Statistical Analysis Plan I4V-MC-JAJA

A Randomized, Active-Controlled, Parallel-Group, Phase 3b/4 Study of Baricitinib in Patients with Rheumatoid Arthritis

NCT03915964

Approval date: 21-Jul-2025

1. Statistical Analysis Plan: I4V-MC-JAJA (RA-BRIDGE): A Randomized, Active-Controlled, Parallel-Group, Phase 3b/4 Study of Baricitinib in Patients with Rheumatoid Arthritis

Baricitinib (LY3009104) Rheumatoid Arthritis

Study I4V-MC-JAJA is a Phase 3b/4, multicenter, randomized, parallel-group, active-controlled, outpatient, event-driven study to evaluate the effects of baricitinib (2-mg and 4-mg daily) compared with tumor necrosis factor (TNF) inhibitors (administered according to approved local labeling) on risk of VTE and other key safety endpoints in patients with RA. Treatment will be open-label for patients and study-site personnel.

Eli Lilly and Company
Indianapolis, Indiana USA 46285

Protocol I4V-MC-JAJA
Phase 3b/4

VV-CLIN-167399

Approved on 21 Jul 2025 GMT

2. Table of Contents

Section **Page**

Table of Contents

Table **Page**

Table of Contents

Figure **Page**

3. Revision History

SAP Version	Approval Date	Change
1	17 April 2019	Not applicable, original version
2	13 May 2020	The major changes incorporated in Version 2 were as follows: definition of venous thromboembolism (VTE) was updated to clarify that deep vein thrombosis (DVT) will include nonsuperficial below-knee thrombosis events added a sensitivity analysis consisting of a Cox proportional hazards regression model with study and randomization stratification factors as stratifying variables and treatment group as an explanatory variable per the Food and Drug] α ‘)Aec]) request
3	13 May 2020	The revision of Statistical Analysis Plan Version 3 was to allow the use of biosimilars within tumor necrosis factor (TNF) inhibitor treatment arm. No other changes were made to the study design and analysis. Removed brand names Humira® and Enbrel® and replaced with the active ingredients names to allow switch to biosimilars.
4	See date on cover	Statistical Analysis Plan Version 4 is the final version and was approved before database lock. The major changes incorporated in Version 4 are as follows: The per protocol population and associated analyses were removed in alignment with the study protocol and the main safety objectives. The per-protocol analysis compares the response for treatment A for those who adhere to treatment A with the response for treatment B for those who adhere to treatment B. This estimate does not correspond to a causal estimand (Lipkovich et al. 2020) and is not consistent with the recommendation in ICH E9 (R1). Therefore, we moved this analysis. The wording regarding interim

		and ad hoc analysis has been aligned with the updates specified in protocol amendment (e) The Statistical Analysis Plans for studies JAJA and JAJD were homogenized Additional safety analyses have been included in the Safety assessments, as per the Program Safety Analysis Plan for Baricitinib Version 7.

4. Study Objectives

4.1. Primary Objective

The primary objective of the study is the following, where venous thromboembolism (VTE) is defined as deep vein thrombosis (DVT, including non-superficial below knee thrombosis) and pulmonary embolism (PE):

Objective	Endpoint
• To compare baricitinib (combined dose groups) to TNF inhibitors with respect to VTE.	• Time from first dose of study treatment to first event of VTE.

Abbreviations: TNF = tumor necrosis factor; VTE = venous thromboembolism that consists of DVT, including nonsuperficial below-knee thrombosis, and PE.

If the accumulation of events for VTE is lagging considerably behind that anticipated when planning study enrollment, then an unblinded assessment will occur after the CCI of the expected patient-years of exposure (PYE) for the final analysis, which is expected to account for at least CCI patient-years (PY) or CCI of the anticipated total PYE. When this unblinded assessment confirms that a 1-sided 95% confidence interval (CI) for the aggregate incidence rate excludes a rate of CCI, the study's primary objective will be modified to the following to rule out an incidence rate difference between groups of CCI provided CCI patients with events have been observed.

Objective	Endpoint
• To compare baricitinib (combined dose groups) to TNF inhibitors with respect to VTE.	• Difference in incidence rate of VTE.

Abbreviations: TNF = tumor necrosis factor; VTE = venous thromboembolism that consists of DVT, including nonsuperficial below-knee thrombosis, and PE.

4.2. Secondary Objectives

The secondary objectives of the study are the following:

Objectives (Not Multiplicity Controlled)	Endpoints
• To compare baricitinib (combined dose groups) to TNF inhibitors with respect to key safety outcomes.	• Time from first dose of study treatment to first event of o ATE o MACE o Malignancy (excluding NMSC) o Opportunistic infection o Serious infection
• To compare each baricitinib dose to TNF inhibitors with respect to key safety outcomes.	• Time from first dose of study treatment to first event of o VTE o ATE o MACE o Malignancy (excluding NMSC) o Opportunistic infection o Serious infection

Secondary Objectives

Abbreviations: ATE = arterial thromboembolic event; DVT = deep vein thrombosis; MACE = major adverse cardiovascular events; NMSC = nonmelanoma skin cancer; PE = pulmonary embolism; TNF = tumor necrosis factor; VTE = venous thromboembolism that consists of DVT, including nonsuperficial below-knee thrombosis, and PE.

4.3. Exploratory Objectives

The exploratory objectives of the study are the following:

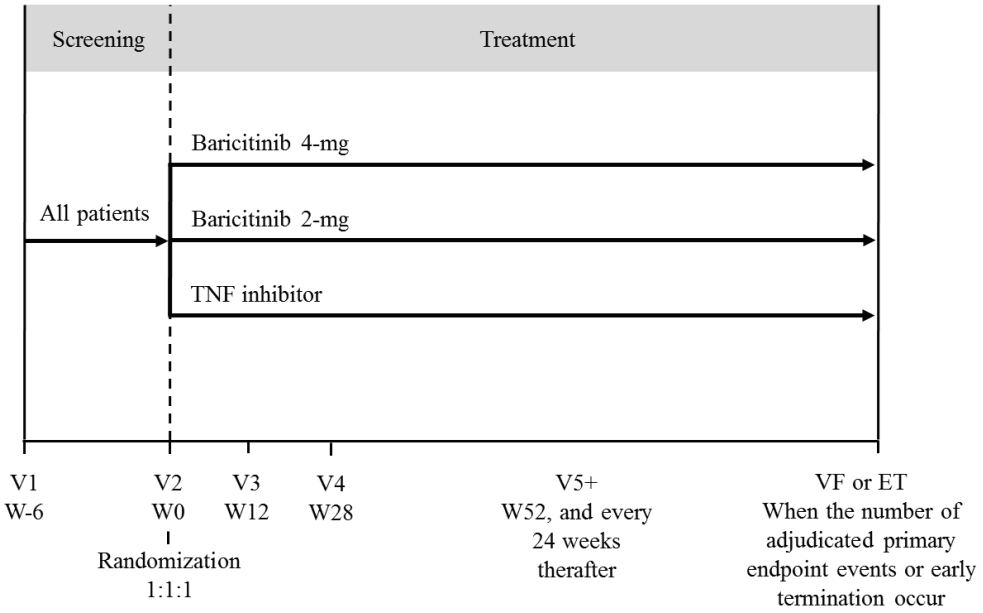
Objectives	Endpoints
To describe the effects of baricitinib (individual and combined dose groups) and TNF inhibitors with respect to other safety outcomes.	Individual components of key safety outcomes o VTE o MACE ATE Malignancy (excluding NMSC) Opportunistic infection Serious infection SAEs TEAEs AEs leading to permanent study drug discontinuation
To compare baricitinib (individual dose groups) to TNF inhibitors with respect to their effects on disease activity measures and patient-reported outcomes.	ACR20, ACR50, and ACR70 DAS28-CRP SDAI CDAI HAQ-DI Tender joint counts Swollen joint courts o‘ <)Global Assessment of Disease Activity o „ ‘)f ’ ‘)] < <))c <‘ <)] „ RAID pain domain score CRP PROMIS measures: Pain Interference, Depression, Cognitive Function, Satisfaction with Social Roles/Activities, Sleep Disturbance RAID

Abbreviations: ACR20/50/70 = American College of Rheumatology 20%/50%/70% response rate; AE = adverse event; ATE = arterial thromboembolic event; CDAI = Clinical Disease Activity Index; CRP = C-reactive protein; DAS28 = Disease Activity Score based on 28-joint assessment; DVT = deep vein thrombosis; HAQ-DI = Health Assessment Questionnaire-Disability Index; MACE = major adverse cardiovascular events; PE = pulmonary embolism; PROMIS = Patient-Reported Outcome Measurement Information System; RAID = Rheumatoid Arthritis Impact of Disease; SAE = serious adverse event; SDAI = Simplified Disease Activity Index; TEAE = treatment-emergent adverse event; TNF = tumor necrosis factor; VTE = venous thromboembolism that consists of DVT, including nonsuperficial below-knee thrombosis, and PE.

5. Study Design

5.1. Summary of Study Design

Study I4V-MC-JAJA (JAJA) is a Phase 3b/4, multicenter, randomized, parallel-group, active-controlled, outpatient, event-driven study to evaluate the effects of baricitinib (2-mg and 4-mg daily) compared to tumor necrosis factor (TNF) inhibitors (adalimumab or etanercept, administered according to approved labeling) on risk of VTE and other key safety endpoints in patients with rheumatoid arthritis (RA) (Figure JAJA.5.1).



Abbreviations: ET = early termination visit; TNF = tumor necrosis factor; V = visit; VF = final visit; W = week.

Figure JAJA.5.1. Study I4V-MC-JAJA design.

5.2. Determination of Sample Size

The results of Study JAJA will be combined analytically with another randomized study (Study I4V-MC-JAJD [JAJD]) run concurrently and comparing the same 3 treatment groups with the objective of assessing the relative risk of VTE for baricitinib (combined doses and individual dose groups) versus TNF inhibitors. The primary inferential statistical analyses will include data from both studies.

At least[redacted] patients with events from the combined studies' planned final analysis ensures [redacted] power to demonstrate that the hazard ratio for the risk of VTE events in the combined baricitinib group (i.e., across dose groups) relative to the TNF inhibitor group is [redacted] based on the upper limit of the 95% CI of the hazard ratio, assuming a true hazard ratio of [redacted] Following this assumption and prorating to approximately [redacted] events across 2 treatment groups (1 baricitinib group and the TNF inhibitor group), this also ensures [redacted] power to demonstrate

that the hazard ratio for the risk of VTE in each baricitinib dose group relative to the TNF inhibitor group is CCI based on a similar approach.

Studies JAJA and JAJD together aim to enroll approximately 3900 patients (2600 in JAJA, 1300 or more in JAJD) in a 1:1:1 ratio among baricitinib 4-mg, baricitinib 2-mg, and TNF inhibitor groups in order to observe the planned number of events.

5.3. Method of Assignment to Treatment

Patients who meet all criteria for enrollment will be randomized in a 1:1:1 ratio (baricitinib 4-mg, baricitinib 2-mg, TNF inhibitor) to open-label treatment at Visit 2. Randomization will be stratified based on the following factors and categories:

- history of a VTE: yes versus no
- score derived from age and body mass index (BMI): >3 versus ≤3, where score is calculated as:

  Min(Max(2(Age-45)/30, 0), 2) + Min(Max(3*(BMI-20)/20, 0), 3)*
- prior inadequate response or intolerance to a TNF inhibitor: yes versus no, and
- geographic region: US and Canada versus Europe versus Rest of World.

Assignment to treatment groups will be determined by a computer-generated random sequence using an interactive web-response system (IWRS).

6. A Priori Statistical Methods

6.1. General Considerations

This plan describes a priori statistical analyses for safety, efficacy, and health outcomes that will be performed.

Statistical analysis of this study will be the responsibility of Eli Lilly and Company (Lilly) or its designee. The statistical analyses will be performed using SAS® Version 9.4 or a more recent version.

Not all displays described in this SAP will necessarily be included in the clinical study report 4b r q 5>))m)‘)α ‘)) <,,< ‘)’ <),, <‘ <α)‘)‘) ‘ ,,)α ‘ >))r <) ‘)’ <) incorporated into interactive display tools instead of or in addition to a static display. Any display described in this SAP and not included in the CSR would be available upon request.

Data collected at early termination visits or the final visit at study closure will be mapped to the next scheduled visit number for that patient. For by-visit summaries, only visits in which a measure was scheduled to be collected will be summarized. Any unscheduled visit data will be included in the patient-level listings.

All descriptive data summaries will be based on observed values only and will include means, medians, standard deviation (SD), and ranges for continuous variables and absolute/relative frequencies for categorical data. In addition, for each variable the number (count) of values that are missing may be reported. Missing data are data that are planned and can be collected but are absent. Statistics will be summarized for the study population overall and by treatment arm.

All tests of treatment effects and CIs will be conducted at a 2- α<α)) < <))B>BG8) < s otherwise stated.

For the TNF inhibitor (adalimumab or etanercept administered according to the local product label) treatment group, there will be no distinctions between adalimumab and etanercept (e.g., cycling between adalimumab and etanercept will be ignored). Hence, safety, efficacy, health outcomes, and other data collected on either treatment will be attributed to the TNF inhibitor group. For the combined baricitinib treatment group, there will be no distinction between baricitinib 4 mg and baricitinib 2 mg (e.g., rescue from 2 mg to 4 mg will be ignored, or dose reduction from 4 mg to 2 mg).

Patients with estimated glomerular filtration rate (eGFR) <60 mL/min/1.73 m^2 at screening who are randomized to the baricitinib 4-mg dose will receive a dose of 2 mg once daily but will be included in the 4-mg dose group for analysis.

Estimands (International Conference on Harmonisation [ICH] E9 R1) associated with each endpoint/analysis are documented in the following places:

Population of interest: patients with RA and meeting the protocol inclusion/exclusion criteria. Analysis populations are defined in Section 6.1.1. The analysis population corresponding to each of the safety, efficacy, and health outcomes endpoints is specified in Table JAJA.6.5 and Table JAJA.6.22.

Endpoints/variables: primary, secondary, and exploratory endpoints/variables are listed in Table JAJA.6.4 and Table JAJA.6.21.

Population-level summary: difference in time from first dose of study treatment to first event of interest, difference in response proportions for binary variables, and difference in means of continuous variables between treatment groups. Details are given in Table JAJA.6.5 and Table JAJA.6.22.

Handling of intercurrent event(s): Specific statistical methods to be used for handling intercurrent events under different strategies are described in Section 6.3. Statistical methods corresponding to safety, efficacy, and health outcomes variables are summarized in Table JAJA.6.5 and Table JAJA.6.22.

Analyses specified in this document will be performed based on data collected in Study JAJA alone, unless otherwise specified.

6.1.1. Analysis Populations

Efficacy and health outcomes analyses will be conducted on the intent-to-treat (ITT) population, whereas the primary, secondary, and exploratory safety analyses will be conducted on the safety population, as described in Table JAJA.6.1.

In the rare situation where a patient is lost to follow-up at the first postbaseline visit but some safety data exist (e.g., unscheduled laboratory assessments) after first dose of study drug, a listing of the data or a patient profile will be provided.

Table JAJA.6.1. Analysis Populations

Population	Description
Intent-to-Treat (ITT) Population	All patients who are randomized will be included in the ITT population. Patients will be analyzed according to the investigational product to which they were randomized.
Safety Population	The safety population is defined as all randomized patients who receive at least 1 dose of investigational product and who did not discontinue from the α)) <) <‘) k))e -)‘) <)) ’ ‘ < <) >)) Patients will be analyzed according to the investigational product to which they were randomized.

6.1.2. Definition of Baseline and Postbaseline Measures

6.1.2.1. Baseline and Postbaseline Definition for Safety Measures

Table JAJA.6.2 describes the baseline and postbaseline definitions for the different types of safety analyses.

Table JAJA.6.2. Baseline and Postbaseline Definition for Safety Measures

Analysis Type	Baseline	Postbaseline
Treatment-emergent adverse events (TEAEs) including special topics	The entire period up to date of first study drug administration (typically at Week 0).	For the final lock, first day of study drug administration up to 30 days after the last dose of the study drug. For interim locks, first day of study drug administration up to data cut for ongoing patients, and up to 30 days after the last dose of the study drug or up to the cutoff date, whichever occurs first, for patients who discontinued drug prior to the data cut.
Adverse events (AEs) leading to study drug discontinuation, AEs leading to study drug interruption, and serious adverse events (SAEs)	NA	For the final lock, first day of study drug administration up to 30 days after the last dose of the study drug. For interim locks, first day of study drug administration up to data cut for ongoing patients, and up to 30 days after the last dose of the study drug or up to the cutoff date, whichever occurs first, for patients who discontinued drug prior to the data cut.
Treatment-emergent (TE) low and high summaries in labs (including Common Terminology Criteria for Adverse Events (CTCAE) shift summaries), vital signs	The entire period (including „ <α <α) ‘ <α)‘ α) „ <α <α) ‘ <α) measurements) up to date of first study drug administration (including baseline pre-dose measurements).	For the final lock, post-first dose of study drug <‘ < <)4 „ α) ‘ <α)‘ α) ‘ <α) <‘ < < 5 up to 30 days after the last dose of the study drug. For interim locks, post-first dose of study drug <‘ < <)4 „ α) ‘ <α)‘ α) ‘ <α) <‘ < < 5) up to data cut for ongoing patients, and up to 30 days after the last dose of the study drug or up to the cutoff date, whichever occurs first, for patients who discontinued drug prior to the data cut. Note that for vital signs, there might not be any ‘ <α) <‘ < < 8))) ‘ <α) measurements will be used.

Baseline and Postbaseline Definition for Safety Measures

Analysis Type	Baseline	Postbaseline
Analyses by time points, change from last baseline to last post-baseline observation in labs and vital signs	Last scheduled "planned" measurements prior to or at date of first study drug dose administration.	For the final lock, post-first dose of study drug "planned" measurements up to the last visit. For interim locks, post-first dose of study drug "planned" measurements up to data cut for ongoing patients, and up to the last visit for patients who discontinued prior to the data cut. The early termination visits (ETV) and the last visits at study closure are considered "planned" visits.

6.1.2.2. Baseline and Postbaseline Definition for Efficacy and Health Outcomes Measures

Baseline will be defined as the last available value on or prior to the date of the first dose of investigational product for efficacy and health outcomes measures. In most cases, this will be the measure recorded at Week 0 (Visit 2).

Postbaseline measurements are collected after study drug administration. For data collected in the electronic Clinical Outcomes Assessment (eCOA) tablet (including Patient-Reported Outcomes [PRO] and Clinician-Reported Outcomes [ClinRO]) and related to efficacy assessments, postbaseline data collected outside of the scheduled visit but within the visit window defined by Lilly will be summarized in the by-visit analyses if there are no data available at the scheduled visit. That is, a +7 day window is applied to all postbaseline visits. If there is more than 1 set of data within the defined visit window and no data available at the scheduled visit, the data collected closest to the scheduled visit date will be used.

6.1.3. Analysis Methods

The main analysis of time-to-event variables will use a Cox proportional hazards regression model stratified by study with treatment group and randomization stratification factors (VTE history, age/BMI combination, prior inadequate response or intolerance to a TNF inhibitor, and geographic region) as explanatory variables. Hazard ratios, their 95% CIs, and associated p-values will be presented for treatment comparisons and assessments of superiority/inferiority and noninferiority, where applicable. In addition, a sensitivity analysis consisting of a Cox proportional hazards regression model, with study and randomization stratification factors as stratifying variables and treatment group as an explanatory variable, will be performed. In the event of a low number of observed events or substantial imbalance in event distribution across randomization stratification factors at study completion, strata may be combined or certain stratification variables excluded from the Cox proportional hazards model to enhance model stability and ensure convergence. Additionally, a pre-specified hierarchical approach will be implemented if the primary Cox proportional hazards model fails to converge. CCI [redacted]

CCI [redacted]

This same hierarchical convergence procedure will be applied consistently to individual study analyses (JAJA and JAJD analyzed separately), while Study JAJD excludes geographical region entirely (US-only study), and any supplementary analyses conducted per study to ensure uniformity in model convergence approaches across all analytical components**.** All convergence attempts and model modifications will be documented, with the final model specification clearly identified and any deviations from the primary analysis treated as sensitivity analyses.

Poisson regression will be used to evaluate incidence rates of adverse events (AEs), serious adverse events (SAEs), discontinuations, and other categorical safety data (Section 6.10.3.5) for between-treatment-group comparisons.

For other safety measures, the following statistical methods will be used unless otherwise noted:

- Fisher's Exact test will be used to compare percentages, and odds ratios will be displayed. Odds ratios will be created with baricitinib treatment as the numerator and TNF inhibitor as the denominator. When Study JAJA and JAJD are combined for analyses, the Cochran–Mantel–Haenszel (CMH) test, stratified by study, will be used.
- Treatment differences in mean change for continuous measurements will be assessed using an analysis of covariance (ANCOVA) model containing terms for treatment and the continuous covariate of baseline measurement. Type 3 sums of squares will be used.

For document writing purposes for safety (other than the primary safety analysis), tests with 2-sided p-values <0.05 will be referred to as having strong statistical evidence for a treatment difference, unless otherwise noted. However, p-values should not be over-interpreted for these safety analyses. Except for prespecified hypotheses, they correspond to data-driven hypotheses and hence are only useful as a flagging mechanism.

Categorical efficacy variables and health outcomes variables may be analyzed using a logistic regression model with geographic region, prior inadequate response or intolerance to a TNF inhibitor, BMI, and treatment group as explanatory factors, providing the p-value and 95% CI for the odds ratio for treatment group differences. The difference in unadjusted percentages and 95% CI of the difference using the Newcombe–Wilson method without continuity correction are used for descriptive purposes, unless otherwise specified. When there are quasi-complete or complete separation warnings from the logistic regression analyses, Firth's penalized likelihood approach may be used. Missing data may be imputed using a nonresponder imputation (NRI) method (Section 6.3.2).

Continuous efficacy and health outcomes variables may be analyzed using a restricted maximum likelihood-based mixed model for repeated measures (MMRM). The model will include geographic region, prior inadequate response or intolerance to a TNF inhibitor, BMI, treatment, visit, and treatment-by-visit interaction as fixed categorical effects, and baseline value and baseline value-by-visit interaction as fixed continuous effects to estimate change from baseline across postbaseline visits. An unstructured covariance matrix will be used to model the

between- and within-patient errors. If the unstructured covariance configuration results in a lack of convergence, then the following variance-covariance matrix choices will be used (in order) until 1 converges: Heterogeneous Toeplitz; Heterogeneous First Order Autoregressive; Heterogeneous Compound Symmetry; Toeplitz; First Order Autoregressive; Compound Symmetry. If all variance-covariance matrices selected fail to converge, Variance Components will be used as the next attempt at convergence. The Kenward-Roger method will be used to estimate the denominator degrees of freedom. Type III tests for the least-squares (LS) means will be used for the statistical comparisons. The LS mean difference, standard error, p-value, and CIs will also be reported. Treatment group comparisons at specific study visits will be tested using the t-test obtained from the MMRM results.

6.2. Adjustments for Covariates

Unless otherwise specified, the following general approach will be used with regard to inclusion of covariates:

> For time-to-event analyses of the primary and secondary objectives, the treatment, VTE history, age/BMI combination, prior inadequate response or intolerance to a TNF inhibitor, and geographic region will be included as covariates. In addition, when Study JAJA and JAJD are combined for the primary and secondary objective, the analysis will be stratified by study.

When ANCOVA is used for laboratory results, treatment and baseline value will be included as covariates in the model.

When MMRM is used for efficacy and health outcomes analyses, geographic region, prior inadequate response or intolerance to a TNF inhibitor, BMI, treatment, visit, treatment-by-visit interaction, baseline value, and baseline value-by-visit interaction will be included as covariates.

When a logistic regression model is used for efficacy and health outcomes analyses, geographic region, prior inadequate response or intolerance to a TNF inhibitor, BMI, and treatment will be included as covariates in the model.

6.3. Handling of Dropouts or Missing Data

Intercurrent events (ICH E9 R1) are events which occur after the treatment initiation and make it impossible to measure a variable or influence how it would be interpreted. Potential intercurrent events include: rescue from baricitinib 2 mg to 4 mg, start of a nonpermitted RA treatment (i.e., initiation of any targeted synthetic disease-modifying antirheumatic drugs [tsDMARDs] or biologic disease-modifying antirheumatic drugs [bDMARDs] other than adalimumab or etanercept in the TNF inhibitor group, or initiation of any bDMARD or tsDMARD other than baricitinib in the baricitinib group), treatment discontinuation due to death or AEs, lost to follow-up, etc. The methods for handling these intercurrent events are described below.

6.3.1. Censoring Rules for Time-to-Event Safety Analyses

For analysis of primary and secondary endpoints other than malignancy, events will be attributed to patients according to their randomized treatment group assignment through the date of their

last dose of randomized study medicine plus up to 30 days, censored after rescue from baricitinib 2 mg to 4 mg or initiation of a nonpermitted RA treatment. For analyses of combined baricitinib dose groups, censoring will not occur upon rescue from baricitinib 2 mg to 4 mg, and all events occurring on either dose will be attributed to the combined dose group. The window of 30 days allows for the accumulation of events that take time to accrue, while censoring is applied to avoid the attribution of effects of subsequent therapies.

Because malignancies may result from an extended biological process, such events will be attributed to patients through the date of their last available study visit, without regard for last dose of randomized study medicine, but censored after rescue to baricitinib 4 mg within the baricitinib 2-mg group (for analyses of individual baricitinib dose groups) or initiation of a nonpermitted RA treatment.

These censoring rules can be justified based on the while-on-treatment strategy (ICH E9 R1) for handling intercurrent events.

Supportive analyses will be conducted in which all events during the study are attributed to the ‘ <) ‘ α <α) <‘ <)) < ‘ α <))‘) ’ < <) < „ <)) ‘))‘) nonpermitted RA treatment. In this case, the ICH E9 R1-defined treatment policy strategy is being followed. However, patients who discontinue the study early will be censored at the time of discontinuation and thus have missing data, which follows the while-on-treatment strategy. So, this supportive analysis can be justified based on the combination of the treatment policy and while-on-treatment strategies for handling intercurrent events.

6.3.2. Handling of Dropouts or Missing data for Efficacy and Health Outcomes Endpoints

As with any clinical study, patient dropouts and, consequently, missing data are expected. Patients who discontinue assigned treatment, unless they withdraw consent, will be encouraged to remain in the study and follow the scheduled visits to provide additional safety data. Although every effort will be made to reduce missing data, the methods described below may be used to provide a conservative approach for assessing efficacy and health outcomes endpoints when patients have missing data.

The following rules may be used to handle dropouts or missing data:

> Nonresponder imputation: For analysis of categorical efficacy and health outcomes data, patients who discontinue study treatment early, who are rescued from baricitinib 2 mg to 4 mg, or who initiate nonpermitted RA treatments will be defined as nonresponders after discontinuation, rescue, or new medicine initiation, respectively. This method can be justified based on the composite strategy (ICH E9 R1) for handling of intercurrent events. In this strategy, discontinuation of study treatment and intake of rescue or nonpermitted RA treatment are being treated as unfavorable outcomes in the analysis.

- Mixed-model repeated measures: For continuous data, MMRM analysis will be applied with a missing-at-random assumption for handling missing data. This analysis takes into account both missingness of data and the correlation of the repeated measurements. No imputation methods will be applied to the MMRM analysis. Observed data will be analyzed with the exception that any data that occur after rescue from baricitinib 2 mg to 4 mg, study treatment discontinuation, or initiation of nonpermitted RA treatments will be censored in order to estimate the effects of the randomized treatment. This method can be justified based on the hypothetical strategy (ICH E9 R1) for handling of intercurrent events. In this strategy, the scientific question of interest is to assess the effect of study treatment in a hypothetical trial where all patients have complete data and continue to take study treatment without intercurrent events.

6.4. Multiple Comparisons/Multiplicity

No multiplicity control will be conducted for the analyses described in this SAP other than the evaluation of the primary objective over repeated interim and final analyses (Section 6.14).

6.5. Active-Control Studies Intended to Show Equivalence

The primary objective of this study is to demonstrate that the hazard ratio for the risk of VTE events in the combined baricitinib group (i.e., across dose groups) relative to the TNF inhibitor group is CCI based on the CCI CI of the hazard ratio for time to first event.

6.6. Patient Disposition

Patient flow will be summarized from entry to randomization to first dosing to completion, and analysis population will be listed and summarized by treatment group.

Patient disposition from study will be listed and summarized with reasons for disposition for the ITT population. The number of patients rescued from baricitinib 2 mg to 4 mg, dose reduction from 4mg to 2mg or cycled between adalimumab and etanercept will be summarized as well.

6.7. Patient Characteristics

Patient demographic variables and baseline characteristics are listed in Table JAJA.6.3 and will be summarized with all treatment groups combined and by treatment groups for the safety population. No formal statistical comparisons will be made among treatment groups. In addition, a by-patient listing will be provided.

Table JAJA.6.3. Patient Characteristics

Variable	Quantitative Summary	Categorical Summary	Derivation
Demographics			
Age	Yes	< 50, ≥50 years	Age (year) = (number of months between first dose date and July 1st of birth year) / 12, and truncated to a whole-year (integer) age. If first dose date is missing, randomization date will be used.
		<60, ≥60 years	
		<65, ≥65 years	
		<75, ≥75 years	
		<85, ≥85 years	
		<65, ≥65 and <75, ≥75 and <85, ≥85 years	
Sex	No	Male, Female	
Ethnicity	No	Hispanic/Latino, Non-Hispanic/Non-Latino, Not Applicable	
Race	No	American Indian/Alaska Native, Asian, Black/African American, Native Hawaiian or other Pacific Islander, White, or Multiple	
Country	No	Will summarize based on final countries that participated	
Region	No	United States and Canada, Europe, Rest of the world United States and Outside United States	
Baseline Anthropometric Measures			
Height (cm)	Yes	None	Height (cm) = height (in) * 2.54
Weight (kg)	Yes	<60 kg, ≥60 kg to <100 kg, ≥100 kg	Weight (kg) = weight (lbs) * 0.454
BMI	Yes	<18.5, ≥18.5 and <25, ≥25 and <30, ≥30 and <35, ≥35 and <40, ≥40 kg/m^2	BMI (kg/m^2) = Weight (kg)/(Height (m))^2
CCI			

Patient Characteristics

Variable	Quantitative Summary	Categorical Summary	Derivation
Baseline RA history			
Duration of RA from diagnosis (year)	Yes	none	Duration of RA from diagnosis (year) = [(Date of first dose Date of RA diagnosis) +1]/ 365.25 If the year of RA diagnosis is missing, the duration will be set as missing. Otherwise, an unknown month will be taken as January, and an unknown day will be taken as 01. If first dose date is missing, randomization date will be used.
Current use of corticosteroid	Yes	Yes, No	
Number of csDMARDs previously used	No	0, 1, 2, 3	
Number of csDMARDs currently used	No	0, 1, 2, 3	
Type of csDMARDs currently used	No	no csDMARDs, MTX only, MTX + 1 other csDMARD, 1 non-MTX csDMARD only, 2 non-MTX csDMARDs only, 3 or more csDMARDs	
Number of bDMARDs previously used	No	0, 1, 2, 3	
Prior inadequate response or intolerance to TNFi	No	Yes, No	
bDMARDs previously used	No	Humira®, Enbrel®, etc.	
Other risk factors			
Tobacco use	No	Never, current, former	eCRF specific entry
Diabetes	No	Yes, No	Prior medical history pre-specified preferred terms and lowest level terms
Prior VTE	No	Yes, No	eCRF specific entries
Prior ATE	No	Yes, No	
NSAID use	No	Yes, No	Prior and concomitant medications with pre-specified ATC codes
COX2 inhibitor use	No	Yes, No	

Patient Characteristics

Variable	Quantitative Summary	Categorical Summary	Derivation
OCP/SERM use	No	Yes, No	Prior and concomitant medications with pre-specified ATC codes
HRT	No	Yes, No	
Baseline RA clinical characteristics			
Tender joint count based on 28 joints (TJC28)	Yes	None	Number of tender joints in the 28 joints that are evaluated. When there are nonevaluable joints, tender joints will be prorated to 28 joints.
Swollen joint count based on 28 joints (SJC28)	Yes	None	Number of swollen joints in the 28 joints that are evaluated. When there are nonevaluable joints, tender joints will be prorated to 28 joints.
Health Assessment Questionnaire Disability Index (HAQ-DI)	Yes	None	P‘ <)‘ < <)) „‘) function is measured by HAQ-DI, which consists of 24 questions referring to 8 domains with a range of possible scores (0-3): dressing and grooming, arising, eating, walking, hygiene, reach, grip, and activities. Scoring will follow the index manual. Missing when there are less than 6 domains completed.
o „ ‘) global assessment of disease activity (0-10 mm) (PhGA)	Yes	None	PhGA is measured using a 0-100 numeric rating scale and will be divided by ten for the purpose of summarizing the data. Single item score reported by physician, missing if missing.
o‘ <) ’ ‘) assessment of disease activity (0-10 mm) (PtGAm)	Yes	None	PtGAm is measured using a 0-100 numeric rating scale and will be divided by ten for the purpose of summarizing the data. Single item score reported by patient, missing if missing.
Rheumatoid Arthritis Impact of Disease (RAID) pain domain (Numeric Rating Scale)	Yes	None	

Patient Characteristics

Variable	Quantitative Summary	Categorical Summary	Derivation
Acute phase reactant as measured by high-sensitivity C-reactive protein (hsCRP) (mg/L)	Yes	None	
Simplified Disease Activity Index (SDAI)	Yes	k)α < <)‘ „)4 CC58) α< ‘ <) (>11 - DH58))4SDH5	Measure of disease activity that integrates measures of physical examination, acute phase response, patient self-assessment, and evaluator assessment using the following variables: TJC28, SJC28, hsCRP, PtGAm, and PhGA. Calculated as: TJC28 + SJC28 + hsCRP (0.1-10.0 mg/dL) + PtGAm (0-10) + PhGA (0-10).[a] Missing if any of the component score is missing.
Clinical Disease Activity Index (CDAI)	Yes	k)α < <)‘ „)4 CB58) α< ‘ <) (>10 - DD58))4SDD5	Measure of disease activity that integrates measures of physical examination, patient self-assessment, and evaluator assessment using the following variables: TJC28, SJC28, PtGAm, and PhGA. Calculated as: TJC28 + SJC28 + PtGAm (0-10) + PhGA (0-10).[a] Missing if any of the component score is missing.
Disease activity score (DAS28-hsCRP)	Yes	Low disease activity E>D8)moderate disease activity SE>D)) G<C8)high disease activity >5.1	Measure of disease activity based on 28 joints that consists of a composite numeric score of the following variables: TJC28, SJC28, hsCRP, and PtGAm. Calculated as: $0.56(TJC28)^{1/2} + 0.28(SJC28)^{1/2} + 0.36(\ln[hsCRP+1]) + 0.014(PtGAm) + 0.96$.[b] Missing if any of the component score is missing.
Rheumatoid factor (RF) (IU/mL)	Yes	Negative, Positive	If 0 RF 14 IU/mL then set to 'Negative'; else if RF>14 IU/mL then set to 'Positive'.

Patient Characteristics

Variable	Quantitative Summary	Categorical Summary	Derivation
ACPA (U/mL)	No	Negative, Positive, Indeterminate	If 0 ACPA 7 U/mL then set to 'Negative'; else if 7<ACPA 10 U/mL then set to 'Indeterminate'; else if ACPA>10 U/mL then set to 'Positive'.
Seropositivity status	No	RF negative and ACPA negative, RF positive and ACPA negative, RF negative and ACPA positive, RF positive and ACPA positive	For seropositivity status category, the ACPA negative group includes patients with negative and indeterminate values: 0 ACPA 10 U/mL.

Abbreviations: ACPA = anticitrullinated peptide antibodies; ATC = Anatomical Therapeutic Chemical; ATE = arterial thromboembolism; bDMARDs = biologic DMARDs; BMI = body mass index; csDMARD = conventional synthetic DMARD; COX = cyclooxygenase; DMARDs = disease-modifying antirheumatic drugs; DVT = deep vein thrombosis; HRT = hormone replacement therapy; MTX = methotrexate; OCP = oral contraceptive pills; PE = pulmonary embolism; RA = rheumatoid arthritis; SERM = selective estrogen receptor modulators; VTE = venous thromboembolism that consists of DVT, including nonsuperficial below-knee thrombosis, and PE.

[a] Aletaha and Smolen 2005.

[b] Vander Cruyssen et al. 2005.

6.8. Treatment Compliance

The compliance with baricitinib will be assessed through counts of dispensed and returned investigational product tablets for baricitinib. Treatment compliance (%) for each patient in the baricitinib group will be calculated as:

$$\text{Compliance} = \frac{\text{total number of tablets dispensed} - \text{total number of tablets returned}}{\text{expected number of total tablets}}$$

where

total number of tablets dispensed: sum of tablets dispensed in the period of interest prior to Visit *x*

total number of tablets returned: sum of the tablets returned in the period of interest prior to and including Visit *x*

expected number of tablets: number of days in the period of interest*number of tablets taken per day = ([date of visit date of first dose + 1] number of days of temporary drug interruption)*number of tablets taken per day

The compliance with TNF inhibitors will be assessed at each visit through data present in the electronic case report form (eCRF) .

For patients who had their treatment temporarily interrupted by the investigator, the period of time that dose was withheld will be taken into account in the compliance evaluation.

A patient will be considered significantly noncompliant if he or she misses more than 20% of the prescribed doses during the study or during key study visit intervals, that is, compliance is <80%. Similarly, a patient will be considered significantly noncompliant if he or she is judged by the investigator to have intentionally or repeatedly taken more than the prescribed amount of medication.

Descriptive statistics for percent compliance will be summarized for the safety population for the baricitinib group at a yearly basis and across the entire study.

The proportion of patients significantly noncompliant will be summarized at a yearly basis and across the entire study for both treatment groups.

6.9. Concomitant Therapy

Concomitant therapy is defined as the therapy that starts before, on, or after the first dose of study treatment and continues into the treatment period.

Concomitant therapy will be recorded at each visit and will be summarized using frequency counts and percentages by the World Health Organization (WHO) drug Anatomic Therapeutic Chemical (ATC) classification and treatment group for the safety population.

In addition, summaries of concomitant therapy will be provided for each treatment group and the combined baricitinib group, using data up to rescue, initiation of nonpermitted RA treatment, or 30 days after their last dose date of the originally randomized study treatment, with an exception that data collected after rescue will be used for the combined baricitinib group, for the following categories:

concomitant medications used for RA

concomitant medications: statins

concomitant medications excluding RA and statin therapies

6.10. Safety Analyses

6.10.1. Extent of Exposure

Duration of exposure to each study drug will be summarized for the safety population. Cumulative exposure and duration of exposure will be summarized in terms of frequency counts and percentages by category and treatment group.

Duration of exposure will be calculated as follows:

baricitinib 4-mg exposure = date of last baricitinib 4-mg treatment date of first baricitinib 4-mg dose + 1. The baricitinib 4-mg exposure gained after rescue from patients who were originally randomized to baricitinib 2-mg will not be included.

baricitinib 2-mg exposure = date of last baricitinib 2-mg treatment date of first baricitinib 2 mg dose + 1. For patients who were rescued from baricitinib 2-mg dose to 4-mg dose, their drug exposure after rescue will not be counted as baricitinib 2-mg exposure.

TNF inhibitor exposure = date of last TNF inhibitor treatment (Adalimumab or Etanercept) date of first TNF inhibitor dose (Adalimumab or Etanercept) + 1, regardless of treatment cycling from Adalimumab to Etanercept or vice versa.

Total baricitinib exposure = date of last baricitinib (2-mg or 4-mg) treatment date of first baricitinib (2-mg or 4-mg) dose + 1, regardless of rescue from baricitinib 2-mg to baricitinib 4-mg.

For patients randomized to the baricitinib 4-mg dose group who receive a dose of 2 mg once daily due to having an eGFR <60 mL/min/1.73 m^2 at screening, exposure will be considered as exposure to 4 mg and computed as total baricitinib exposure, regardless of actual dose taken.

Missing date handling will follow the compound-level safety standards.

The number and percentage of patients in each of the following exposure ranges will be included in the summaries:

6 months (24 weeks), 12 months (52 weeks), DF) (104 weeks)8) EHmonths (156 weeks), and so on. Note that patients may be included in more than 1 category.

>0 to <6 months, 6 months to <12 months, CD) to <DF) 8) DF) to <36 months, and so on

The summaries will also include the overall exposure in total PY, calculated as:

exposure in PY (PYE) = sum of duration of exposure in days (for all patients in treatment group) / 365.25

No p-values will be reported in these tables as they are intended to describe the characteristics of the study sets.

6.10.2. Primary, Secondary, Sensitivity, and Supplementary Analyses

The general methods used to summarize safety data, including analysis population definition and analysis methods, are described in Section 6.1.

Primary endpoint events (VTE) and selected secondary endpoint events, including certain arterial thromboembolic events (ATEs), cardiovascular (CV) events incorporating major adverse cardiovascular events (MACE), and opportunistic infections, will undergo adjudication by independent, blinded clinical endpoint committees (CECs). Only CEC-confirmed events will be included in the following endpoint analyses. Table JAJA.6.4 includes the descriptions and derivations of the primary and secondary safety measures.

Table JAJA.6.5 provides the detailed analyses including analysis type, population, and comparisons for safety analyses.

Table JAJA.6.4. Description and Derivation of Primary and Secondary Safety Measures

Measure	Description	Calculation of targeted analysis event time	Notes	Censoring type
Time to first VTE, Time to first ATE, Time to first MACE, Time to first Malignancy (excluding NMSC), Time to first opportunistic infection, Time to first serious infection	When a targeted analysis event occurs between first dose date and within 30 days of their last dose date of originally randomized treatment[a] (or date of their last available study visit for malignancies)	Event date first randomized dose date	The targeted analysis event and intercurrent event: rescue, initiation of nonpermitted RA treatment, discontinuation/completion of the study, whichever occurs first, the corresponding time will be used in the analysis. For analysis of malignancy, the last visit date will be used. **Note:** rescue only applies to analysis of individual baricitinib 2-mg dose vs. TNFi. In the combined baricitinib dose vs TNFi analysis, rescue will not be considered in the analysis. That is, events occurring on either dose will be attributed to the combined dose group.	Primary (in consideration of intercurrent events)
	When there is no targeted analysis event	(last dose date of the originally randomized treatment[a] + 30 days or last visit date, whichever the earliest) first randomized dose date		
	When nonpermitted RA treatment is initiated	Initiation date of the nonpermitted RA treatment first randomized dose date		
	When there is rescue	Rescue date first randomized dose date		
Time to first VTE, Time to first ATE, Time to first MACE, Time to first Malignancy (excluding NMSC), Time to first opportunistic infection, Time to first serious infection	When an event occurs between first dose date and date of their last available study visit	Event date first randomized dose date	All events occurring during the study will be ‘ ’ <α)) <) ‘ <) randomized treatment group regardless of any subsequent rescue, treatment discontinuation, or initiation of a nonpermitted RA treatment.	Supportive (regardless of intercurrent events)
	When there is no event	Last visit date - first randomized dose date		

Description and Derivation of Primary and Secondary Safety Measures

Abbreviations: ATE = arterial thromboembolic event; DVT = deep vein thrombosis; JAJD = I4V-MC-JAJD; MACE = major adverse cardiovascular event; NMSC = nonmelanoma skin cancer; PE = pulmonary embolism; TNFi = tumor necrosis factor inhibitor; RA = rheumatoid arthritis; VTE = venous thromboembolism that consists of DVT, including nonsuperficial below-knee thrombosis, and PE.

[a] When evaluating baricitinib vs. TNFi, the originally randomized treatment refers to baricitinib. When evaluating baricitinib 4 mg or 2 mg vs. TNFi, the originally randomized treatment is that specific dose of baricitinib.

Table JAJA.6.5. Description of Primary, Secondary, Sensitivity, and Supplementary Analyses

Measure	Analysis Method (Section 6.1.3.)	Population (Section 6.1.1)	Comparison	Censoring type	Analysis type
Time to first VTE	Cox proportional hazards regression model	Safety (JAJA and JAJD combined)	Combined baricitinib vs TNF inhibitor	Primary	Primary
Time to first VTE	Cox proportional hazards regression model	Safety (JAJA and JAJD combined)	Combined baricitinib vs TNF inhibitor	Supportive	Sensitivity
Time to first VTE	Cox proportional hazards regression model	Safety (JAJA alone)	Combined baricitinib vs TNF inhibitor	Primary	Supplementary
Time to first VTE	Cox proportional hazards regression model with study and stratification factors as stratifying variables	Safety (JAJA and JAJD combined)	Combined baricitinib vs TNF inhibitor	Primary	Sensitivity
Time to first ATE, Time to first MACE, Time to first Malignancy (excluding NMSC), Time to first opportunistic infection, Time to first serious infection	Cox proportional hazards regression model	Safety (JAJA and JAJD combined)	Combined baricitinib vs TNF inhibitor	Primary	Secondary

Description of Primary, Secondary, Sensitivity, and Supplementary Analyses

Measure	Analysis Method (Section 6.1.3.)	Population (Section 6.1.1)	Comparison	Censoring type	Analysis type
Time to first ATE, Time to first MACE, Time to first Malignancy (excluding NMSC), Time to first opportunistic infection, Time to first serious infection	Cox proportional hazards regression model	Safety (JAJA and JAJD combined)	Combined baricitinib vs TNF inhibitor	Supportive	Sensitivity
Time to first ATE, Time to first MACE, Time to first Malignancy (excluding NMSC), Time to first opportunistic infection, Time to first serious infection	Cox proportional hazards regression model	Safety (JAJA alone)	Combined baricitinib vs TNF inhibitor	Primary	Supplementary
Time to first VTE, Time to first ATE, Time to first MACE, Time to first Malignancy (excluding NMSC), Time to first opportunistic infection, Time to first serious infection	Cox proportional hazards regression model	Safety (JAJA and JAJD combined)	Baricitinib 4-mg vs TNF inhibitor	Primary	Secondary
Time to first VTE, Time to first ATE, Time to first MACE, Time to first Malignancy (excluding NMSC), Time to first opportunistic infection, Time to first serious infection	Cox proportional hazards regression model	Safety (JAJA alone)	Baricitinib 4-mg vs TNF inhibitor	Primary	Supplementary

Description of Primary, Secondary, Sensitivity, and Supplementary Analyses

Measure	Analysis Method (Section 6.1.3.)	Population (Section 6.1.1)	Comparison	Censoring type	Analysis type
Time to first VTE, Time to first ATE, Time to first MACE, Time to first Malignancy (excluding NMSC), Time to first opportunistic infection, Time to first serious infection	Cox proportional hazards regression model	Safety (JAJA and JAJD combined)	Baricitinib 2-mg vs. TNF inhibitor	Primary	Secondary
Time to first VTE, Time to first ATE, Time to first MACE, Time to first Malignancy (excluding NMSC), Time to first opportunistic infection, Time to first serious infection	Cox proportional hazards regression model	Safety (JAJA alone)	Baricitinib 2-mg vs. TNF inhibitor	Primary	Supplementary

Abbreviations: ATE = arterial thromboembolic event; DVT = deep vein thrombosis; JAJA = I4V-MC-JAJA; JAJD = I4V-MC-JAJD; MACE = major adverse cardiovascular event; NMSC = nonmelanoma skin cancer; PE = pulmonary embolism; TNFi = tumor necrosis factor inhibitor; VTE = venous thromboembolism that consists of DVT, including nonsuperficial below-knee thrombosis, and PE.

6.10.3. Other Safety Analyses (Exploratory)

Other safety assessments will include AEs, laboratory analytes, and vital signs. These safety data will be summarized by treatment group and analyzed using the safety population. Analyses will focus on comparisons of each baricitinib dose group to the TNF inhibitor group using the safety analysis methods described in Section 6.1. Baseline and postbaseline are defined in Table JAJA.6.2. Two analysis periods will be used depending on the analysis type: treatment period is defined as up to 30 days after the last dose of the study treatment; study period is defined as up to the last visit of the study.

Studies JAJA and JAJD may be combined for analyses including SAEs and AEs related to permanent or temporary study drug discontinuation. General AEs will be reported individually for study JAJA.

6.10.3.1. Adverse Events

The planned summaries for AEs are provided in Table JAJA.6.6 and are described more fully in the compound-level safety standards.

A treatment-emergent adverse event (TEAE) is defined as an event that either first occurred or worsened in severity after the first dose of study treatment and on or prior to the last visit date during the analysis period. Refer to the compound-level safety standards for handling of missing severity.

Table JAJA.6.6. Summary Tables Related to Adverse Events

Analysis	Analysis Period
An overview table, with the number and percentage of patients who experienced a TEAE, SAE, death, discontinuation from the study due to an AE, permanent discontinuation from study drug due to an AE, or a severe TEAE	Treatment period
An overview table, with the number and percentage of patients who experienced a TE AESI including VTE, MACE, ATE, malignancy (excluding NMSC), opportunistic and serious infections	Treatment period
The number and percentage of patients with TEAEs using MedDRA Preferred Term nested within System Organ Class	Treatment period
The number and percentage of patients with TEAEs using MedDRA Preferred Term (without regard to System Organ Class)	Treatment period
The number and percentage of patients with TEAEs by maximum severity using MedDRA Preferred Term	Treatment period
The number and percentage of patients with TEAEs using MedDRA Preferred Term for the common TEAEs (occurring in 2% [before rounding] of treated patients)	Treatment period
The number and percentage of patients who experienced an SAE (including deaths and SAEs temporally associated or preceding deaths) during the treatment period using MedDRA Preferred Term nested within System Organ Class	Treatment period
A listing of SAEs	Study period
A listing of deaths	Study period
The number and percentage of patients who permanently discontinued from study treatment due to an adverse event (including adverse events that led to death) during the treatment period using MedDRA Preferred Term nested within System Organ Class	Treatment period
The number and percentage of patients who temporarily interrupted study treatment due to an adverse event during the treatment period using MedDRA Preferred Term nested within System Organ Class	Treatment period
Number and percentage of patients with TEAEs using MedDRA Preferred Term nested within pre-defined event clusters	Treatment period

Abbreviations: AE = adverse event; MedDRA = Medical Dictionary for Regulatory Activities; SAE = serious adverse event; TEAE = treatment-emergent adverse event.

6.10.3.2. Deaths, Other Serious Adverse Events, and Temporary Study Drug Interruptions

The planned table and listing for SAEs are included in the previous section. A listing of deaths, regardless of when they occurred during the study, will also be provided as specified in the previous section.

A summary of temporary interruptions of study drug will also be provided, showing the number of patients who reported at least 1 temporary interruption and the number of temporary interruptions per patient who had an interruption. Further, the duration of each temporary interruption (in days), the cumulative duration of dose interruption (in days) using basic descriptive statistics (n, mean, standard deviation [SD], minimum, first quartile, median, third quartile, and maximum), and the reason for dose interruption will be provided.

6.10.3.3. Clinical Laboratory Evaluation

The planned summaries for clinical laboratory evaluations are provided in Table JAJA.6.7 and are described more fully in compound-level safety standards and in the laboratory-related PhUSE white papers (Analyses and Displays Associated with Measures of Central Tendency Focus on Vital Sign, Electrocardiogram, and Laboratory Analyte Measurements in Phase 2-4 Clinical Trials and Integrated Submission Documents [PhUSE 2013] and Analyses and Displays Associated with Outliers or Shifts from Normal to Abnormal: Focus on Vital Signs, Electrocardiogram, and Laboratory Analyte Measurements in Phase 2-4 Clinical Trials and Integrated Summary Documents [PhUSE 2015]).

For the categorical laboratory analyses (shift and treatment-emergent), the treatment period is used as the analysis period. For the continuous laboratory analyses (e.g., change from baseline by time point), the study period is used as the analysis period. Refer to the compound-level safety standards for more analysis details and the definition of treatment-emergent abnormal result for a laboratory analyte.

Table JAJA.6.7. Summary Tables Related to Clinical Laboratory Evaluations

Analysis	Analysis Period
Box plots for observed values	Study period
Box plots for change values	Study period
Tables with the number and percentage of patients who shift from normal/high to low (i.e., treatment-emergent low) and the number and percentage of patients who shift from normal/low to high (i.e., treatment-emergent high)	Treatment period
Listing of abnormal findings for laboratory analyte measurements, including qualitative measures	Study period

6.10.3.4. Vital Signs and Other Physical Findings

The planned summaries for vital signs (systolic blood pressure [BP], diastolic BP, pulse, weight, BMI, temperature) are provided in Table JAJA.6.8 and are described more fully in compound-level safety standards and in the vitals-related PhUSE white papers (Analyses and Displays Associated with Measures of Central Tendency Focus on Vital Sign, Electrocardiogram, and Laboratory Analyte Measurements in Phase 2-4 Clinical Trials and Integrated Submission

Documents [PhUSE 2013] and Analyses and Displays Associated with Outliers or Shifts from Normal to Abnormal: Focus on Vital Signs, Electrocardiogram, and Laboratory Analyte Measurements in Phase 2-4 Clinical Trials and Integrated Summary Documents [PhUSE 2015]).

For the treatment-emergent categorical analyses, the treatment period is used as the analysis period. For the continuous analyses (e.g., change from baseline), the study period is used as the analysis period.

Table JAJA.6.8. Summary Tables Related to Vital Signs

Analysis	Analysis Period
Box plots for observed values	Study period
Box plots for change values	Study period
Tables with the number and percentage of patients who shift from normal/high to low (i.e., treatment-emergent low) and the number and percentage of patients who shift from normal/low to high (i.e., treatment-emergent high); the limits are defined in the compound-level safety standards and are based on literature.	Treatment period

6.10.3.5. Special Safety Topics

6.10.3.5.1. Abnormal Hepatic Tests

Hepatic labs include alanine aminotransferase (ALT), aspartate aminotransferase (AST), total bilirubin (TBL) and serum alkaline phosphatase (ALP). When criteria are met for hepatic evaluations, investigators will complete a follow-up hepatic safety eCRF. The planned summaries are provided in Table JAJA.6.9. The analysis details are described more fully in compound-level safety standards.

Table JAJA.6.9. Summary Tables Related to Hepatic Safety

Analysis	Analysis Period
ALT and AST: The percentages of patients with a measurement greater than or equal to 3 times (3X), 5 times (5X), and 10 times (10X) the central lab upper limit of normal (ULN) during the treatment period for all patients with a post-baseline value and for subsets based on various levels of baseline value.	Treatment period
TBL: The percentages of patients with a measurement greater than or equal to 2 times (2X) the central lab ULN during the treatment period will be summarized for all patients with a post-baseline value and for subsets based on various levels of baseline value.	Treatment period
ALP: The percentages of patients with a measurement greater than or equal to 1.5 times (1.5X) the central lab ULN during the treatment period will be summarized for all patients with a post-baseline value and for subsets based on various levels of baseline value.	Treatment period
Plot of maximum post-baseline ALT vs. maximum post-baseline total bilirubin.	Study period
Patient profiles including demographics, disposition, information collected on the hepatic-safety CRF (where applicable) and a display of study drug exposure, adverse events, medications, blood pressure, heart rate, and the liver-related measurements over time will be provided for patients with information collected on the hepatic-safety CRF and any additional patients meeting ALT or AST measurement greater than or equal to 5X ULN (on a single measurement) or ALP measurement greater than or equal to 2X ULN (on a single measurement).	Study period

Summary Tables Related to Hepatic Safety

Analysis	Analysis Period
A listing of the information collected on the hepatic-safety CRF will be generated. In addition, a graphical patient profile including demographics, disposition, and a display of study drug exposure, adverse events, medications, and the liver-related measurements over time will be provided for these patients and any additional patients meeting ALT or AST measurement greater than or equal to 5× ULN (on a single measurement) or ALP measurement greater than or equal to 2× ULN (on a single measurement). The review for these patients includes an assessment of the time course of elevations of ALT, AST, ALP, TBL, gamma-glutamyl transferase (GGT) levels, temporal association with potential causes, and events such as nausea, vomiting, anorexia, abdominal pain, or fatigue.	Study period

Abbreviations: ALP = alkaline phosphatase; ALT = alanine aminotransferase; AST = aspartate aminotransferase; CRF = clinical report form; TBL = total bilirubin.

6.10.3.5.2. Hematologic Changes

Hematologic changes will be assessed through analysis of hemoglobin, white blood cell count, absolute neutrophil count, lymphocyte count, and platelet count. The planned summaries are provided in Table JAJA.6.10 and are described more fully in compound-level safety standards.

Table JAJA.6.10. Summary Tables Related to Hematologic Changes

Analysis	Analysis Period
Shift tables showing the number and percentage of patients based on baseline to maximum during the treatment period will be created, with baseline depicted by the most extreme CTCAE grade during the baseline period. With each shift table, a summary displaying the number and percentage of patients who decreased, increased, or stayed the same in CTCAE grade will be presented.	Treatment period
The percentages of patients with treatment-emergent laboratory abnormalities at any time during the treatment period will be summarized, based on any increase in postbaseline CTCAE grade, increase to CTCAE Grade 1 or above, Grade 2 or above, Grade 3 or above, and Grade 4.	Treatment period
The percentages of patients with treatment-emergent thrombocytosis will be summarized, defined as an increase in platelet count from a maximum baseline value HBB billion/L to any postbaseline value >600 billion/L. Increase in platelet count from a ‘)’ ‘ < <) ‘ <) FBB)’ Ak))‘) ’ ‘ < <) ‘ <)SFBB)’ Ak))’ < also presented.	Treatment period
Listing of patients with treatment-emergent thrombocytosis may be provided.	Study period

Abbreviation: CTCAE = Common Terminology Criteria for Adverse Events.

6.10.3.5.3. Lipid Effects

Lipid effects will be assessed through analysis of elevated total cholesterol, elevated low-density lipoprotein (LDL) cholesterol, decreased and increased high-density lipoprotein (HDL) cholesterol, and elevated triglycerides as described in Section 6.10.3.3. Planned summaries are provided in Table JAJA.6.11. In addition, the summary of treatment-emergent potential hyperlipidemia will be provided with more analysis details provided in the compound-level safety standards.

Table JAJA.6.11. Summary Tables Related to Lipid Effects

Analysis	Analysis Period
Treatment-emergent laboratory abnormalities related to elevated total cholesterol, elevated triglycerides, elevated LDL cholesterol, and decreased and increased HDL cholesterol occurring at any time during the treatment period will be tabulated. Shift tables will show the number and percentage of patients based on baseline to the least desirable category during the treatment period, with baseline depicted by the least desirable category during the baseline period.	Treatment period
The percentages of patients with treatment-emergent potential hyperlipidemia will be summarized using a predefined MedDRA list of PTs that is a subset of the narrow scope os)) <)l <αc q])r l p) c α< ‘ >))	Treatment period

Abbreviations: MedDRA = Medical Dictionary for Regulatory Activities; PTs = Preferred Terms; SMQ = Standardized MedDRA Query.

6.10.3.5.4. Renal Function Effects

Effects on renal function will be assessed through analysis of elevated creatinine. Renal events will be identified using terms from the acute renal failure Standardized Medical Dictionary for Regulatory Activities Query (SMQ). The planned summaries are provided in Table JAJA.6.12 and are described more fully in compound-level safety standards.

Table JAJA.6.12. Summary Tables Related to Effects on Renal Function

Analysis	Analysis Period
Shift tables showing the number and percentage of patients based on baseline to maximum during the treatment period will be created, with baseline depicted by the most extreme CTCAE grade during the baseline period. With each shift table, a summary displaying the number and percentage of patients who decreased, increased, or stayed the same in CTCAE grade will be presented.	Treatment period
The percentages of patients with treatment-emergent laboratory abnormalities related to elevated creatinine at any time during the treatment period will be summarized, based on any increase in postbaseline CTCAE grade, increase to CTCAE Grade 1 or above, Grade 2 or above, Grade 3 or above, and Grade 4.	Treatment period

Abbreviation: CTCAE = Common Terminology Criteria for Adverse Events.

6.10.3.5.5. Elevations in Creatine Phosphokinase (CPK)

The planned summaries are provided in Table JAJA.6.13 and are described more fully in compound-level safety standards.

Table JAJA.6.13. Summary Tables Related to CPK

Analysis	Analysis Period
Shift tables showing the number and percentage of patients based on baseline to maximum during the treatment period will be created, with baseline depicted by the most extreme CTCAE grade during the baseline period. With each shift table, a summary displaying the number and percentage of patients who decreased, increased, or stayed the same in CTCAE grade will be presented.	Treatment period

Summary Tables Related to CPK

Analysis	Analysis Period
The percentages of patients with treatment-emergent laboratory abnormalities related to elevated CPK at any time during the treatment period will be summarized, based on any increase in postbaseline CTCAE grade, increase to CTCAE Grade 1 or above, Grade 2 or above, Grade 3 or above, and Grade 4.	Treatment period
Treatment-emergent adverse events potentially related to muscle symptoms may also be analyzed based on reported AEs. The Muscle Symptoms special search category is a predefined MedDRA search criteria list that contains the narrow scope terms from the Rhabdomyolysis/Myopathy SMQ plus selected terms from the Musculoskeletal SOC.	Treatment period

Abbreviations: AEs = adverse events; CTCAE = Common Terminology Criteria for Adverse Events; MedDRA = Medical Dictionary for Regulatory Activities; SMQ = Standardized MedDRA Query; SOC = System Organ Class.

6.10.3.5.6. Infections

Infections will be defined using all the Preferred Terms (PTs) from the Infections and Infestations System Organ Class (SOC) as defined in the Medical Dictionary for Regulatory Activities (MedDRA). Opportunistic infections will be adjudicated by an independent, external adjudication committee and will be summarized by PT nested under infection pathogen (or presentation of pathogens). Events will be ordered by decreasing frequency of pathogen nested under its pathogen species (mycobacteria, bacteria, fungi, virus, parasites, other).

The planned summaries are provided in Table JAJA.6.14 and are described more fully in compound-level safety standards.

Infections will be reported separately for Studies JAJA and JAJD.

Table JAJA.6.14. Summary Tables Related to Infections

Analysis	Analysis Period
The number and percentage of patients with treatment-emergent infections, serious infections, opportunistic infections and infections resulting in permanent study drug discontinuation using MedDRA PTs	Treatment period
The number and percentage of patients with TEAEs of infections by maximum severity using MedDRA PTs	Treatment period
The exposure-adjusted incidence rate (EAIR) and 95% confidence intervals by the overall period for infections of special interest (i.e., serious infections, herpes simplex, and herpes zoster)	Treatment period
Listing of patients experiencing TEAE infections will be provided. The listing will include patient demographics, treatment group, treatment start and stop dates, infectious PT event, event start and stop dates, total leukocytes, total lymphocytes, absolute neutrophils, event seriousness, and event outcome.	Treatment period
The number and percentage of patients with opportunistic infection, as positively adjudicated, will be summarized by treatment group.	Treatment period
A listing of the events sent for opportunistic infection adjudication will be provided to include data concerning the MedDRA PT related to the event, lab results (for HBV reactivation only), the seriousness of the event, and the event outcome, along with the adjudicated result.	Study period

Summary Tables Related to Infections

Analysis	Analysis Period
A summary table of herpes zoster will be provided, including event maximum severity, seriousness, whether resulting in temporary study drug interruption, whether resulting in study drug discontinuation, whether treated with antiviral medication, and event outcome. The incidence rate adjusted for observation time will also be provided.	Treatment period
A summary table of herpes simplex will be provided, including event maximum severity, seriousness, whether resulting in temporary study drug interruption, whether resulting in study drug discontinuation, whether treated with antiviral medication, and event outcome. The incidence rate adjusted for observation time will also be provided.	Treatment period
A listing of patients with detectable HBV DNA	Study period
HBV DNA status (not detectable, detectable but not quantifiable [i.e., < lower limit of detection (LLOD)], quantifiable [i.e.8) k k n c ≥5) ‘ <α)’)‘ „‘ ’ <)’ ‘ < <)g a u) serology status	Treatment period

Abbreviations: HBV = hepatitis B virus; MedDRA = Medical Dictionary for Regulatory Activities; PT = Preferred Term; TEAE = treatment-emergent adverse event.

6.10.3.5.7. Major Adverse Cardiovascular Events (MACE) and Other Cardiovascular Events

Major adverse cardiovascular events and other CV events will be adjudicated by an independent, external adjudication committee. All confirmed events after adjudication will be used for the analysis.

The planned summaries are provided in Table JAJA.6.15 and are described more fully in compound-level safety standards.

Table JAJA.6.15. Summary Tables Related to MACE and Other Cardiovascular Events

Analysis	Analysis Period
The number and percentage of patients with MACE, other cardiovascular events, noncardiovascular death, and all-cause death, as positively adjudicated, will be summarized by treatment group based on the categories and subcategories as defined in the compound-level safety standards.	Treatment period
A listing of the events sent for cardiovascular adjudication will be provided to include data concerning the MedDRA PT related to the event, the seriousness of the event, and the event outcome, along with the adjudicated result.	Study period
EAIRs and 95% confidence intervals by the overall period for positively adjudicated MACE	Treatment period

Abbreviations: EAIRs = exposure-adjusted incidence rates; MACE = major adverse cardiovascular events; MedDRA = Medical Dictionary for Regulatory Activities; PT = Preferred Term.

6.10.3.5.8. Venous and Pulmonary Artery Thromboembolic (VTE) Events

Venous thromboembolic events will be adjudicated by an independent, external adjudication committee. Venous and pulmonary artery thromboembolic events will be classified as DVT (excluding nonsuperficial below-knee thrombosis), PE, or other peripheral venous thrombosis (including nonsuperficial below-knee thrombosis). All confirmed events after adjudication will be used for the analysis.

The planned summaries are provided in Table JAJA.6.16 and are described more fully in compound-level safety standards.

Table JAJA.6.16. Summary Tables Related to VTE Events

Analysis	Analysis Period
The number and percentage of patients with a VTE, DVT/PE, DVT, PE, and other peripheral venous thrombosis, as positively adjudicated, will be summarized separately and by treatment group.	Treatment period
A listing of the VTE events sent for adjudication will be provided to include data concerning the MedDRA PT related to the event, the seriousness of the event, and the event outcome, along with the adjudicated result.	Study period
EAIRs and 95% confidence intervals by the overall period for positively adjudicated VTE	Treatment period

Abbreviations: EAIRs = exposure-adjusted incidence rates; DVT = deep vein thrombosis; MedDRA = Medical Dictionary for Regulatory Activities; PE = pulmonary embolism; PT = Preferred Term; VTE = venous thromboembolism that consists of DVT, including nonsuperficial below-knee thrombosis, and PE.

6.10.3.5.9. Arterial Thromboembolic Events (ATEs)

Arterial thromboembolic events will be adjudicated by an independent, external adjudication committee.

The planned summaries are provided in Table JAJA.6.17 and are described more fully in compound-level safety standards.

Table JAJA.6.17. Summary Tables Related to ATE Events

Analysis	Analysis Period
The number and percentage of patients with an ATE, as positively adjudicated, will be summarized by treatment group.	Treatment period
A listing of the ATE events sent for adjudication will be provided to include data concerning the MedDRA PT related to the event, the seriousness of the event, and the event outcome, along with the adjudicated result.	Study period
EAIRs and 95% confidence intervals by the overall period for positively adjudicated ATE.	Treatment period

Abbreviations: ATE = arterial thromboembolism; EAIRs = exposure-adjusted incidence rates; MedDRA = Medical Dictionary for Regulatory Activities; PT = Preferred Term.

6.10.3.5.10. Malignancies

Malignancies will be identified using terms from the Malignant Tumors SMQ. Malignancies excluding nonmelanoma skin cancers (NMSC) and NMSC will be reported separately. All the cases identified by the malignant tumors SMQ will be assessed through medical review to determine *confirmed* NMSC cases.

The planned summaries are provided in Table JAJA.6.18 and are described more fully in compound-level safety standards.

Table JAJA.6.18. Summary Tables Related to Malignancies

Analysis	Analysis Period
The number and percentage of patients with treatment-emergent malignancies excluding NMSC and NMSC will be summarized by treatment group.	Treatment period
Listing of all malignancy cases, with an NMSC flag.	Study period
EAIRs and 95% confidence intervals by the overall period for treatment-emergent malignancies excluding NMSC and NMSC. Of note, investigator-reported time will be used for the EAIR calculation.	Treatment period

Abbreviations: EAIRs = exposure-adjusted incidence rates; NMSC = nonmelanoma skin cancer.

6.10.3.5.11. Allergic Reactions/Hypersensitivities

A search for relevant events related to allergic reaction and hypersensitivity will be performed using the following SMQs:

Anaphylactic Reaction SMQ

Hypersensitivity SMQ

Angioedema SMQ

Events that satisfy the queries will be listed, by temporal order within patient identification (ID) and will include SOC, PT, SMQ event categorization including detail on the scope, reported AE term, AE onset and end dates, severity, seriousness, outcome, etc.

The summaries described in Table JAJA.6.19 will be created if there are sufficient numbers of events to warrant further examination beyond the listing specified above. More details on the event definition are included in the compound-level safety standards.

Table JAJA.6.19. Summary Tables Related to Allergic Reactions/Hypersensitivities

Analysis	Analysis Period
The number and percentage of patients with treatment-emergent allergic reactions/hypersensitivities will be summarized by treatment.	Treatment period

6.10.3.5.12. Lower Gastrointestinal Perforations

Potential lower gastrointestinal (GI) perforations will be identified using terms from the GI perforations SMQ. Potential GI perforations identified by the SMQ search will be provided as a listing for internal review by the medical safety team. Each case will be assessed to determine whether it is lower GI perforation. All confirmed events after medical review will be used for

the analysis. The planned summaries are provided in Table JAJA.6.20 and are described more fully in compound-level safety standards.

Table JAJA.6.20. Summary Tables Related to Gastrointestinal Perforations

Analysis	Analysis Period
The number and percentage of patients with treatment-emergent lower gastrointestinal perforations will be summarized by treatment.	Treatment period

6.10.3.6. Criteria for Notable Patients

o‘ <) ‘ ‘ <))’<) α<α))‘) ‘ <))< < < „<),,< ‘) ‘ ’ <)< < >)r <<) compound level safety standards for list of criteria.

6.11. Efficacy and Health Outcomes Analyses (Exploratory)

The general methods used to summarize efficacy and health outcomes, including the definition of baseline value for assessments, are described in Section 6.1.

Table JAJA.6.21 includes the descriptions and derivations of the efficacy and health outcomes measures.

Table JAJA.6.22 provides the detailed analyses including analysis type, method and imputation, population, time point, and comparisons for efficacy and health outcomes measures. The treatment comparisons will be conducted up to 52 weeks, in which the rescue treatment is not allowed, providing the most clean treatment comparisons.

Additional long-term analyses may be explored and the analysis plan will be documented in a separate document.

Table JAJA.6.21. Description and Derivation of Efficacy and Health Outcomes Measures

Measure	Description	Variable	Derivation/Comment	Processing of Missing Components
Tender joint count (TJC28)	Number of tender joints in the 28 joints that are evaluated	TJC28	Assessed based on 28 joints	When there are nonevaluable joints, tender joints will be prorated to 28 joints. Handling of other medical procedures/ conditions will follow compound-level standards.
		TJC28 change from baseline	Calculated as: observed TJC28 baseline TJC28	Missing if baseline or observed value is missing

Description and Derivation of Efficacy and Health Outcomes Measures

Measure	Description	Variable	Derivation/Comment	Processing of Missing Components
Swollen joint count (SJC28)	Number of swollen joints in the 28 joints that are evaluated	SJC28	Assessed based on 28 joints	When there are nonevaluable joints, swollen joints will be prorated to 28 joints. Handling of other medical procedures/ conditions will follow compound-level standards.
		SJC28 change from baseline	Calculated as: observed SJC28 baseline SJC28	Missing if baseline or observed value is missing
o‘ <) Global Assessment of Disease Activity (PtGA)	PtGA is measured using 0 to 100 mm Visual Analog Scale [VAS]. For analysis the value will be divided by 10.	PtGA	A single item score reported by patients.	Single item, missing if missing.
		PtGA change from baseline	Calculated as: observed PtGA baseline PtGA	Missing if baseline or observed value is missing.
o „ ‘) Global Assessment of Disease Activity (PhGA)	PhGA is measured using 0 to 100 mm Visual Analog Scale [VAS]. For analysis the value will be divided by 10.	PhGA	A single item score reported by physician.	Single item, missing if missing.
		PhGA change from baseline	Calculated as: observed PhGA baseline PhGA	Missing if baseline or observed value is missing.

Description and Derivation of Efficacy and Health Outcomes Measures

Measure	Description	Variable	Derivation/Comment	Processing of Missing Components
Health Assessment Questionnaire Disability Index (HAQ-DI)	P‘ <) assessment of physical function is measured by HAQ-DI, which consists of 24 questions referring to 8 domains with a range of possible scores (0-3): dressing and grooming, arising, eating, walking, hygiene, reach, grip, and activities	HAQ-DI	Will follow the scoring manual	Missing when there are H) domains completed
		HAQ-DI change from baseline	Calculated as: observed HAQ-DI baseline HAQ-DI	Missing if baseline or observed value is missing
		HAQ-DI improvement from baseline	Calculated as: baseline HAQ-DI observed HAQ-DI	Missing if baseline or observed value is missing
Rheumatoid Arthritis Impact of Disease (RAID) pain	RAID Pain is from the RAID pain domain.	RAID pain score	A single item score from the numeric rating scale (0-10)	Single item, missing if missing
		RAID pain score change from baseline	Calculated as: observed RAID pain score baseline RAID pain score	Missing if baseline or observed value is missing
Acute phase reactant as measured by high-sensitivity C-reactive protein (hsCRP)	hsCRP is a lab measure.	hsCRP	A collected lab value	Single item, missing if missing
		hsCRP change from baseline	Calculated as: observed hsCRP baseline hsCRP	Missing if baseline or observed value is missing
American College of Rheumatology (ACR) response	ACR response will be defined based on 7 components: 1. TJC28 2. SJC28 3. RAID pain 4. PtGA 5. PhGA 6. HAQ-DI 7. hsCRP	ACR20	At least 20% improvement from baseline in the following ACR Core Set values: TJC28, SJC28, and an improvement of at least 20% in at least 3 of the following 5 assessments: RAID pain, PtGA, PhGA, HAQ-DI, hsCRP	ACR20 response calculation will follow compound-level standards.
		ACR50	Same as above with at least 50% improvement.	Same as above
		ACR70	Same as above with at least 70% improvement.	Same as above

Description and Derivation of Efficacy and Health Outcomes Measures

Measure	Description	Variable	Derivation/Comment	Processing of Missing Components
Disease activity score	The Disease Activity Score (DAS)28 is a measure of disease activity based on 28 joints that consists of a composite numeric score of the following variables: TJC28, SJC28, hsCRP, and PtGA.	DAS28-hsCRP value	Calculated as: $0.56(TJC28)^{1/2} + 0.28(SJC28)^{1/2} + 0.36(\ln(hsCRP+1)) + 0.014(VAS) + 0.96$ (Vander Cruyssen et al. 2005)	Missing if any of the component score is missing
		DAS28-hsCRP change from baseline	Calculated as: observed DAS28-hsCRP baseline DAS28-hsCRP	Missing if baseline or observed value is missing
		DAS28-hsCRP E≥2	A DAS28-hsCRP E≥2.	Missing if DAS28-hsCRP is missing
		DAS28-hsCRP <2.6	A DAS28-hsCRP <2.6.	Missing if DAS28-hsCRP is missing
Simplified Disease Activity Index	The Simplified Disease Activity Index (SDAI) is a tool for measurement of disease activity in RA that integrates measures of physical examination, acute phase response, patient self-assessment, and evaluator assessment.	SDAI value	Calculated as: TJC28 + SJC28 + hsCRP (0.1 to 10.0 mg/dL) + PtGA (0 to 10.0 cm) + PhGA (0 to 10.0 cm) (Aletaha and Smolen 2005)	Missing if any of the component score is missing
		SDAI change from baseline	Calculated as: observed SDAI baseline SDAI	Missing if baseline or observed value is missing
		SDAI 11	An SDAI 11.	Missing if SDAI is missing
		SDAI 3.3	An SDAI 3.3.	Missing if SDAI is missing
Clinical Disease Activity Index	The Clinical Disease Activity Index (CDAI) is a tool for measurement of disease activity in RA that integrates measures of physical examination, patient self-assessment, and evaluator assessment.	CDAI value	Calculated as: TJC28 + SJC28 + PtGA (0 to 10.0 cm) + PhGA (0 to 10.0 cm) (Aletaha and Smolen 2005)	Missing if any of the component score is missing
		CDAI change from baseline	Calculated as: observed CDAI baseline CDAI	Missing if baseline or observed value is missing
		CDAI 10	A CDAI 10	Missing if CDAI is missing
		CDAI 2.8	A CDAI 2.8.	Missing if CDAI is missing.

Description and Derivation of Efficacy and Health Outcomes Measures

Description and Derivation of Efficacy and Health Outcomes MeasuresMeasure	Description	Variable	Derivation/Comment	Processing of Missing Components
PROMIS Pain Interference score	The questionnaire consists of 8 questions and uses a 5-point Likert scale (1 = not at all; 5 = very much).	Pain interference score	Scores are totaled and higher scores indicate more interference.	Calculation will be done by the HealthMeasures Scoring Service, powered by Assessment Center℠.
		Pain interference change from baseline	Calculated as: observed pain interference baseline pain interference	Missing if baseline or observed value is missing
PROMIS Depression score	The questionnaire consists of 8 questions and uses a 5-point Likert scale (1 = never; 5 = always).	Depression score	Scores are totaled and higher scores indicate more depressive symptoms.	Calculation will be done by the HealthMeasures Scoring Service, powered by Assessment Center℠.
		Depression score change from baseline	Calculated as: observed depression score baseline depression score	Missing if baseline or observed value is missing
PROMIS Cognitive Function score	The questionnaire consists of 8 questions and uses a 5-point Likert scale 1 = very often; 5 = never).	Cognitive function score	Scores are totaled and higher scores indicate better cognitive ability.	Calculation will be done by the HealthMeasures Scoring Service, powered by Assessment Center℠.
		Cognitive function score change from baseline	Calculated as: observed cognitive function score baseline cognitive function score	Missing if baseline or observed value is missing

Description and Derivation of Efficacy and Health Outcomes Measures

Description and Derivation of Efficacy and Health Outcomes MeasuresMeasure	Description	Variable	Derivation/Comment	Processing of Missing Components
PROMIS Satisfaction with Social Roles/Activities score	The questionnaire consists of 8 questions and uses a 5-point Likert scale (1 = not at all; 5 = very much).	Satisfaction score	Scores are totaled and higher scores indicate more satisfaction.	Calculation will be done by the HealthMeasures Scoring Service, powered by Assessment Center℠.
		Satisfaction score change from baseline	Calculated as: observed satisfaction score baseline satisfaction score	Missing if baseline or observed value is missing
PROMIS Sleep Disturbance score	The questionnaire consists of 8 questions and uses a 5-point Likert scale (1=very good/very much; 5=very poor/not at all).	Sleep disturbance score	Scores are totaled and higher scores indicate worse sleep.	Calculation will be done by the HealthMeasures Scoring Service, powered by Assessment Center℠.
		Sleep disturbance score change from baseline	Calculated as: observed sleep disturbance score baseline sleep disturbance score	Missing if baseline or observed value is missing

Description and Derivation of Efficacy and Health Outcomes Measures

Measure	Description	Variable	Derivation/Comment	Processing of Missing Components
RAID (Gossec et al. 2011)	The RAID is a composite instrument that is used to assess the impact of RA on 7 domains of importance to patients: fatigue, sleep disturbance, pain, physical well-being, function, psychological well-being, and coping using a multiple 0-10 NRS.	Fatigue score	A single item score from the numeric rating scale (0-10).	Single item, missing if missing.
		Fatigue score change from baseline	Calculated as: observed fatigue score baseline fatigue score	Missing if baseline or observed value is missing
		Sleep disturbance score	A single item score from the numeric rating scale (0-10)	Single item, missing if missing
		Sleep disturbance score change from baseline	Calculated as: observed sleep disturbance score baseline sleep disturbance score	Missing if baseline or observed value is missing
		Physical well-being score	A single item score from the numeric rating scale (0-10)	Single item, missing if missing
		Physical well-being score change from baseline	Calculated as: observed physical well-being score baseline physical well-being score	Missing if baseline or observed value is missing
		Function score	A single item score from the numeric rating scale (0-10)	Single item, missing if missing
		Function score change from baseline	Calculated as: observed function score baseline function score	Missing if baseline or observed value is missing
		Psychological well-being score	A single item score from the numeric rating scale (0-10)	Single item, missing if missing
		Psychological well-being score change from baseline	Calculated as: observed psychological well-being score baseline psychological well-being score	Missing if baseline or observed value is missing

Description and Derivation of Efficacy and Health Outcomes Measures

Measure	Description	Variable	Derivation/Comment	Processing of Missing Components
		Coping score	A single item score from the numeric rating scale (0-10)	Single item, missing if missing.
		Coping score change from baseline	Calculated as: observed coping score baseline coping score	Missing if baseline or observed value is missing.
		Pain Score	Included at beginning of this table	
		Pain score change from baseline	Included at beginning of this table	
		RAID total score	Calculated as: (pain x 0.21 + function x 0.16 + fatigue x 0.15 + physical well-being x 0.12 + sleep x 0.12 + emotional well-being x 0.12 + coping x 0.12. RAID total score is 0-10 where higher values indicate worse status.	If 1 of the 7 domain scores is missing, impute as follows: a. calculate the mean value of the 6 other nonmissing values; b. impute the missing domain score using this value; c. then calculate the total score. If 2 domains missing, RAID total score will be missing.

Abbreviations: NRS = numeric rating scale; PROMIS = Patient-Reported Outcome Measurement Information System; RA = rheumatoid arthritis; RAID = Rheumatoid Arthritis Impact of Disease.

Table JAJA.6.22. Description of Efficacy and Health Outcomes Analyses

Measure	Variable	Analysis Method (Section 6.1.3)	Population (Section 6.1.1)	Comparison/Time Period	Missing data imputation
Tender joint count (TJC28)	TJC28 change from baseline	Descriptive summary	ITT	No comparison/all time points where data are collected	None
		MMRM	ITT	Each baricitinib dose vs. TNFi/up to 52 weeks	None
Swollen joint count (SJC28)	SJC28 change from baseline	Descriptive summary	ITT	No comparison/all time points where data are collected	None
		MMRM	ITT	Each baricitinib dose vs. TNFi/up to 52 weeks	None
o‘ <)f ’‘) Assessment of Disease Activity (PtGA)	PtGA change from baseline	Descriptive summary	ITT	No comparison/all time points where data are collected	None
		MMRM	ITT	Each baricitinib dose vs. TNFi/up to 52 weeks	None
o „‘)f ’‘) Assessment of Disease Activity (PhGA)	PhGA change from baseline	Descriptive summary	ITT	No comparison/all time points where data are collected	None
		MMRM	ITT	Each baricitinib dose vs. TNFi/up to 52 weeks	None
Health Assessment Questionnaire Disability Index (HAQ-DI)	HAQ-DI change from baseline	Descriptive summary	ITT	No comparison/all time points where data are collected	None
		MMRM	ITT	Each baricitinib dose vs. TNFi/up to 52 weeks	None
	HAQ-DI improvement 0.22	Descriptive statistics	ITT	No comparison/all time points where data are collected	None
		Logistic regression	ITT	Each baricitinib dose vs. TNFi/up to 52 weeks	NRI
	HAQ-DI improvement 0.3	Descriptive statistics	ITT	No comparison/all time points where data are collected	None

Description of Efficacy and Health Outcomes Analyses

Measure	Variable	Analysis Method (Section 6.1.3)	Population (Section 6.1.1)	Comparison/Time Period	Missing data imputation
		Logistic regression	ITT	Each baricitinib dose vs. TNFi/up to 52 weeks	NRI
Acute phase reactant as measured by high-sensitivity C-reactive protein (hsCRP)	hsCRP change from baseline	Descriptive summary	ITT	No comparison/all time points where data are collected	None
		MMRM	ITT	Each baricitinib dose vs. TNFi/up to 52 weeks	None
ACR response	ACR20	Descriptive statistics	ITT	No comparison/all time points where data are collected	None
		Logistic regression	ITT	Each baricitinib dose vs. TNFi/up to 52 weeks	NRI
	ACR50	Descriptive statistics	ITT	No comparison/all time points where data are collected	None
		Logistic regression	ITT	Each baricitinib dose vs. TNFi/up to 52 weeks	NRI
	ACR70	Descriptive statistics	ITT	No comparison/all time points where data are collected	None
		Logistic regression	ITT	Each baricitinib dose vs. TNFi/up to 52 weeks	NRI
Disease Activity Score (DAS28)	DAS28-hsCRP change from baseline	Descriptive summary	ITT	No comparison/all time points where data are collected	None
		MMRM	ITT	Each baricitinib dose vs. TNFi/up to 52 weeks	None
	DAS28-hsCRP E≥2	Descriptive statistics	ITT	No comparison/all time points where data are collected	None
		Logistic regression	ITT	Each baricitinib dose vs. TNFi/up to 52 weeks	NRI

Description of Efficacy and Health Outcomes Analyses

Measure	Variable	Analysis Method (Section 6.1.3)	Population (Section 6.1.1)	Comparison/Time Period	Missing data imputation
	DAS28-hsCRP <2.6	Descriptive statistics	ITT	No comparison/all time points where data are collected	None
		Logistic regression	ITT	Each baricitinib dose vs. TNFi/up to 52 weeks	NRI
Simplified Disease Activity Index (SDAI)	SDAI value	Descriptive summary	ITT	No comparison/all time points where data are collected	None
	SDAI change from baseline	MMRM	ITT	Each baricitinib dose vs. TNFi/up to 52 weeks	None
	SDAI 11	Descriptive statistics	ITT	No comparison/all time points where data are collected	None
		Logistic regression	ITT	Each baricitinib dose vs. TNFi/up to 52 weeks	NRI
	SDAI 3.3	Descriptive statistics	ITT	No comparison/all time points where data are collected	None
		Logistic regression	ITT	Each baricitinib dose vs. TNFi/up to 52 weeks	NRI
Clinical Disease Activity Index (CDAI)	CDAI change from baseline	Descriptive summary	ITT	No comparison/all time points where data are collected	None
		MMRM	ITT	Each baricitinib dose vs. TNFi/up to 52 weeks	None
	CDAI 10	Descriptive statistics	ITT	No comparison/all time points where data are collected	None
		Logistic regression	ITT	Each baricitinib dose vs. TNFi/up to 52 weeks	NRI
	CDAI 2.8	Descriptive statistics	ITT	No comparison/all time points where data are collected	None
		Logistic regression	ITT	Each baricitinib dose vs. TNFi/up to 52 weeks	NRI

Description of Efficacy and Health Outcomes Analyses

Measure	Variable	Analysis Method (Section 6.1.3)	Population (Section 6.1.1)	Comparison/Time Period	Missing data imputation
PROMIS Pain Interference score	Pain interference change from baseline	Descriptive summary	ITT patients with PROMIS data	No comparison/all time points where data are collected	None
		MMRM	ITT patients with PROMIS data	Each baricitinib dose vs. TNFi/up to 52 weeks	None
PROMIS Depression score	Depression score change from baseline	Descriptive summary	ITT patients with PROMIS data	No comparison/all time points where data are collected	None
		MMRM	ITT patients with PROMIS data	Each baricitinib dose vs. TNFi/up to 52 weeks	None
PROMIS Cognitive Function score	Cognitive function score change from baseline	Descriptive summary	ITT patients with PROMIS data	No comparison/all time points where data are collected	None
		MMRM	ITT patients with PROMIS data	Each baricitinib dose vs. TNFi/up to 52 weeks	None
PROMIS Satisfaction with Social Roles/Activities score	Satisfaction score change from baseline	Descriptive summary	ITT patients with PROMIS data	No comparison/all time points where data are collected	None
		MMRM	ITT patients with PROMIS data	Each baricitinib dose vs. TNFi/up to 52 weeks	None

Description of Efficacy and Health Outcomes Analyses

Measure	Variable	Analysis Method (Section 6.1.3)	Population (Section 6.1.1)	Comparison/Time Period	Missing data imputation
PROMIS Sleep Disturbance score	Sleep disturbance score change from baseline	Descriptive summary	ITT patients with PROMIS data	No comparison/all time points where data are collected	None
		MMRM	ITT patients with PROMIS data	Each baricitinib dose vs. TNFi/up to 52 weeks	None
RAID	Fatigue score change from baseline	Descriptive summary	ITT patients with PROMIS data	No comparison/all time points where data are collected	None
		MMRM	ITT patients with PROMIS data	Each baricitinib dose vs. TNFi/up to 52 weeks	None
	Sleep disturbance score change from baseline	Descriptive summary	ITT patients with PROMIS data	No comparison/all time points where data are collected	None
		MMRM	ITT patients with PROMIS data	Each baricitinib dose vs. TNFi/up to 52 weeks	None
	Physical well-being score change from baseline	Descriptive summary	ITT patients with PROMIS data	No comparison/all time points where data are collected	None
		MMRM	ITT patients with PROMIS data	Each baricitinib dose vs. TNFi/up to 52 weeks	None

Description of Efficacy and Health Outcomes Analyses

Measure	Variable	Analysis Method (Section 6.1.3)	Population (Section 6.1.1)	Comparison/Time Period	Missing data imputation
	Function score change from baseline	Descriptive summary	ITT patients with PROMIS data	No comparison/all time points where data are collected	None
		MMRM	ITT patients with PROMIS data	Each baricitinib dose vs. TNFi/up to 52 weeks	None
	Psychological well-being score change from baseline	Descriptive summary	ITT patients with PROMIS data	No comparison/all time points where data are collected	None
		MMRM	ITT patients with PROMIS data	Each baricitinib dose vs. TNFi/up to 52 weeks	None
	Coping score change from baseline	Descriptive summary	ITT patients with PROMIS data	No comparison/all time points where data are collected	None
		MMRM	ITT patients with PROMIS data	Each baricitinib dose vs. TNFi/up to 52 weeks	None
	Pain score change from baseline	Descriptive summary	ITT patients with PROMIS data	No comparison/all time points where data are collected	None
		MMRM	ITT patients with PROMIS data	Each baricitinib dose vs. TNFi/up to 52 weeks	None

Description of Efficacy and Health Outcomes Analyses

Measure	Variable	Analysis Method (Section 6.1.3)	Population (Section 6.1.1)	Comparison/Time Period	Missing data imputation
	RAID total score	Descriptive summary	ITT patients with PROMIS data	No comparison/all time points where data are collected	None
		MMRM	ITT patients with PROMIS data	Each baricitinib dose vs. TNFi/up to 52 weeks	None

Abbreviations: ACR = American College of Rheumatology; ACR20/50/70 = American College of Rheumatology 20%/50%/70% response rate; ITT = intent-to-treat; MMRM = mixed model for repeated measures; NRI = nonresponder imputation; PROMIS = Patient-Reported Outcome Measurement Information System; RAID = Rheumatoid Arthritis Impact of Disease; TNFi = tumor necrosis factor inhibitor.

6.12. Protocol Violations

Protocol deviations will be tracked by the clinical team, and their importance will be assessed by key team members during protocol deviation review meetings.

The categories and subcategories of IPDs, the source of identification for the deviation, and the statistical programming guidance for the CSR are included in the JAJA Trial Issue Management Plan document.

The number and percentage of patients having IPDs will be summarized within category and subcategory of deviation by treatment groups. Individual patient listings of IPDs will be provided.

6.13. Interim Analyses and Data Monitoring

A data monitoring committee (DMC) will oversee the conduct of this study in conjunction with Study JAJD. The DMC will consist of members external to Lilly who have rheumatology, CV, coagulation, or statistics expertise. This DMC will follow the rules defined in the DMC charter.

The data from Studies JAJA and JAJD will be evaluated at interim analyses in order to determine if the combined evidence is sufficient to meet the studies' primary objective with respect to VTE. CCI interim data analyses are planned to occur based on the cumulative number of patients with events of VTE observed in these CCI trials: CCI analyses to evaluate safety with no consideration of the studies' primary objective and CCI subsequent analyses to evaluate safety and whether the primary objective has been met. The first CCI analyses are planned to occur when approximately CCI of the planned total patients with VTE events are observed (across studies). The CCI analyses to evaluate the primary objective will occur when approximately CCI of the planned total across studies) and CCI of the planned total across studies) patients with VTE events are observed. Ad hoc interim analyses may be conducted to address regulatory requests.

CCI [redacted]

The studies can be stopped early at the CCI [redacted] interim analyses for success based on the primary safety analysis. CCI [redacted]

When the accumulation of events for VTE is lagging considerably behind, the primary analysis will be modified to an incidence rate difference analysis. See more details in Section 4.1.

If the study is not stopped by these rules or by the rules for early stopping for success, it is continued until the full planned number of patients with events is observed.

More details are given in the DMC charter.

6.14. Planned Exploratory Analyses

The planned exploratory analyses are described in Sections 6.10.3 and 6.10.3.6. Health Technology Assessment (HTA) toolkit analyses, which may be produced, will be documented in the supplemental SAP.

6.15. Annual Report Analyses

Annual report analyses, such as for the Development Safety Update Report (DSUR), will be documented in a separate document.

6.16. Clinical Trial Registry Analyses

Additional analyses will be performed for the purpose of fulfilling the Clinical Trial Registry (CTR) requirements.

Analyses provided for the CTR requirements include the following:

Summary of AEs, provided as a dataset which will be converted to an XML file. Both SAEs and 'Other' AEs are summarized: by treatment group, by MedDRA PT.

- An AE is considered 'Serious' whether or not it is a TEAE.
- An AE is considered in the 'Other' category if it is both a TEAE and is not serious. For each 'Serious' AE and 'Other' AE, for each term and treatment group, the following are provided:
    - the number of participants at risk of an event
    - the number of participants who experienced each event term
    - the number of events experienced

b <)) >b „‘ s ‘ >)< < < 8) n <)] d) ‘) „„))< <) than 5% of patients in every treatment group may not be included if a 5% threshold is chosen (5% is the minimum threshold).

AE reporting is consistent with other document disclosures for example, the CSR, manuscripts, and so forth.

Similar methods will be used to satisfy the European Clinical Trials Database (EudraCT) requirements.

7. Unblinding Plan

Refer to a separate blinding and unblinding plan document for details.

8. References

Aletaha D, Smolen J. The Simplified Disease Activity Index (SDAI) and the Clinical Disease Activity Index (CDAI): a review of their usefulness and validity in rheumatoid arthritis. *Clin Exp Rheumatol.* 2005;23(5 Suppl 39):S100-108.

Gossec L, Paternotte S, Aanerud GJ, Balanescu A, Boumpas DT, Carmona L, de Wit M, Dijkmans BA, Dougados M, Englbrecht M, Gogus F, Heiberg T, Hernandez C, Kirwan JR, Mola EM, Cerinic MM, Otsa K, Schett G, Scholte-Voshaar M, Sokka T, von Krause G, Wells GA, Kvien TK. Finalisation and validation of the rheumatoid arthritis impact of disease score, a patient-derived composite measure of impact of rheumatoid arthritis: a EULAR initiative. *Ann Rheum Dis.* 2011;70(6):935-942.

Haybittle JL. Repeated assessments of results in clinical trials of cancer treatment. *Br J Radiol.* 1971;44(526):793-797.

Jennison C, Turnbull BW. *Group sequential methods with applications to clinical trials*. New York: Chapman & Hall/CRC; 2000.

Lipkovich, I., Ratitch, B., & Mallinckrodt, C. H. (2020). Causal inference and estimands in clinical trials. *Statistics in Biopharmaceutical Research*, *12*(1), 54-67.

[PhUSE] Analyses and displays associated with measures of central tendency focus on vital sign, electrocardiogram, and laboratory analyte measurements in phase 2-4 clinical trials and integrated submission documents. Available at: https://www.phuse.eu/documents//working-groups/deliverables/analyses-displays-associated-with-measures-of-central-tendancy-version-10-10-oct-13-11816.pdf. Published October 10, 2013. Accessed May 5, 2020.

[PhUSE] Analyses and displays associated with outliers or shifts from normal to abnormal: focus on vital signs, electrocardiogram, and laboratory analyte measurements in Phase 2-4 clinical trials and integrated summary documents. Available at: https://www.phuse.eu/documents/working-groups/deliverables/analyses-displays-associated-with-outliers-or-shifts-from-normal-to-abnormal-version-10-10-sept-15-11815.pdf. Published September 9, 2015. Accessed May 5, 2020.

Vander Cruyssen B, Van Looy S, Wyns B, Westhovens R, Durez P, Van den Bosch F, Veys EM, Mielants H, De Clerck L, Peretz A, Malaise M, Verbruggen L, Vastesaeger N, Geldhof A, Boullart L, De Keyser F. DAS28 best reflects the physician's clinical judgment of response to infliximab therapy in rheumatoid arthritis patients: validation of the DAS28 score in patients under infliximab treatment. *Arthritis Res Ther*. 2005;7(5):R1063-R1071.

Signature Page for VV-CLIN-167399 v2.0

Approval	PPD 21-Jul-2025 14:41:21 GMT+0000

Signature Page for VV-CLIN-167399 v2.0