Statistical Analysis Plan for Study M20-466

A Randomized, Double-Blind, Placebo-Controlled Study to Evaluate the Safety and Efficacy of ABBV-154 in Subjects with Moderately to Severely Active Rheumatoid Arthritis with Inadequate Response to Biologic and/or Targeted Synthetic Disease-Modifying Anti-Rheumatic Drugs (b/tsDMARDs)

Date: 17 August 2022

Version 2.0

Table of Contents

List of Tables

List of Figures

List of Appendices

1.0 Introduction

This Statistical Analysis Plan (SAP) describes the statistical analyses for ABBV-154 Study M20-466: A Randomized, Double-Blind, Placebo-Controlled Study to Evaluate the Safety and Efficacy of ABBV-154 in Subjects with Moderately to Severely Active Rheumatoid Arthritis with Inadequate Response to Biologic and/or Targeted Synthetic Disease-Modifying Anti-Rheumatic Drugs (b/tsDMARDs).

Study M20-466 assesses the efficacy and safety of ABBV-154 in subjects with moderately to severely active rheumatoid arthritis with inadequate response to biologic and/or targeted synthetic disease-modifying anti-rheumatic drugs (b/tsDMARDs).

The analyses of pharmacokinetic endpoints and pharmacodynamic biomarker endpoints will not be covered in this SAP.

The SAP will not be updated in case of administrative changes or amendments to the protocol that do not 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, Hypotheses and Estimands

The primary objective is to assess the safety, tolerability, and efficacy of ABBV-154 vs. placebo on background methotrexate (MTX) for the treatment of subjects with moderately to severely active rheumatoid arthritis (RA) who have had an inadequate response to at least one prior b/tsDMARD.

The primary efficacy objective of the study is to demonstrate a higher rate of achieving 50% improvement as measured by American College of Rheumatology response criteria (ACR50) after 12 weeks of treatment with ABBV-154 when compared to placebo in the Intent-to-Treat (ITT) Population (further described in Section 4.0), which consists of all

randomized subjects who received at least one dose of study drug. The hypothesis corresponding to the primary efficacy objective and endpoint is that the proportion of subjects achieving ACR50 with ABBV-154 is greater than that with placebo at Week 12. The estimand for the primary endpoint is defined as the difference in the proportion of subjects achieving ACR50 at Week 12 in each of the ABBV-154 dose groups compared with the placebo group in the ITT Population.

The secondary efficacy objectives of the study are to demonstrate higher efficacy of treatment with ABBV-154 compared to placebo with respect to the secondary endpoints specified in Section 3.2 in the ITT Population. The estimands corresponding to the secondary efficacy objectives are as follow.

- For each of the categorical secondary endpoints, the estimands corresponding to the secondary efficacy objectives are defined as follows: Difference in the proportion of subjects achieving the endpoint in each of the ABBV-154 groups in comparison with the placebo group in the ITT Population.
- For each of the continuous secondary endpoints, the estimands corresponding to the secondary efficacy objectives are defined as follows: Difference in the mean change from baseline of the endpoint in each of the ABBV-154 dose groups in comparison with the placebo group in the ITT Population.

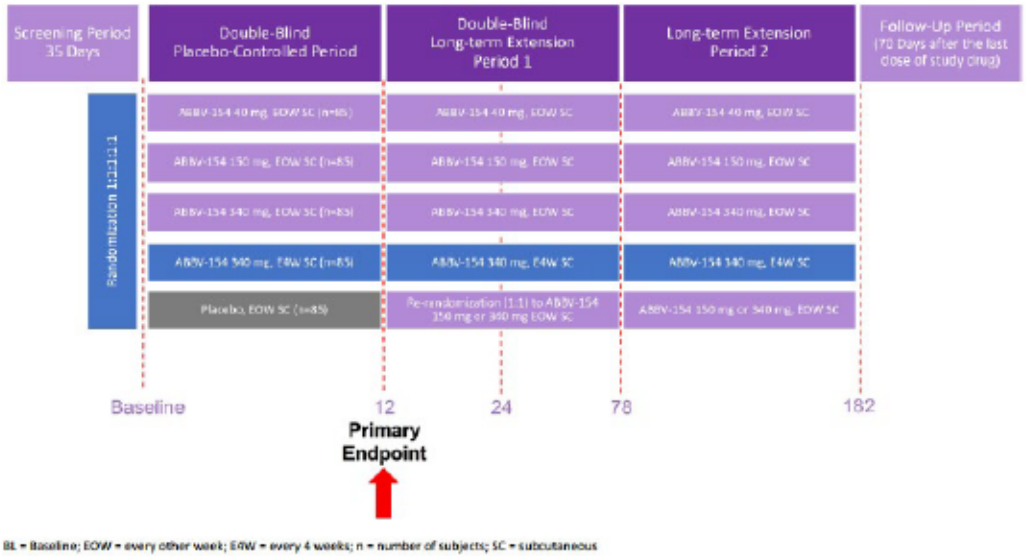
2.2 Study Design Overview

This is a randomized, double-blind, Phase 2b, dose-ranging study with a 12-week Placebo-Controlled Period, a 66-week Long-Term Extension (LTE) Period 1, and a 104-week LTE Period 2 in adults with moderately to severely active RA and on background MTX, who have had an inadequate response to b/tsDMARD treatment.

The Primary Analysis will be conducted after all ongoing subjects have completed Week 12 or withdrawn from the study. After the Primary Analysis, additional analyses may be conducted as needed. A final analysis will be conducted after all subjects have completed the LTE Period 2 and a safety follow-up visit 70 days after the last dose of

study drug, or have withdrawn from the study. The schematic of the study is shown in Figure 1.

Figure 1. Study Schematic



BL = Baseline; EOW = every other week; E4W = every 4 weeks; n = number of subjects; SC = subcutaneous

2.3 Treatment Assignment and Blinding

Subjects will remain on a stable dose of MTX until Week 24 (unless qualifying for rescue therapy as described in Protocol Section 5.4); thereafter MTX may be adjusted within the range of 7.5 – 25 mg/week. At Baseline, approximately 425 subjects will be randomized to 5 treatment groups in a 1:1:1:1:1 ratio to receive ABBV-154: 40 mg, 150 mg, 340 mg SC EOW; 340 mg SC E4W, or placebo. Randomization will be stratified by Baseline glucocorticoid (yes/no); number of prior failed b/tsDMARDs (1; 2 or more); and prior anti-TNF failure (yes/no) and if yes further stratification by prior adalimumab use (yes/no). At Week 12, subjects assigned to placebo will be re-randomized in a 1:1 ratio to receive ABBV-154 150 mg or 340 mg SC EOW; subjects from the other dose groups will continue with their respective dose and dosing regimen. Re-randomization will be stratified by Baseline glucocorticoid use (yes/no) and prior adalimumab use (yes/no). The

enrollment for subjects with prior adalimumab use is capped at approximately 30% of the total subjects enrolled. Systemic glucocorticoids (dose of ≤ 7.5 mg/d prednisone equivalent at Baseline) must be tapered off within 4 weeks of the start of study drug.

All AbbVie personnel with direct oversight of the conduct and management of the trial (except for AbbVie Drug Supply Management Team) will be blinded to each subject's treatment through Week 12. The AbbVie study team will be unblinded to treatment assignment for the Primary Analysis. This unblinding will take place after all subjects have completed the Week 12 visit or have withdrawn from the study prior to Week 12. To maintain integrity of the trial, the study site personnel and the subjects will remain blinded until all subjects have either completed the last visit of LTE Period 1 or have withdrawn from the study. Blinded dosing regimens will be continued until the previous criteria are met. To maintain the blind, the ABBV-154 and placebo pre-filled syringes provided for the study will be identical in appearance. The IRT will provide access to unblinded subject treatment information in the case of a medical emergency.

2.4 Sample Size Determination

The planned total sample size of 425 subjects (approximately 85 subjects per treatment group) provides at least 90% power to detect a 22% difference between ABBV-154 and placebo in ACR50 response rate at Week 12 (assuming a placebo ACR50 response rate of 13%), at a 2-sided significance level of 0.1 based on Chi-square test.

3.0 Endpoints

3.1 Primary Endpoint(s)

The primary endpoint is the achievement of ACR50 at Week 12.

3.2 Secondary Endpoint(s)

All secondary endpoints are at Week 12.

- Change in Disease Activity Score (DAS) 28 (CRP) from Baseline;

- Change in Clinical Disease Activity Index (CDAI) from Baseline;
- Achievement of ACR20;
- Achievement of ACR70;
- Achievement of low disease activity (LDA) defined by DAS28 (CRP) ≤ 3.2;
- Achievement of LDA defined by CDAI ≤ 10;
- Achievement of clinical remission (CR) defined by DAS28 (CRP) < 2.6;
- Achievement of CR defined by CDAI ≤ 2.8;
- Change in the Health Assessment Questionnaire Disability Index (HAQ-DI) from Baseline.

3.3 Other Efficacy Endpoint(s)

The primary endpoint and all secondary endpoints will be analyzed at all visits assessed other than Week 12. In addition, the following endpoints will be analyzed at all visits assessed.

- Change from Baseline in individual components of ACR
- Change from Baseline in simplified disease activity index (SDAI)
- Achievement of LDA defined as SDAI ≤ 11
- Achievement of CR defined as SDAI ≤ 3.3
- Achievement of Boolean remission based on a 28-joint count and defined as swollen joint count (SJC) ≤ 1, tender joint count (TJC) ≤ 1, high sensitivity C-reactive protein (hsCRP) ≤ 1 mg/dL, and Patient Global Assessment of Disease Activity ≤ 1 (on a 0-10 numeric rating scale [NRS])
- Change from Baseline in PROMIS Pain Interference
- Change from Baseline in FLARE-RA

3.4 Safety Endpoint(s)

Safety evaluations include adverse event (AE) monitoring, physical examinations, vital signs measurements, electrocardiograms (ECGs), and clinical laboratory testing (chemistry, hematology, and urinalysis).

Safety endpoints include the occurrence of the following:

1. Treatment-emergent adverse events, serious adverse events (SAEs), and AEs leading to discontinuation of the study drug
2. Potentially glucocorticoid related AEs
3. Potentially clinically important laboratory, vital signs, and electrocardiogram (ECG) parameters

3.5 Additional Endpoint(s)

Pharmacokinetic, immunogenicity, and biomarker research endpoints will be analyzed separately, and the corresponding analysis plan is not covered in this SAP.

4.0 Analysis Populations

The following population sets will be used for the analyses.

Intent-to-Treat (ITT) Population

The Intent-to-Treat (ITT) Population includes all randomized subjects who received at least one dose of study drug. The ITT Population will be used for all efficacy analyses. Subjects will be included in the analysis according to the treatment groups to which they are randomized.

Safety Population

The Safety Population in Placebo-Controlled Period (Safety_PC) includes all subjects randomized and received at least 1 dose of study drug in Placebo-Controlled Period. This population will be used for the safety analysis in the Placebo-Controlled Period.

The All ABBV-154 Treated Population (ALL_154) includes all subjects who receive at least 1 dose of ABBV-154 during the study. This population will be used to provide a comprehensive summary of safety.

For these safety populations (Safety_PC and ALL_154), subjects are assigned to treatment groups based on the treatments they receive during the majority of their drug exposure time in the analysis period for each subject in the respective safety population.

5.0 Subject Disposition

The total number of subjects who are screened, randomized, and treated will be summarized. Reasons for exclusion, including screen failure, will be summarized.

A summary of subject accountability will be provided where the number of subjects in each of the following categories will be summarized for each treatment group:

- Subjects randomized in the study;
- Subjects who take at least one dose of study drug;
- Subjects who complete protocol-specified treatment for Placebo-Controlled and LTE Periods 1 and 2, respectively;
- Subjects who prematurely discontinue study drug for Placebo-Controlled and LTE Periods 1 and 2, respectively;
- Subjects who complete study participation for Placebo-Controlled and LTE Periods, 1 and 2, respectively;
- Subjects who prematurely discontinue study participation for Placebo-Controlled and LTE Periods 1 and 2, respectively;
- Subjects in each analysis population, as applicable.

6.0 Study Drug Duration and Compliance

For the Placebo-Controlled period, LTE Period 1 and LTE Period 2, as well as study overall, duration of treatment will be summarized for each treatment group and for all ABBV-154 dose groups combined. Duration of treatment is defined for each subject as last dose date minus first dose date plus 14 days. Duration of treatment will be summarized using the number of subjects treated, mean, standard deviation, median, minimum and maximum. In addition, the number and percentage of subjects in each treatment duration interval (< 2 weeks [14 days], ≥ 2 weeks [14 days], ≥ 4 weeks [28 days], ≥ 8 weeks [56 days], ≥ 12 weeks [84 days], ≥ 24 weeks [168 days], ≥ 36 weeks [252 days], ≥ 52 weeks [364 days], ≥ 78 weeks [546 days], 104 weeks [728 days], 130 weeks [910 days], and 156 weeks [1,092 days]) will be summarized.

Treatment compliance will be summarized by treatment group and total ABBV-154 group for the safety populations. Treatment compliance is defined as the number of injections actually taken divided by the number of injections that should have been taken. All placebo injections will be included in the calculation for treatment compliance. Percent compliance will be summarized.

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

Demographics, baseline or disease characteristics, medical history, and prior and concomitant medications will be summarized for the ITT population overall and by treatment group. 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).

7.1 Demographics and Baseline Characteristics

Continuous demographic variables include age, weight, height, and body mass index (BMI). Categorical demographic variables include sex, ethnicity, race, age (< 40, 40 – 65, ≥ 65 years), weight (< 60, 60 – 100, ≥ 100 kg), BMI (< 25, ≥ 25 - < 30, ≥ 30 kg/m^2), region (Asia, North America, or rest of world), tobacco user (current, former, never, unknown), and alcohol user (current, former, never, unknown).

Disease characteristics that will be summarized as continuous and categorical variables are listed below:

RA Medical History and Characteristics

- Duration of RA Symptoms in years
- Duration of RA Diagnosis in years
- Duration of RA Diagnosis Categories (< 5, 5 – 10 or ≥ 10 years)

ACR and/or DAS Components at Baseline

- Tender joint count (TJC28) defined as the number of tender joints out of 28 assessed joints used for DAS28 calculation
- Swollen joint count (SJC28) defined as the number of swollen joints out of 28 assessed joints used for DAS28 calculation
- Tender joint count (TJC68) defined as the number of tender joints out of 68 assessed joints used for ACR20/50/70 calculation
- Swollen joint count (SJC66) defined as the number of swollen joints out of 66 assessed joints used for ACR20/50/70 calculation
- Physician's global assessment of disease activity (on a 0-10 NRS)
- Patient's assessment of pain (on a 0-10 NRS)
- Patient's global assessment of disease activity (on a 0-10 NRS)
- Health Assessment Questionnaire Disability Index of the (HAQ – DI) (range: 0 to 3)
- High sensitivity C-reactive protein (hsCRP) (mg/L)

Other Baseline RA Disease Characteristics

- DAS28 (CRP)
- DAS28 Categories:
  - DAS28 > 5.1 (High Disease Activity)
  - DAS28 3.2 to ≤ 5.1 (Moderate Disease Activity)
  - DAS28 < 3.2 (Low Disease Activity)
- Clinical Disease Activity Index (CDAI)
- CDAI categories:
  - CDAI > 22 (High Disease Activity)
  - CDAI >10 to ≤ 22 (Moderate Disease Activity)
  - CDAI ≤ 10 (Low Disease Activity)
- Simplified Disease Activity Index (SDAI)
- SDAI categories:
  - SDAI > 26 (High Disease Activity)
  - SDAI > 11 to ≤ 26 (Moderate Disease Activity)
  - SDAI ≤ 11 (Low Disease Activity)
- PROMIS Pain Interference Total Score and item level scores
- FLARE-RA (overall score, Joint and General subdomain scores, whether the subject experienced an RA flare in the last 3 months or since their last clinic visit [No/Yes, Once/Yes, several times], and whether the subject is currently flaring at baseline [No/Yes])
- Rheumatoid Factor (positive: ≥ 10 IU/mL, negative: < 10 IU/mL)
- Anti-citrullinated protein antibody (ACPA) (positive, negative)

Prior and Concomitant Treatment Use

- Baseline glucocorticoid use (yes/no)
- Glucocorticoid dose at Baseline
- Prior failed Biologic and/or Targeted Synthetic DMARDS (1; 2 or more)
- Prior anti-TNF failure (yes/no)

- Prior anti-TNF use (yes/no)
- Prior adalimumab use (yes/no)
- Prior non-anti-TNF biologic DMARD use (yes/no)
- Prior synthetic (conventional and targeted) DMARD use (yes/no)
  - Prior conventional synthetic DMARD use (yes/no)
  - Prior targeted synthetic DMARD use (yes/no)
- Duration of prior glucocorticoid use (month)
- Duration of prior glucocorticoid-free days before randomization (day)
- Prior cumulative glucocorticoid dose (mg)
- Prior cumulative glucocorticoid dose (mg) (within one year before randomization)

7.2 Medical History

Medical history data will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). The actual version of the MedDRA coding dictionary will be noted in the statistical tables and clinical study report. The number and percentage of subjects in each medical history category (by MedDRA system organ class and preferred term) will be summarized overall and by treatment group. The system organ class (SOC) will be presented in alphabetical order, and the preferred terms will be presented in alphabetical order within each SOC. Subjects reporting more than one condition/diagnosis will be counted only once in each row (SOC or preferred term).

7.3 Prior and Concomitant Medications

Prior and concomitant medications will be summarized by generic name. A prior medication is defined as any medication taken prior to the date of the first dose of study drug. A concomitant medication is defined as any medication that started prior to the date of the first dose of study drug and continued to be taken after the first dose of study drug or any medication that started on or after the date of the first dose of study drug, but not after the date of the last dose of study drug plus 70 days. The number and percentage of

subjects taking medications will be summarized by generic drug name based on the World Health Organization (WHO) Drug Dictionary for both prior and concomitant medications. The same summary will be provided specifically for DMARD medications, as well as for medications used for rescue therapy per protocol criterion.

8.0 Handling of Potential Intercurrent Events for the Primary and Key Secondary Endpoints

No intercurrent events are considered for the primary and key secondary endpoints.

9.0 Efficacy Analyses

9.1 General Considerations

All efficacy analyses will be conducted in the ITT Population. All tests will be 2-sided at an α level of 0.1.

The Primary Analysis will be performed after all ongoing subjects have completed the Placebo-Controlled Period, and the database has been locked. This will be the only and final analysis for the primary and secondary efficacy endpoints.

Unless otherwise specified, the categorical endpoints and continuous endpoints will be analyzed by Cochran-Mantel-Haenszel (CMH) and Mixed Model for Repeated Measures (MMRM), respectively, and the corresponding analyses are specified in Section 9.2.

Two of the baseline randomization stratification factors, prior anti-TNF failure and prior adalimumab use, will be combined into a single prior anti-TNF failure/adalimumab use factor (prior anti-TNF failure with adalimumab use; prior anti-TNF failure without adalimumab use; no prior anti-TNF failure) in the analysis methods. Other randomization stratification factors will be used without combination.

Unless otherwise specified, any subject who is randomized based on a wrong stratum will be analyzed according to the actual stratum the subject belongs to.

"Baseline" refers to the last non-missing observation before or on the day of the first administration of study drug or randomization if no study drug is given. Observations on the same day as the first administration of study drug will be considered for baseline, regardless of whether observation time (if available) was after the first administration of study drug time.

9.2 Handling of Intercurrent Event and Missing Data

9.2.1 Missing Data due to COVID-19 and Logistical Problems (Geo-political Restrictions)

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, as well as due to logistical problems (geo-political restrictions).

The COVID-19 pandemic is interfering with the conduct of many ongoing trials, with potential impacts on treatment duration and the 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. Impacted visits due to COVID-19 will be recorded in the database. 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. Sensitivity analyses will be performed to assess the impact of missing data and the robustness of the conclusion. The same logic is likely applicable to data missing due to logistical problems (geo-political restrictions) so that the MAR assumptions is also applicable.

9.2.2 Handling of missing data in Categorical Endpoints

- Non-Responder Imputation incorporating multiple imputation to handle COVID-19 and logistical problems (geo-political restrictions) (NRI-MI) will be the primary approach to handle missing data for categorical endpoints. The

NRI-MI will categorize any subject who does not have an evaluation during a pre-specified visit window (either due to missing assessment or due to early withdrawal from the study) as a non-responder for the visit. The only exceptions are: 1) when the subject is a responder both before and after the visit window, where the subject will be categorized as a responder for the visit, and 2) when the data is missing due to COVID-19 infection or logistical restriction, or logistical problems (geo-political restrictions), which will then be handled by Multiple Imputation. Subjects whose change/percent change from Baseline cannot be calculated because of a missing Baseline will be considered as a non-responder at all post-baseline visits in NRI-MI approach.

- Multiple Imputation (MI): a sensitivity analysis for the primary endpoint: Markov Chain Monte Carlo (MCMC) will be first applied to augment data into monotonic missing pattern and PROC MI will be used to generate 30 datasets using the regression method. The variables to be included in the imputation model for the Placebo-Controlled Period are treatment group, stratification factors at randomization [Baseline glucocorticoid (yes/no), number of prior failed b/tsDMARDs (1; 2 or more), and a combined prior anti-TNF failure/adalimumab use factor (prior anti-TNF failure with adalimumab use; prior anti-TNF failure without adalimumab use; no prior anti-TNF failure)], Baseline, and measurements at each visit in the Placebo-Controlled Period. The variables to be included in the imputation model for the LTE periods are treatment group, major stratum [baseline glucocorticoid use (yes/no) and prior adalimumab use (yes/no)], Baseline, and measurements at each visit in the LTE periods. The random seed for MCMC and the random seed for PROC MI are specified in Appendix C. The imputed post-baseline measurements will be rounded to the same precision as the observed data before the determination of responder status. Subjects will be characterized as responders or non-responders based on the MI imputed datasets. Using the Cochran-Mantel-Haenszel (CMH) model adjusted by stratification factors [Baseline glucocorticoid (yes/no), number of prior failed b/tsDMARDs (1; 2 or more), and a combined prior anti-TNF failure/adalimumab use factor (prior anti-TNF failure with adalimumab use; prior anti-TNF failure without adalimumab use; no prior anti-TNF failure)], the imputed endpoints will be analyzed using each of the 30 datasets. SAS PROC MIANALYZE will be used to generate the

final inferences of the risk difference between each ABBV-154 group and placebo.

9.2.3 Handling of missing data in Continuous Endpoints

- Mixed Model for Repeated Measures (MMRM): The repeated measures analysis will be conducted using a mixed model including observed measurements at all visits. For the Placebo-Controlled Period, the mixed model includes the categorical fixed effects of treatment, visit and treatment-by-visit interaction, stratification factors at randomization [Baseline glucocorticoid (yes/no), number of prior failed b/tsDMARDs (1; 2 or more), and a combined prior anti-TNF failure/adalimumab use factor (prior anti-TNF failure with adalimumab use; prior anti-TNF failure without adalimumab use; no prior anti-TNF failure)], and the continuous fixed covariates of baseline measurement. For LTE periods, the mixed model includes the categorical fixed effects of treatment, visit and treatment-by-visit interaction, major stratum [baseline glucocorticoid use (yes/no) and prior adalimumab use (yes/no)], and the continuous fixed covariates of baseline measurement. An unstructured variance-covariance matrix (UN) will be used. If the model cannot converge, an appropriate covariance structure matrix (e.g., autoregressive (1) or compound symmetry) will be used. Parameter estimation is based on the assumption of data being missing at random and using the method of restricted maximum likelihood (REML).

MMRM will be the primary approach to the analysis of continuous variables

9.2.4 Summary of Long-Term Efficacy

Long-term efficacy in the LTE periods will be summarized using the As Observed case approach by treatment sequence (ABBV-154 40 mg EOW, ABBV-154 150 mg EOW, ABBV-154 340 mg EOW, ABBV-154 340 mg E4W, ABBV-154 150 mg EOW re-randomized from Placebo, ABBV-154 340 mg EOW re-randomized from Placebo).

- As Observed (AO): The AO analysis will be used for the summaries of long-term efficacy, 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 AO analysis for that visit.

- NRI-MI and MMRM will be used to summarize selected categorical and continuous efficacy variables, respectively, for long-term efficacy.

9.3 Primary Efficacy Endpoint(s) and Analyses

9.3.1 Primary Efficacy Endpoint(s)

The primary endpoint is the achievement of ACR50 at Week 12.

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

ACR50 is defined as at least 50% improvement (compared to baseline values) in tender and swollen joint counts and at least 50% improvement in 3 of the remaining 5 core set measures (patient's assessment of pain, patient's global assessment of disease activity, physician's global assessment of disease activity, HAQ-DI and hsCRP).

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

1. $\geq$ 50% improvement from baseline in tender joint count (TJC68) and
2. $\geq$ 50% improvement from baseline in swollen joint count (SJC66) and
3. $\geq$ 50% improvement from baseline in at least 3 of the following 5:
   - patient's assessment of pain (on a 0-10 NRS)
   - patient's global assessment of disease activity (PtGA) (on a 0-10 NRS)
   - physician's global assessment of disease activity (PhGA) (on a 0-10 NRS)
   - HAQ-DI
   - hsCRP

9.3.2 Analysis of Primary Efficacy Endpoint(s)

The statistical null hypothesis corresponding to the primary endpoint is that there is no difference between an ABBV-154 group and the placebo group response rate in the proportion of subjects achieving ACR50 at Week 12.

Analysis of the primary endpoint will be conducted in the ITT Population based on treatment as randomized using NRI-MI. Comparison of the primary endpoint will be made between each ABBV-154 group and the placebo group using the CMH test adjusting for Baseline glucocorticoid (yes/no), number of prior failed b/tsDMARDs (1; 2 or more), and a combined prior anti-TNF failure/adalimumab use factor (prior anti-TNF failure with adalimumab use; prior anti-TNF failure without adalimumab use; no prior anti-TNF failure)].

The attributes of the estimands corresponding to the primary efficacy endpoint are summarized in Table 1.

Table 1. Summary of the Estimand Attributes of the Primary Efficacy Endpoint(s)

	Attributes of the Estimand				
Estimand Label	**Treatment**	**Endpoint**	**Population**	**Handling of Intercurrent Events (IE)**	**Statistical Summary**
Primary: Composite Estimand	ABBV-154 vs. Placebo	Achievement of ACR50 at Week 12	All randomized subjects who received at least one dose of study drug (ITT Population)	IE: none considered	Difference between ABBV-154 and placebo in the proportion of subjects achieving ACR50 at Week 12

9.3.3 Sensitivity and Supplementary Analyses of the Primary Efficacy Endpoint(s)

A sensitivity analysis of the primary analysis of the primary efficacy endpoint will be conducted using the MI method for handling missing data.

An additional analysis will be conducted where the baseline definition for patient's assessment of pain, PtGA, PhGA, and HAQ-DI is updated to where values collected up to 7 days after the day of first dose may be considered for the baseline value.

A separate analysis may be performed in the ITT Population excluding subjects who have major protocol deviations in the Placebo-Controlled Period. The subject exclusion criteria for this analysis will be finalized before unblinding data for the Primary Analysis.

9.4 Secondary Efficacy Endpoints and Analyses

9.4.1 Secondary Efficacy Endpoints

The continuous secondary endpoints are, at Week 12,

- Change in DAS 28 (CRP) from Baseline: DAS28 (CRP) is a composite index to assess disease activity in RA patients using hsCRP measurement. The DAS28 (CRP) provides a score between 0.96 and 10, indicating how active the rheumatoid arthritis is at the time of measurement.
- DAS28 (CRP) is calculated based on Tender Joint Count, Swollen Joint Count, Patient's Global Assessment of Disease Activity (PtGA) (on a 0-10 NRS), and hsCRP (in mg/L).

  $DAS28\ (CRP) = 0.56 \times \sqrt{}(TJC28^{*}) + 0.28 \times \sqrt{}(SJC28^{**}) + 0.36 \times \ln(hsCRP^{\&} + 1) + 0.14 \times PtGA^{\gg} + 0.96$

  * 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.

  » PtGA refers to the Patient's Global Assessment of Disease Activity.

  Where $\sqrt{}$ is square root and ln is natural log.

- Change in CDAI from Baseline: 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) (on a 0-10 NRS) and Physician's Global Assessment of Disease Activity (PhGA) (on a 0-10 NRS). It can be derived as follows:

  CDAI = TJC28 + SJC28 + PtGA+ PhGA.
- Change in the HAQ-DI from Baseline

The categorical secondary endpoints are, at Week 12,

- Achievement of ACR20 and Achievement of ACR70: The definition is similar to ACR50 with at least 20% and 70% improvement, respectively
- Achievement of LDA defined by DAS28 (CRP) ≤ 3.2
- Achievement of LDA defined by CDAI ≤ 10
- Achievement of clinical remission (CR) defined by DAS28 (CRP) < 2.6
- Achievement of CR defined by CDAI ≤ 2.8

9.4.2 Analyses of Secondary Efficacy Endpoints

The statistical null hypothesis corresponding to the secondary endpoint is that there is no difference between an ABBV-154 group and the placebo group at Week 12 in terms of response rate for a categorical endpoint and change from baseline for a continuous endpoint.

Analysis of the secondary endpoint will be conducted in the ITT Population based on treatment as randomized. Comparison of the secondary endpoint will be made between each ABBV-154 group and the placebo group using:

- For categorical variables: CMH test based on NRI-MI adjusted by Baseline glucocorticoid (yes/no), number of prior failed b/tsDMARDs (1; 2 or more), and a combined prior anti-TNF failure/adalimumab use factor (prior anti-TNF failure with adalimumab use; prior anti-TNF failure without adalimumab use; no prior anti-TNF failure).

- For continuous variables: MMRM mentioned in Section 9.2

The attributes of the estimands corresponding to the key secondary efficacy endpoints are summarized in Table 2.

Table 2. Summary of the Estimand Attributes of the Secondary Efficacy Endpoints

For binary endpoint:

	Attributes of the Estimand				
Estimand Label	**Treatment**	**Endpoint**	**Population**	**Handling of Intercurrent Events**	**Statistical Summary**
Binary Secondary: Composite Estimand	ABBV-154 vs. Placebo	Achievement of each binary secondary endpoint respectively	All randomized subjects who received at least one dose of study drug (ITT Population)	IE: none considered	Difference in proportion of subjects achieving each binary secondary endpoint

For continuous endpoint:

	Attributes of the Estimand				
Estimand Label	**Treatment**	**Endpoint**	**Population**	**Handling of Intercurrent Events**	**Statistical Summary**
Continuous Secondary: Composite Estimand	ABBV-154 vs Placebo	Change from Baseline in each respective continuous secondary endpoint	All randomized subjects who received at least one dose of study drug (ITT Population)	IE: none considered	Difference in the mean change from Baseline in each continuous secondary endpoint

9.4.3 Sensitivity and Supplementary Analyses for Secondary Efficacy Endpoints

Not applicable.

9.4.4 Supportive Secondary Efficacy Endpoints and Analyses

Not applicable.

9.5 Additional Efficacy Analyses

The primary and all secondary efficacy variables will be analyzed at all visits assessed other than Week 12, including in the LTE periods, using methods mentioned above. At time of the Primary Analysis, NRI will be used after Week 12 for all binary endpoints.

In addition, other efficacy endpoints described in Section 3.3 will be analyzed at all visits assessed using the same methods as secondary endpoints.

9.6 Dose-Response Modeling for ABBV-154

The dose-response relationship among ABBV-154 dose groups and the placebo group will be characterized for the primary endpoint using the Multiple Comparison Procedure - Modeling (MCP-Mod) method [Pinheiro 2006, Bretz 2005]. The summary level response rates based on the primary analysis approach above will be used, and ADDPLAN DF software will be used to perform the MCP-Mod analyses.

A set of 6 pre-specified standardized candidate dose-response models, as described in Table 3 will be utilized to examine the dose-response relationship. A statistically significant dose-response relationship will be declared if at least one model is identified by the MCP-Mod method to be statistically significant at two-sided $\alpha = 0.1$. The fitted dose-response curves will be presented graphically for all statistically significant models along with 95% confidence bands. The minimum effective dose (MED) will be identified for each statistically significant model based on the pre-specified clinical meaningful target of 22%. The weighted MED across all significant models will be calculated, with weight being the inverse of each candidate dose-response model AIC.

Table 3. Candidate Models

Model	$f(d,\theta)$ d = dose, θ = Model Parameters	$f^0(d,\theta)$ Standardized Model	Initial Value(s) for Parameter(s)*
Linear	$E_0 + \delta d$	d	NA
Exponential	$E_0 + E_1\left[exp\left(\frac{d}{\delta}\right) - 1\right]$	$exp\left(\frac{d}{\delta}\right) - 1$	$\delta = 680$
Logistic	$E_0 + \frac{E_{max}}{1 + exp\left(\frac{ED_{50} - d}{\delta}\right)}$	$\frac{1}{1 + exp\left(\frac{ED_{50} - d}{\delta}\right)}$	$ED_{50} = 150$, $\delta = 64.5$
EMax	$E_0 + \frac{E_{max} d}{ED_{50} + d}$	$\frac{d}{ED_{50} + d}$	$ED_{50} = 83.9$
sigEMax	$E_0 + \frac{E_{max} d^h}{{ED_{50}}^h + d^h}$	$\frac{d^h}{{ED_{50}}^h + d^h}$	$ED_{50} = 150$, $h = 3.6$
Quadratic	$E_0 + \beta_1 d + \beta_2 d^2$	$d + \frac{\beta_2}{\lvert\beta_1\rvert} d^2$	$\delta = -0.001$

* For Exponential model, the initial value was determined based on the assumption that ABBV-154 340 mg EOW will achieve 95% of the maximum efficacy of ABBV-154. For Logistic, EMax, sigEMax, and Quadratic model, the initial values were determined based on the assumption that ABBV-154 340 mg EOW and ABBV-154 140 mg EOW will achieve 95% and 50% of the maximum efficacy of ABBV-154, respectively.

Steps of MCP-Mod:

1. Choose a candidate set of models as in Table 3.
2. Compute the optimum contrast for each model.
3. Use contrast test to find all significant models while preserving family-wise error rate (FWER).
4. Use all significant models to make inference about the weighted target dose of interest.

9.7 Efficacy Subgroup Analyses

Subgroup analysis for the primary endpoint will be conducted by the subgroups specified below. The difference in the primary efficacy endpoint between the treatment groups in each subgroup will be assessed using CMH approach. Treatment difference and corresponding confidence interval will be used to summarize the result.

- Number of prior failed b/tsDMARDs (1; 2 or more)
- Prior anti-TNF failure (yes/no)
- Prior use of adalimumab (yes/no)
- Age (< median; ≥ median years)
- Sex (female; male)
- Race (white; non-white)
- Geographic region (North America, Asia, Rest of World)
- BMI (< 25, ≥ 25 - < 30, ≥ 30 kg/m^2)
- Baseline glucocorticoid use (yes/no)

If any subgroup has fewer than 10% of the total subjects, then the analysis for this subgroup will not be performed.

10.0 Safety Analyses

10.1 General Considerations

Safety data will be summarized for each safety population (Safety_PC and ALL_154). Safety summaries will be presented by treatment group, including a total group for all subjects on active study drug. For the safety analysis, subjects are assigned to a treatment group based on the treatment actually received, regardless of the treatment randomized.

"Baseline" refers to the last non-missing observation before or on the day of the first administration of study drug or randomization if no study drug is given. Observations on the same day as the first administration of study drug will be considered for baseline,

regardless if observation time (if available) was after the first administration of study drug time.

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. Specific adverse events will be counted once for each subject for calculating percentages, unless stated otherwise. In addition, if the same adverse event occurs multiple times within a subject, the highest severity and level of relationship to investigational product will be reported.

10.2.1 Treatment-Emergent Adverse Events

Treatment-emergent AEs are defined as any AE with the onset that is after the first dose of study drug and with an onset date no more than 5 half-lives of the study drug (i.e., 70 days). Events where the onset date is the same as the study drug start date are assumed to be treatment-emergent, unless the study drug start time and the AE start time are collected and the AE start time is prior to the study drug start time. All treatment-emergent AEs will be summarized overall, as well as by primary MedDRA SOC and Preferred Term. The SOCs will be presented in alphabetical order, and the PTs will be presented in alphabetical order within each SOC.

The number and percentage of subjects experiencing treatment-emergent AEs will be summarized.

In the ALL_154 population, treatment-emergent AEs are defined based on the first dose of ABBV-154.

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 treatment-emergent AE
- Any treatment-emergent AE related to study drug according to the investigator
- Any severe (CTCAE grade ≥ 3) treatment-emergent AE
- Any serious treatment-emergent AE
- Any treatment-emergent AE leading to discontinuation of study drug
- Any treatment-emergent AE leading to death
- Adverse Events of Special Interest (formally known as Safety Topics of Interest)
- All deaths
  - Deaths occurring ≤ 70 days after last dose of study drug
  - Deaths occurring > 70 days after last dose of study drug.

The overview of AEs per 100 patient-years of exposure will also be provided for the above categories.

10.2.3 Treatment-Emergent Adverse Events by SOC and/or PT

Treatment-emergent adverse events will be summarized by SOC and PT; by maximum relationship to study drug 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. Specific adverse events will be counted once for each subject for calculating percentages unless stated otherwise. In addition, if the same adverse event occurs multiple times within a subject, the highest severity and level of relationship to investigational product will be reported.

In addition, treatment-emergent adverse events will be summarized by PT and sorted by decreasing frequency for the total ABBV-154 group.

10.2.4 Treatment-Emergent Adverse Events per Patient-Years of Exposure

Exposure-adjusted AEs per 100 patient-years will be provided, where AEs per 100 patient-years of exposure are defined as the number of AEs divided by the total exposure in 100 patient-years. The study drug exposure is defined as the last dose date plus 5 × half-life (70 days), minus the first dose date.

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

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

In addition, the event rate per 100 patient-years of exposure will also be provided by SOC and PT for each treatment-emergent SAE and TEAE leading to study drug discontinuation.

10.2.6 Adverse Events of Special Interest

Adverse Events of Special Interest (AESI, formerly known as Safety topics of interest) will be summarized by SOC and PT and will be based on standardized or company MedDRA queries (SMQs or CMQs), or based on adjudication results. AESIs to be summarized will follow those indicated in the latest version of the Product Safety Statistical Analysis Plan (PSSAP) for ABBV-154 in the AESI/Safety Topics of Interest section. Detailed information about search criteria for the AESIs are also provided in a table in the same section.

Tabular listings of AESIs will be provided.

In addition, the event rate per 100 patient-years of exposure will also be provided by SOC and PT for each AESI.

10.3 Analysis of Laboratory Data

Data collected from central and local laboratories, including additional laboratory testing due to an SAE, will be used in all analyses, except for Baseline where SAE-related laboratory assessments on or before the first dose of study drug will be excluded. The clinical laboratory tests defined in the protocol operations manual (e.g., hematology and clinical chemistry) will be summarized.

Each laboratory variable will be summarized for all time points (starting with Baseline) with the number of non-missing observations, mean and standard deviation, median, minimum and maximum. Mean change from baseline to each applicable post-baseline visit will be summarized for selected laboratory variables, with the number of observations, baseline mean, and visit mean. The change from baseline mean, standard error, and 95% confidence interval will be presented for the mean change from baseline within each treatment group and difference between treatment groups (ABBV-154 vs. placebo) as applicable. This summary and change from baseline summary will be presented once by treatment sequence.

Changes in laboratory parameters will be tabulated using shift tables either by NCI CTC criteria or other criteria listed in Appendix B. A shift table from baseline either to the worse value (based on NCI CTC criteria) during treatment or to minimum and maximum value (based on normal range), will be created. A similar shift table will be provided to summarize shifts from baseline to the final post-baseline value.

Laboratory abnormalities meeting CTC criteria grade 3 and 4 will be summarized.

Laboratory abnormalities will be evaluated based on Potentially Clinically Important (PCI) criteria (based on NCI CTC criteria, except for those listed in Appendix B). For each laboratory PCI criterion, the number and percentage of subjects who have a laboratory value meeting the criteria will be summarized. Listings will be provided to summarize subject-level laboratory data for subjects meeting PCI criteria. In the evaluation of PCI laboratory values, the baseline value for the ALL_154 population will

be determined by the last non-missing measurement recorded on or before the date of the first dose of study drug in the corresponding ABBV-154 treatment group (which may be different than the first dose of study drug received in the study).

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

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

A listing of possible Hy's Law cases, defined as those who meet all of the following conditions simultaneously, will be provided: ALT > 3 × ULN or AST > 3 × ULN that is associated with an increase in bilirubin > 2 × ULN.

10.4 Analysis of Vital Signs

Vital sign measurements of systolic and diastolic blood pressure, respiratory rate, pulse rate, and body temperature will be summarized. Weight will be included in the summaries.

Each vital sign variable will be summarized for all time points (starting with Baseline) with the number of non-missing observations, mean and standard deviation, median, minimum and maximum. Mean change from baseline to each applicable post-baseline visit will be summarized for each vital sign variable, with the number of observations, baseline mean, and visit mean. The change from baseline mean, standard error, and 95% confidence interval will be presented for the mean change from baseline within each treatment group and difference between treatment groups (ABBV-154 vs. placebo) as applicable. This summary and change from baseline summary will be presented once by treatment sequence.

Vital sign variables will be evaluated based on PCI criteria. Pulse rate, body temperature, and respiratory rate will follow the criteria in Appendix B. All other vital signs will follow the criteria in the PSSAP. For each vital sign PCI criterion, the number and percentage of subjects who have a vital sign value meeting the criteria will be summarized. Listings will be provided to summarize subject-level vital sign data for subjects meeting PCI criteria. In the evaluation of PCI vital sign values for the ALL_154 population, 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 the corresponding ABBV-154 treatment group.

10.5 Safety Subgroup Analyses

The AE Overview and AE by SOC and PT summaries will be provided for the stratification variables:

- Baseline glucocorticoid dose (yes/no)
- Number of prior failed b/tsDMARDs (1; 2 or more)

10.6 Other Safety Analyses

ECG is collected at screening, Week 12, Week 24, Week 36, Week 78, Week 134, and Week 182. ECG findings will be summarized by treatment group for each parameter and visit.

Dual-energy X-ray absorptimometry (DXA) is collected at screening, Week 24, Week 78, Week 134, and Week 182. DXA measurements, including bone mineral density and corresponding T and Z scores, will be summarized by treatment group for each visit and also in the following groups:

- Female patients' post-menopausal status at baseline (Yes/No)
- Bone protective agent status at baseline (Yes/No)

Body composition measures of total fat mass (g), total lean mass (g), total body mass (g), fat (%), and lean (%) will also be summarized. Additionally, the ratio of total fat mass divided by total lean mass will be included.

DXA summaries will be presented once including all subjects and visits where the same machine as screening was used, and again including all subjects and visits (regardless of machine switching). The summaries will be provided for both centrally read and locally read results.

11.0 Other Analyses

Not applicable.

12.0 Interim Analyses

Not applicable.

12.1 Data Monitoring Committee

An external DMC will be established to safeguard the interest of trial subjects by assessing the safety of the interventions during the trial and well as for monitoring the integrity and interpretability of the trial.

A separate DMC charter describes the roles and responsibilities of the DMC members, frequency of data reviews, relevant data to be assessed, and general operations.

Since there are no efficacy analyses for early stopping, no alpha adjustment is needed.

An internal independent glucocorticoid adjudication committee will be established to adjudicate systemic glucocorticoid AEs. A separate charter describes the roles and responsibilities of the adjudication committee members, frequency of data reviews, and expectations for blinded communications.

13.0 Overall Type-I Error Control

Overall Type-I error control for multiple comparisons is not planned in this Phase 2b study.

Since there are no efficacy analyses for early stopping planned for the DMC review and IERC review, no alpha spending is needed due to analyses for the DMC review and the IERC review.

14.0 Version History

Table 4. SAP Version History Summary

Version	Date	Summary
1.0	04 August 2021	Original version

2.0	17 August 2022	• Updated to reflect protocol version 2 changes, including:
		  o Removal of intercurrent events
		  o Addition of LTE Period 2
		  o ITT population now requires at least 1 dose of study treatment
		• Changed categories for some baseline demographics/characteristics
		• Added summary of prior and concomitant DMARD, in addition to all medications
		• Added logistical problems (geo-political restrictions) to NRI-MI, in addition to COVID-19
		• Changed Observed Cases long-term efficacy to As Observed
		• Changed categories for efficacy and safety subgroup analyses for baseline glucocorticoid dose
		• Clarified treatment emergence definition for adverse events
		• Changed Safety Topics of Interest to Adverse Events of Special Interest
		• Clarified Potentially Clinical Important criteria for labs and vital signs
		• Corrected Hy's Law cases
		• Clarified included variables in vital signs analysis
		• Added details for DXA analysis

15.0 References

1. Rubin DB, Schenker N. Interval estimation from multiply-imputed data: a case study using agriculture industry codes. J Am Stat Assoc. 1987;81:366-74.
2. FDA Guidance on Conduct of Clinical Trials of Medical Products during COVID-19 Pandemic - Guidance for Industry, Investigators, and Institutional Review Boards. FDA. 2020.

3. Statistical Considerations for Clinical Trials During the COVID-19 Public Health Emergency - Guidance for Industry. FDA. 2020.
4. Points to consider on implications of Coronavirus disease (COVID-19) on methodological aspects of ongoing clinical trials. EMA. 2020.
5. Graham JW, Olchowski AE, Gilreath TD. How many imputations are really needed? Some practical clarifications of multiple imputation theory. Prev Sci. 2007;8(3):206-13.
6. Rubin DB, Schenker N. Interval estimation from multiply-imputed data: a case study using agriculture industry codes. J Am Stat Assoc. 1987;81:366-74.
7. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. Stat Med. 2011;30(4):377-99.
8. Pinheiro J, Bornkamp B, Bretz F. Design and analysis of dose-finding studies combining multiple comparisons and modeling procedures. J Biopharm Stat. 2006;16(5):639-56.
9. Bretz F, Pinheiro JC, Branson M. Combining multiple comparisons and modeling techniques in dose-response studies. Biometrics. 2005;61(3):738-48.

Appendix A. Protocol Deviations

The number and percentage of subjects who reported at least one of the following protocol deviation categories will be provided.

- Subject entered into the study even though s/he did not satisfy entry criteria.
- Subject developed withdrawal criteria during the study and was not withdrawn.
- Subject received wrong treatment or incorrect dose of study.
- Subject took prohibited concomitant medication.

Appendix B. Potentially Clinically Important Criteria for Safety Endpoints

The PCI criteria for selected laboratory variables are described in Table B1 instead of using NCI CTC.

Table B1. Potentially Clinically Important Criteria for Selected Laboratory Values

Laboratory Test	Category
HbA1c	< 6.5%
	≥ 6.5% to < 8%
	≥ 8%
HDL	< 1.03 mmol/L
	≥ 1.03 mmol/L
LDL	< 3.36 mmol/L
	≥ 3.36 to < 4.14 mmol/L
	≥ 4.14 mmol/L

The PCI criteria for selected vital sign variables (pulse rate, body temperature, and respiratory rate) are described in Table B2.

Table B2. Criteria for Potentially Clinically Important Vital Sign Values

Vital Sign	Category	Criteria for Potential Clinically Important Vital Signs
Pulse	Low	Value ≤ 50 bpm and decrease ≥ 15 bpm from Baseline
	High	Value ≥ 120 bpm and increase ≥ 15 bpm from Baseline
Body temperature	High	> 39.0 degrees C (102.3 degrees F)
Respiratory rate	Low	< 10 breaths/min
	High	> 24 breaths/min

Appendix C. Random Seeds

In case of non-convergence, the random seed will be updated by adding 10000 at each attempt until convergence of model happens.

A. Random Seeds for NRI-MI

Endpoints	Random Seed	
	MCMC procedure	PROC MI
TJC68	10001	20001
SJC66	10002	20002
TJC28	10003	20003
SJC28	10004	20004
Patient's Assessment of Pain	10005	20005
Patient's Global Assessment of Disease Activity	10006	20006
Physician's Global Assessment of Disease Activity	10007	20007
HAQ-DI	10008	20008
hsCRP	10009	20009

B. Random Seeds for MI

Endpoints	Random Seed	
	MCMC procedure	PROC MI
TJC68	40001	50001
SJC66	40002	50002
Patient's Assessment of Pain	40005	50005
Patient's Global Assessment of Disease Activity	40006	50006
Physician's Global Assessment of Disease Activity	40007	50007
HAQ-DI	40008	50008
hsCRP	40009	50009

final inferences of the risk difference between each ABBV-154 group and placebo.

9.2.3 Handling of missing data in Continuous Endpoints

- Mixed Model for Repeated Measures (MMRM): The repeated measures analysis will be conducted using a mixed model including observed measurements at all visits. For the Placebo-Controlled Period, the mixed model includes the categorical fixed effects of treatment, visit and treatment-by-visit interaction, stratification factors at randomization [Baseline glucocorticoid (yes/no), number of prior failed b/tsDMARDs (1; 2 or more), and a combined prior anti-TNF failure/adalimumab use factor (prior anti-TNF failure with adalimumab use; prior anti-TNF failure without adalimumab use; no prior anti-TNF failure)], and the continuous fixed covariates of baseline measurement. For LTE periods, the mixed model includes the categorical fixed effects of treatment, visit and treatment-by-visit interaction, major stratum [baseline glucocorticoid use (yes/no) and prior adalimumab use (yes/no)], and the continuous fixed covariates of baseline measurement. An unstructured variance-covariance matrix (UN) will be used. If the model cannot converge, an appropriate covariance structure matrix (e.g., autoregressive (1) or compound symmetry) will be used. Parameter estimation is based on the assumption of data being missing at random and using the method of restricted maximum likelihood (REML).

MMRM will be the primary approach to the analysis of continuous variables

9.2.4 Summary of Long-Term Efficacy

Long-term efficacy in the LTE periods will be summarized using the As Observed case approach by treatment sequence (ABBV-154 40 mg EOW, ABBV-154 150 mg EOW, ABBV-154 340 mg EOW, ABBV-154 340 mg E4W, ABBV-154 150 mg EOW re-randomized from Placebo, ABBV-154 340 mg EOW re-randomized from Placebo).

- As Observed (AO): The AO analysis will be used for the summaries of long-term efficacy, 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 AO analysis for that visit.

- NRI-MI and MMRM will be used to summarize selected categorical and continuous efficacy variables, respectively, for long-term efficacy.

9.3 Primary Efficacy Endpoint(s) and Analyses

9.3.1 Primary Efficacy Endpoint(s)

The primary endpoint is the achievement of ACR50 at Week 12.

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

ACR50 is defined as at least 50% improvement (compared to baseline values) in tender and swollen joint counts and at least 50% improvement in 3 of the remaining 5 core set measures (patient's assessment of pain, patient's global assessment of disease activity, physician's global assessment of disease activity, HAQ-DI and hsCRP).

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

1. $\geq$ 50% improvement from baseline in tender joint count (TJC68) and
2. $\geq$ 50% improvement from baseline in swollen joint count (SJC66) and
3. $\geq$ 50% improvement from baseline in at least 3 of the following 5:
   - patient's assessment of pain (on a 0-10 NRS)
   - patient's global assessment of disease activity (PtGA) (on a 0-10 NRS)
   - physician's global assessment of disease activity (PhGA) (on a 0-10 NRS)
   - HAQ-DI
   - hsCRP

9.3.2 Analysis of Primary Efficacy Endpoint(s)

The statistical null hypothesis corresponding to the primary endpoint is that there is no difference between an ABBV-154 group and the placebo group response rate in the proportion of subjects achieving ACR50 at Week 12.

Analysis of the primary endpoint will be conducted in the ITT Population based on treatment as randomized using NRI-MI. Comparison of the primary endpoint will be made between each ABBV-154 group and the placebo group using the CMH test adjusting for Baseline glucocorticoid (yes/no), number of prior failed b/tsDMARDs (1; 2 or more), and a combined prior anti-TNF failure/adalimumab use factor (prior anti-TNF failure with adalimumab use; prior anti-TNF failure without adalimumab use; no prior anti-TNF failure)].

The attributes of the estimands corresponding to the primary efficacy endpoint are summarized in Table 1.

Table 1. Summary of the Estimand Attributes of the Primary Efficacy Endpoint(s)

	Attributes of the Estimand				
Estimand Label	**Treatment**	**Endpoint**	**Population**	**Handling of Intercurrent Events (IE)**	**Statistical Summary**
Primary: Composite Estimand	ABBV-154 vs. Placebo	Achievement of ACR50 at Week 12	All randomized subjects who received at least one dose of study drug (ITT Population)	IE: none considered	Difference between ABBV-154 and placebo in the proportion of subjects achieving ACR50 at Week 12

9.3.3 Sensitivity and Supplementary Analyses of the Primary Efficacy Endpoint(s)

A sensitivity analysis of the primary analysis of the primary efficacy endpoint will be conducted using the MI method for handling missing data.

An additional analysis will be conducted where the baseline definition for patient's assessment of pain, PtGA, PhGA, and HAQ-DI is updated to where values collected up to 7 days after the day of first dose may be considered for the baseline value.

A separate analysis may be performed in the ITT Population excluding subjects who have major protocol deviations in the Placebo-Controlled Period. The subject exclusion criteria for this analysis will be finalized before unblinding data for the Primary Analysis.

9.4 Secondary Efficacy Endpoints and Analyses

9.4.1 Secondary Efficacy Endpoints

The continuous secondary endpoints are, at Week 12,

- Change in DAS 28 (CRP) from Baseline: DAS28 (CRP) is a composite index to assess disease activity in RA patients using hsCRP measurement. The DAS28 (CRP) provides a score between 0.96 and 10, indicating how active the rheumatoid arthritis is at the time of measurement.
- DAS28 (CRP) is calculated based on Tender Joint Count, Swollen Joint Count, Patient's Global Assessment of Disease Activity (PtGA) (on a 0-10 NRS), and hsCRP (in mg/L).

  $DAS28\ (CRP) = 0.56 \times \sqrt{}(TJC28^{*}) + 0.28 \times \sqrt{}(SJC28^{**}) + 0.36 \times \ln(hsCRP^{\&} + 1) + 0.14 \times PtGA^{\gg} + 0.96$

  * 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.

  » PtGA refers to the Patient's Global Assessment of Disease Activity.

  Where $\sqrt{}$ is square root and ln is natural log.

- Change in CDAI from Baseline: 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) (on a 0-10 NRS) and Physician's Global Assessment of Disease Activity (PhGA) (on a 0-10 NRS). It can be derived as follows:

  CDAI = TJC28 + SJC28 + PtGA+ PhGA.
- Change in the HAQ-DI from Baseline

The categorical secondary endpoints are, at Week 12,

- Achievement of ACR20 and Achievement of ACR70: The definition is similar to ACR50 with at least 20% and 70% improvement, respectively
- Achievement of LDA defined by DAS28 (CRP) ≤ 3.2
- Achievement of LDA defined by CDAI ≤ 10
- Achievement of clinical remission (CR) defined by DAS28 (CRP) < 2.6
- Achievement of CR defined by CDAI ≤ 2.8

9.4.2 Analyses of Secondary Efficacy Endpoints

The statistical null hypothesis corresponding to the secondary endpoint is that there is no difference between an ABBV-154 group and the placebo group at Week 12 in terms of response rate for a categorical endpoint and change from baseline for a continuous endpoint.

Analysis of the secondary endpoint will be conducted in the ITT Population based on treatment as randomized. Comparison of the secondary endpoint will be made between each ABBV-154 group and the placebo group using:

- For categorical variables: CMH test based on NRI-MI adjusted by Baseline glucocorticoid (yes/no), number of prior failed b/tsDMARDs (1; 2 or more), and a combined prior anti-TNF failure/adalimumab use factor (prior anti-TNF failure with adalimumab use; prior anti-TNF failure without adalimumab use; no prior anti-TNF failure).

- For continuous variables: MMRM mentioned in Section 9.2

The attributes of the estimands corresponding to the key secondary efficacy endpoints are summarized in Table 2.

Table 2. Summary of the Estimand Attributes of the Secondary Efficacy Endpoints

For binary endpoint:

	Attributes of the Estimand				
Estimand Label	**Treatment**	**Endpoint**	**Population**	**Handling of Intercurrent Events**	**Statistical Summary**
Binary Secondary: Composite Estimand	ABBV-154 vs. Placebo	Achievement of each binary secondary endpoint respectively	All randomized subjects who received at least one dose of study drug (ITT Population)	IE: none considered	Difference in proportion of subjects achieving each binary secondary endpoint

For continuous endpoint:

	Attributes of the Estimand				
Estimand Label	**Treatment**	**Endpoint**	**Population**	**Handling of Intercurrent Events**	**Statistical Summary**
Continuous Secondary: Composite Estimand	ABBV-154 vs Placebo	Change from Baseline in each respective continuous secondary endpoint	All randomized subjects who received at least one dose of study drug (ITT Population)	IE: none considered	Difference in the mean change from Baseline in each continuous secondary endpoint

9.4.3 Sensitivity and Supplementary Analyses for Secondary Efficacy Endpoints

Not applicable.

9.4.4 Supportive Secondary Efficacy Endpoints and Analyses

Not applicable.

9.5 Additional Efficacy Analyses

The primary and all secondary efficacy variables will be analyzed at all visits assessed other than Week 12, including in the LTE periods, using methods mentioned above. At time of the Primary Analysis, NRI will be used after Week 12 for all binary endpoints.

In addition, other efficacy endpoints described in Section 3.3 will be analyzed at all visits assessed using the same methods as secondary endpoints.

9.6 Dose-Response Modeling for ABBV-154

The dose-response relationship among ABBV-154 dose groups and the placebo group will be characterized for the primary endpoint using the Multiple Comparison Procedure - Modeling (MCP-Mod) method [Pinheiro 2006, Bretz 2005]. The summary level response rates based on the primary analysis approach above will be used, and ADDPLAN DF software will be used to perform the MCP-Mod analyses.

A set of 6 pre-specified standardized candidate dose-response models, as described in Table 3 will be utilized to examine the dose-response relationship. A statistically significant dose-response relationship will be declared if at least one model is identified by the MCP-Mod method to be statistically significant at two-sided $\alpha = 0.1$. The fitted dose-response curves will be presented graphically for all statistically significant models along with 95% confidence bands. The minimum effective dose (MED) will be identified for each statistically significant model based on the pre-specified clinical meaningful target of 22%. The weighted MED across all significant models will be calculated, with weight being the inverse of each candidate dose-response model AIC.

Table 3. Candidate Models

Model	$f(d,\theta)$ d = dose, θ = Model Parameters	$f^0(d,\theta)$ Standardized Model	Initial Value(s) for Parameter(s)*
Linear	$E_0 + \delta d$	d	NA
Exponential	$E_0 + E_1\left[exp\left(\frac{d}{\delta}\right) - 1\right]$	$exp\left(\frac{d}{\delta}\right) - 1$	$\delta = 680$
Logistic	$E_0 + \frac{E_{max}}{1 + exp\left(\frac{ED_{50} - d}{\delta}\right)}$	$\frac{1}{1 + exp\left(\frac{ED_{50} - d}{\delta}\right)}$	$ED_{50} = 150$, $\delta = 64.5$
EMax	$E_0 + \frac{E_{max} d}{ED_{50} + d}$	$\frac{d}{ED_{50} + d}$	$ED_{50} = 83.9$
sigEMax	$E_0 + \frac{E_{max} d^h}{{ED_{50}}^h + d^h}$	$\frac{d^h}{{ED_{50}}^h + d^h}$	$ED_{50} = 150$, $h = 3.6$
Quadratic	$E_0 + \beta_1 d + \beta_2 d^2$	$d + \frac{\beta_2}{\lvert\beta_1\rvert} d^2$	$\delta = -0.001$

* For Exponential model, the initial value was determined based on the assumption that ABBV-154 340 mg EOW will achieve 95% of the maximum efficacy of ABBV-154. For Logistic, EMax, sigEMax, and Quadratic model, the initial values were determined based on the assumption that ABBV-154 340 mg EOW and ABBV-154 140 mg EOW will achieve 95% and 50% of the maximum efficacy of ABBV-154, respectively.

Steps of MCP-Mod:

1. Choose a candidate set of models as in Table 3.
2. Compute the optimum contrast for each model.
3. Use contrast test to find all significant models while preserving family-wise error rate (FWER).
4. Use all significant models to make inference about the weighted target dose of interest.

9.7 Efficacy Subgroup Analyses

Subgroup analysis for the primary endpoint will be conducted by the subgroups specified below. The difference in the primary efficacy endpoint between the treatment groups in each subgroup will be assessed using CMH approach. Treatment difference and corresponding confidence interval will be used to summarize the result.

- Number of prior failed b/tsDMARDs (1; 2 or more)
- Prior anti-TNF failure (yes/no)
- Prior use of adalimumab (yes/no)
- Age (< median; ≥ median years)
- Sex (female; male)
- Race (white; non-white)
- Geographic region (North America, Asia, Rest of World)
- BMI (< 25, ≥ 25 - < 30, ≥ 30 kg/m^2)
- Baseline glucocorticoid use (yes/no)

If any subgroup has fewer than 10% of the total subjects, then the analysis for this subgroup will not be performed.

10.0 Safety Analyses

10.1 General Considerations

Safety data will be summarized for each safety population (Safety_PC and ALL_154). Safety summaries will be presented by treatment group, including a total group for all subjects on active study drug. For the safety analysis, subjects are assigned to a treatment group based on the treatment actually received, regardless of the treatment randomized.

"Baseline" refers to the last non-missing observation before or on the day of the first administration of study drug or randomization if no study drug is given. Observations on the same day as the first administration of study drug will be considered for baseline,

regardless if observation time (if available) was after the first administration of study drug time.

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. Specific adverse events will be counted once for each subject for calculating percentages, unless stated otherwise. In addition, if the same adverse event occurs multiple times within a subject, the highest severity and level of relationship to investigational product will be reported.

10.2.1 Treatment-Emergent Adverse Events

Treatment-emergent AEs are defined as any AE with the onset that is after the first dose of study drug and with an onset date no more than 5 half-lives of the study drug (i.e., 70 days). Events where the onset date is the same as the study drug start date are assumed to be treatment-emergent, unless the study drug start time and the AE start time are collected and the AE start time is prior to the study drug start time. All treatment-emergent AEs will be summarized overall, as well as by primary MedDRA SOC and Preferred Term. The SOCs will be presented in alphabetical order, and the PTs will be presented in alphabetical order within each SOC.

The number and percentage of subjects experiencing treatment-emergent AEs will be summarized.

In the ALL_154 population, treatment-emergent AEs are defined based on the first dose of ABBV-154.

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 treatment-emergent AE
- Any treatment-emergent AE related to study drug according to the investigator
- Any severe (CTCAE grade ≥ 3) treatment-emergent AE
- Any serious treatment-emergent AE
- Any treatment-emergent AE leading to discontinuation of study drug
- Any treatment-emergent AE leading to death
- Adverse Events of Special Interest (formally known as Safety Topics of Interest)
- All deaths
  - Deaths occurring ≤ 70 days after last dose of study drug
  - Deaths occurring > 70 days after last dose of study drug.

The overview of AEs per 100 patient-years of exposure will also be provided for the above categories.

10.2.3 Treatment-Emergent Adverse Events by SOC and/or PT

Treatment-emergent adverse events will be summarized by SOC and PT; by maximum relationship to study drug 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. Specific adverse events will be counted once for each subject for calculating percentages unless stated otherwise. In addition, if the same adverse event occurs multiple times within a subject, the highest severity and level of relationship to investigational product will be reported.

In addition, treatment-emergent adverse events will be summarized by PT and sorted by decreasing frequency for the total ABBV-154 group.

10.2.4 Treatment-Emergent Adverse Events per Patient-Years of Exposure

Exposure-adjusted AEs per 100 patient-years will be provided, where AEs per 100 patient-years of exposure are defined as the number of AEs divided by the total exposure in 100 patient-years. The study drug exposure is defined as the last dose date plus 5 × half-life (70 days), minus the first dose date.

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

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

In addition, the event rate per 100 patient-years of exposure will also be provided by SOC and PT for each treatment-emergent SAE and TEAE leading to study drug discontinuation.

10.2.6 Adverse Events of Special Interest

Adverse Events of Special Interest (AESI, formerly known as Safety topics of interest) will be summarized by SOC and PT and will be based on standardized or company MedDRA queries (SMQs or CMQs), or based on adjudication results. AESIs to be summarized will follow those indicated in the latest version of the Product Safety Statistical Analysis Plan (PSSAP) for ABBV-154 in the AESI/Safety Topics of Interest section. Detailed information about search criteria for the AESIs are also provided in a table in the same section.

Tabular listings of AESIs will be provided.

In addition, the event rate per 100 patient-years of exposure will also be provided by SOC and PT for each AESI.

10.3 Analysis of Laboratory Data

Data collected from central and local laboratories, including additional laboratory testing due to an SAE, will be used in all analyses, except for Baseline where SAE-related laboratory assessments on or before the first dose of study drug will be excluded. The clinical laboratory tests defined in the protocol operations manual (e.g., hematology and clinical chemistry) will be summarized.

Each laboratory variable will be summarized for all time points (starting with Baseline) with the number of non-missing observations, mean and standard deviation, median, minimum and maximum. Mean change from baseline to each applicable post-baseline visit will be summarized for selected laboratory variables, with the number of observations, baseline mean, and visit mean. The change from baseline mean, standard error, and 95% confidence interval will be presented for the mean change from baseline within each treatment group and difference between treatment groups (ABBV-154 vs. placebo) as applicable. This summary and change from baseline summary will be presented once by treatment sequence.

Changes in laboratory parameters will be tabulated using shift tables either by NCI CTC criteria or other criteria listed in Appendix B. A shift table from baseline either to the worse value (based on NCI CTC criteria) during treatment or to minimum and maximum value (based on normal range), will be created. A similar shift table will be provided to summarize shifts from baseline to the final post-baseline value.

Laboratory abnormalities meeting CTC criteria grade 3 and 4 will be summarized.

Laboratory abnormalities will be evaluated based on Potentially Clinically Important (PCI) criteria (based on NCI CTC criteria, except for those listed in Appendix B). For each laboratory PCI criterion, the number and percentage of subjects who have a laboratory value meeting the criteria will be summarized. Listings will be provided to summarize subject-level laboratory data for subjects meeting PCI criteria. In the evaluation of PCI laboratory values, the baseline value for the ALL_154 population will

be determined by the last non-missing measurement recorded on or before the date of the first dose of study drug in the corresponding ABBV-154 treatment group (which may be different than the first dose of study drug received in the study).

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

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

A listing of possible Hy's Law cases, defined as those who meet all of the following conditions simultaneously, will be provided: ALT > 3 × ULN or AST > 3 × ULN that is associated with an increase in bilirubin > 2 × ULN.

10.4 Analysis of Vital Signs

Vital sign measurements of systolic and diastolic blood pressure, respiratory rate, pulse rate, and body temperature will be summarized. Weight will be included in the summaries.

Each vital sign variable will be summarized for all time points (starting with Baseline) with the number of non-missing observations, mean and standard deviation, median, minimum and maximum. Mean change from baseline to each applicable post-baseline visit will be summarized for each vital sign variable, with the number of observations, baseline mean, and visit mean. The change from baseline mean, standard error, and 95% confidence interval will be presented for the mean change from baseline within each treatment group and difference between treatment groups (ABBV-154 vs. placebo) as applicable. This summary and change from baseline summary will be presented once by treatment sequence.

Vital sign variables will be evaluated based on PCI criteria. Pulse rate, body temperature, and respiratory rate will follow the criteria in Appendix B. All other vital signs will follow the criteria in the PSSAP. For each vital sign PCI criterion, the number and percentage of subjects who have a vital sign value meeting the criteria will be summarized. Listings will be provided to summarize subject-level vital sign data for subjects meeting PCI criteria. In the evaluation of PCI vital sign values for the ALL_154 population, 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 the corresponding ABBV-154 treatment group.

10.5 Safety Subgroup Analyses

The AE Overview and AE by SOC and PT summaries will be provided for the stratification variables:

- Baseline glucocorticoid dose (yes/no)
- Number of prior failed b/tsDMARDs (1; 2 or more)

10.6 Other Safety Analyses

ECG is collected at screening, Week 12, Week 24, Week 36, Week 78, Week 134, and Week 182. ECG findings will be summarized by treatment group for each parameter and visit.

Dual-energy X-ray absorptimometry (DXA) is collected at screening, Week 24, Week 78, Week 134, and Week 182. DXA measurements, including bone mineral density and corresponding T and Z scores, will be summarized by treatment group for each visit and also in the following groups:

- Female patients' post-menopausal status at baseline (Yes/No)
- Bone protective agent status at baseline (Yes/No)

Body composition measures of total fat mass (g), total lean mass (g), total body mass (g), fat (%), and lean (%) will also be summarized. Additionally, the ratio of total fat mass divided by total lean mass will be included.

DXA summaries will be presented once including all subjects and visits where the same machine as screening was used, and again including all subjects and visits (regardless of machine switching). The summaries will be provided for both centrally read and locally read results.

11.0 Other Analyses

Not applicable.

12.0 Interim Analyses

Not applicable.

12.1 Data Monitoring Committee

An external DMC will be established to safeguard the interest of trial subjects by assessing the safety of the interventions during the trial and well as for monitoring the integrity and interpretability of the trial.

A separate DMC charter describes the roles and responsibilities of the DMC members, frequency of data reviews, relevant data to be assessed, and general operations.

Since there are no efficacy analyses for early stopping, no alpha adjustment is needed.

An internal independent glucocorticoid adjudication committee will be established to adjudicate systemic glucocorticoid AEs. A separate charter describes the roles and responsibilities of the adjudication committee members, frequency of data reviews, and expectations for blinded communications.

13.0 Overall Type-I Error Control

Overall Type-I error control for multiple comparisons is not planned in this Phase 2b study.

Since there are no efficacy analyses for early stopping planned for the DMC review and IERC review, no alpha spending is needed due to analyses for the DMC review and the IERC review.

14.0 Version History

Table 4. SAP Version History Summary

Version	Date	Summary
1.0	04 August 2021	Original version

2.0	17 August 2022	• Updated to reflect protocol version 2 changes, including:
		  o Removal of intercurrent events
		  o Addition of LTE Period 2
		  o ITT population now requires at least 1 dose of study treatment
		• Changed categories for some baseline demographics/characteristics
		• Added summary of prior and concomitant DMARD, in addition to all medications
		• Added logistical problems (geo-political restrictions) to NRI-MI, in addition to COVID-19
		• Changed Observed Cases long-term efficacy to As Observed
		• Changed categories for efficacy and safety subgroup analyses for baseline glucocorticoid dose
		• Clarified treatment emergence definition for adverse events
		• Changed Safety Topics of Interest to Adverse Events of Special Interest
		• Clarified Potentially Clinical Important criteria for labs and vital signs
		• Corrected Hy's Law cases
		• Clarified included variables in vital signs analysis
		• Added details for DXA analysis

15.0 References

1. Rubin DB, Schenker N. Interval estimation from multiply-imputed data: a case study using agriculture industry codes. J Am Stat Assoc. 1987;81:366-74.
2. FDA Guidance on Conduct of Clinical Trials of Medical Products during COVID-19 Pandemic - Guidance for Industry, Investigators, and Institutional Review Boards. FDA. 2020.

3. Statistical Considerations for Clinical Trials During the COVID-19 Public Health Emergency - Guidance for Industry. FDA. 2020.
4. Points to consider on implications of Coronavirus disease (COVID-19) on methodological aspects of ongoing clinical trials. EMA. 2020.
5. Graham JW, Olchowski AE, Gilreath TD. How many imputations are really needed? Some practical clarifications of multiple imputation theory. Prev Sci. 2007;8(3):206-13.
6. Rubin DB, Schenker N. Interval estimation from multiply-imputed data: a case study using agriculture industry codes. J Am Stat Assoc. 1987;81:366-74.
7. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. Stat Med. 2011;30(4):377-99.
8. Pinheiro J, Bornkamp B, Bretz F. Design and analysis of dose-finding studies combining multiple comparisons and modeling procedures. J Biopharm Stat. 2006;16(5):639-56.
9. Bretz F, Pinheiro JC, Branson M. Combining multiple comparisons and modeling techniques in dose-response studies. Biometrics. 2005;61(3):738-48.

Appendix A. Protocol Deviations

The number and percentage of subjects who reported at least one of the following protocol deviation categories will be provided.

- Subject entered into the study even though s/he did not satisfy entry criteria.
- Subject developed withdrawal criteria during the study and was not withdrawn.
- Subject received wrong treatment or incorrect dose of study.
- Subject took prohibited concomitant medication.

Appendix B. Potentially Clinically Important Criteria for Safety Endpoints

The PCI criteria for selected laboratory variables are described in Table B1 instead of using NCI CTC.

Table B1. Potentially Clinically Important Criteria for Selected Laboratory Values

Laboratory Test	Category
HbA1c	< 6.5%
	≥ 6.5% to < 8%
	≥ 8%
HDL	< 1.03 mmol/L
	≥ 1.03 mmol/L
LDL	< 3.36 mmol/L
	≥ 3.36 to < 4.14 mmol/L
	≥ 4.14 mmol/L

The PCI criteria for selected vital sign variables (pulse rate, body temperature, and respiratory rate) are described in Table B2.

Table B2. Criteria for Potentially Clinically Important Vital Sign Values

Vital Sign	Category	Criteria for Potential Clinically Important Vital Signs
Pulse	Low	Value ≤ 50 bpm and decrease ≥ 15 bpm from Baseline
	High	Value ≥ 120 bpm and increase ≥ 15 bpm from Baseline
Body temperature	High	> 39.0 degrees C (102.3 degrees F)
Respiratory rate	Low	< 10 breaths/min
	High	> 24 breaths/min

Appendix C. Random Seeds

In case of non-convergence, the random seed will be updated by adding 10000 at each attempt until convergence of model happens.

A. Random Seeds for NRI-MI

Endpoints	Random Seed	
	MCMC procedure	PROC MI
TJC68	10001	20001
SJC66	10002	20002
TJC28	10003	20003
SJC28	10004	20004
Patient's Assessment of Pain	10005	20005
Patient's Global Assessment of Disease Activity	10006	20006
Physician's Global Assessment of Disease Activity	10007	20007
HAQ-DI	10008	20008
hsCRP	10009	20009

B. Random Seeds for MI

Endpoints	Random Seed	
	MCMC procedure	PROC MI
TJC68	40001	50001
SJC66	40002	50002
Patient's Assessment of Pain	40005	50005
Patient's Global Assessment of Disease Activity	40006	50006
Physician's Global Assessment of Disease Activity	40007	50007
HAQ-DI	40008	50008
hsCRP	40009	50009