format.

10.3 Analysis of Laboratory Data

Not applicable. The clinical laboratory tests defined in the protocol (e.g., hematology and clinical chemistry) will be summarized for the main study.

10.4 Analysis of Vital Signs

Not applicable, vital sign measurements of systolic and diastolic blood pressure, pulse rate, respiratory rate, body temperature and weight will be summarized for the main study.

10.5 Safety Subgroup Analyses

None.

10.6 Other Safety Analyses

Not applicable.

11.0 Other Analyses

Not applicable.

12.0 Interim Analyses

Not applicable.

12.1 Data Monitoring Committee

Not applicable.

13.0 Overall Type-I Error Control

No multiplicity adjustment is planned for any analysis in this vaccine sub-study.

14.0 Summary of Changes

14.1 Summary of Changes between the Previous Version and the Current Version

1. Addition of other efficacy endpoint and the analysis method for this endpoint in Section 3.3 and Section 9.5, respectively.
2. Defined the Safety Analysis Set in Section 4.0.
3. Summary of concomitant medication in Section 8.3 has been removed from analysis.
4. The period for AE summaries has been updated to up to 30 days post-vaccination in Section 10.0.
5. Clarified the summary of AEs including potentially vaccine-related adverse events in Section 10.2.3.

6. Laboratory tests in Section 10.3 and Vital signs in Section 10.4 have been removed from the safety analyses.

15.0 Version History

Table 2. SAP Version History Summary

Version	Date	Summary
1.0	05 Aug 2019	Original version
2.0	06 March 2020	Amendment 1