Statistical Analysis Plan Version 4 I4V-MC-JAIM

A Phase 3, Double-Blind, Multicenter Study to Evaluate the Long-Term Safety and Efficacy of Baricitinib in Patients with Systemic Lupus Erythematosus (SLE)

NCT03843125

Approval Date: 30-Nov-2021

1. Statistical Analysis Plan: I4V-MC-JAIM
A Phase 3, Double-Blind, Multicenter Study to Evaluate the Long-Term Safety and Efficacy of Baricitinib in Patients with Systemic Lupus Erythematosus (SLE)

Baricitinib (LY3009104) Systemic Lupus Erythematosus

Study I4V-MC-JAIM is a Phase 3, multicenter, randomized, long term extension trial to evaluate the long-term safety and efficacy of baricitinib in eligible patients with systemic lupus erythematosus who have completed the treatment period in an originating study (such as, Study I4V-MC-JAHZ or Study I4V-MC-JAIA).

Eli Lilly and Company
Indianapolis, Indiana USA 46285

[Protocol I4V-MC-JAIM]
[Phase 3]

Statistical Analysis Plan Version 1 electronically signed and approved by Lilly on 04 Sep 2019
Statistical Analysis Plan Version 2 electronically signed and approved by Lilly on 26 May 2021
Statistical Analysis Plan Version 3 electronically signed and approved by Lilly on 21 October 2021
Statistical Analysis Plan Version 4 electronically signed and approved by Lilly on date provided below.

Approval Date: 30-Nov-2021 GMT

2. Table of Contents

Section **Page**

Table of Contents

Table **Page**

Table of Contents

Figure **Page**

List of Appendices

Appendix **Page**

3. Revision History

Statistical Analysis Plan (SAP) Version 1 was approved prior to the first visit in which a subject received study drug.

Statistical Analysis Plan Version 2 was approved prior to the first outcome database lock.

The following key changes include

- added tables to summarize coronavirus disease 2019 (COVID-19) impact
- removed exploratory analyses as they will be retained in a separate document, and
- removed publicly available efficacy scale derivations.

Statistical Analysis Plan Version 3 was approved prior to the database lock/unblinding of the first pivotal originating study. The key changes include:

- updated secondary endpoints in Table 4.1 based on Protocol Amendment (a)
- added dose adjustment for renal impairment based on the protocol in Section 5.4, and added how the treatment was analyzed for patients who have decreased renal function
- added the Entered and Randomized populations in Table JAIM.6.2
- clarified the baseline definition
- removed the logistic regression model for the treatment comparisons in Section 6.4, which is not applicable
- added as-observed analysis in Section 6.4.2.1
- added baseline definitions for concomitant medication and prednisone dose in Section 6.10
- added the definition of prohibited concomitant medication
- added more details for the derivation of efficacy/health outcome measures and endpoints, including the missing data imputation rules for items and components in Table JAIM.6.3
- replaced the logistic regression and negative binomial regression analyses with descriptive analysis in Table JAIM.6.4, as no treatment comparisons will be performed for this study
- added analyses for the British Isles Lupus Assessment Group (BILAG) individual organ system, the British Isles Lupus Assessment Group-based Composite Lupus Assessment (BICLA), and the European Quality of Life–5 Dimensions–5 Level (EQ-5D-5L) item score
- updated the population used for the analyses of Systemic Lupus Erythematosus Disease Activity Index 2000 (SLEDAI-2K) and Work Productivity and Activity Impairment (WPAI) for Lupus
- added details to the definition of safety postbaseline time period in Table JAIM.6.5
- added the listing of opportunistic infections and summary of temporary interruption of study drug in Table JAIM.6.6
- added details of the overview of infections in Section 6.14.4
- added details of by-patient COVID-19 listings in Section 6.15

- removed the content of Section 7, Unblinding Plan, by referring to a separate unblinding plan
- updated the Anatomical-Therapeutic Chemical (ATC) classification codes and medical review process of prohibited medication in Appendix 1
- added column "Included in Interim 1 Analysis/Derivation" and updated a few categories of patient characteristics in Appendix 2
- added analysis plan for China patients in Appendix 3, and
- made other minor typographical corrections and clarifications not affecting content.

Statistical Analysis Plan Version 4 was approved prior to the database lock/unblinding of the second pivotal originating study JAHZ and prior to the JAIM first interim database lock and unblinding. The key changes include:

- Removed JAHZ PPD from the modified intent-to-treat (mITT) population and safety population due to confirmed misconduct that is likely to have significant impact on reliability and robustness of data at this particular site.

4. Study Objectives

Table JAIM.4.1 provides the protocol defined primary and major secondary objectives and endpoints of the study.

The estimand (EMA 2020) associated with each endpoint/analysis is documented in the following places:

- The population of interest is described in the protocol inclusion/exclusion criteria and in this document in Table JAIM.6.2 and Section 6.2.
- The endpoints/variables may be found in Table JAIM.6.3 and Table JAIM.6.4.
- The handling of intercurrent events and missing data may be found in Section 6.4.2.
- Population summary measures are described in Table JAIM.6.4.

Table JAIM.4.1. Primary and Secondary Objectives

Objectives	Endpoints (Variables)
Primary	
To evaluate the long-term safety and tolerability of baricitinib in patients with SLE.	Safety and tolerability assessments will include: • Proportion of patients with treatment-emergent adverse events (TEAEs), adverse events of special interest (AESIs), and serious adverse events (SAEs). • Proportion of patients with temporary investigational product interruptions and permanent discontinuations.
Secondary	
To evaluate the long-term effect of baricitinib 4-mg or 2-mg QD and background standard-of-care therapy on SLE disease activity.	• Proportion of patients achieving SRI-4 response through Week 156, defined as: o Reduction of ≥4 points from baseline in SLEDAI-2K score; and o No new British Isles Lupus Assessment Group (BILAG) A or* no more than 1 new BILAG B disease activity score; and o No worsening (defined as an increase of ≥0.3 points [10 mm] from baseline) in the Physician's Global Assessment of Disease Activity. • Proportion of patients achieving an SRI-5, -6, -7, or -8 response through Week 156. • Proportion of patients achieving an LLDAS response through Week 156. • Change from baseline in mean total SLEDAI-2K scores through Week 156. • Change from baseline in Physician's Global Disease Activity score through Week 156.
To evaluate the long-term corticosteroid sparing effect of baricitinib 4-mg or 2-mg QD.	Change from baseline in prednisone dose through Week 156.

Primary and Secondary Objectives

Objectives	Endpoints (Variables)
To evaluate the long-term effect of baricitinib 4-mg or 2-mg QD on SLE flares.	• Annualized mild/moderate flare rate • Annualized severe flare rate • Annualized flare rate (any severity)
To evaluate the long-term effect of baricitinib 4-mg or 2-mg QD on mucocutaneous manifestations of SLE.	• Proportion of patients with CLASI total activity score ≥10 at baseline with ≥50% reduction in CLASI total activity score through Week 156.
To evaluate the long-term effect of baricitinib 4-mg or 2-mg QD on musculoskeletal manifestations of SLE.	• Change from baseline in tender joint count through Week 156. • Change from baseline in swollen joint count through Week 156.
To evaluate the long-term effect of baricitinib 4-mg or 2-mg QD on individual organ system disease activity.	• Proportion of patients with improvement in each SLEDAI-2K organ system vs baseline through Week 156. • Proportion of patients with worsening in each SLEDAI-2K organ system vs baseline through Week 156.
To evaluate the long-term effect of baricitinib 4-mg or 2-mg QD on damage.	• Change from baseline in SLICC/ACR damage index total score through Week 156.
To evaluate the long-term effect of baricitinib 4-mg or 2-mg QD on patient-reported outcomes (PROs).	• Change from baseline in Worst Pain NRS through Week 156. • Change from baseline in Worst Joint Pain NRS through Week 156. • Change from baseline in Worst Fatigue NRS through Week 156. • Change from baseline in Patient Global Impression of Severity through Week 156. • Change from baseline in mental component score (MCS), physical component score (PCS), and domain scores in the Short-Form 36-item health survey version 2 (SF- 36v2) acute through Week 156. • Change from baseline in FACIT-F total score through Week 156. • Change from baseline in the EQ-5D-5L through Week 156. • Change from baseline in the WPAI-Lupus through Week 156.

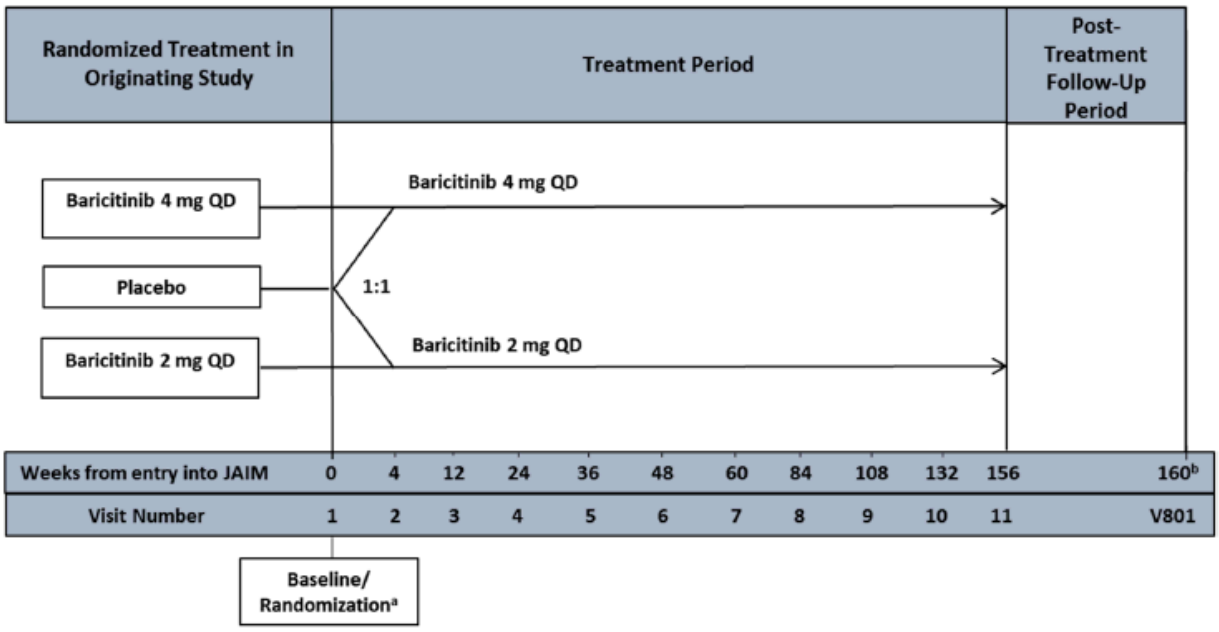
Abbreviations: BILAG = British Isles Lupus Assessment Group; LLDAS= Lupus Low Disease Activity State; NRS = Numeric rating scale; PGA = Physician Global Assessment; PRO = Patient reported outcome; QD = once daily; SLE = systemic lupus erythematosus; SLEDAI-2K = Systemic Lupus Erythematosus Disease Activity Index 2000; SRI = Systemic Lupus Erythematosus Responder Index.

* This is written as it is in protocol amendment (a), but the word “or” should be updated to “and”.

5. Study Design

5.1. Summary of Study Design

Study I4V-MC-JAIM (JAIM) is a Phase 3, multicenter, randomized, outpatient, long term extension trial to evaluate the long-term safety and efficacy of baricitinib 4-mg and 2-mg in eligible patients with systemic lupus erythematosus (SLE) who have completed the treatment period in an originating study (such as, Study I4V-MC-JAHZ [JAHZ] or Study I4V-MC-JAIA [JAIA]). Study JAIM will enroll patients who have completed treatment through week 52 of Study JAHZ or Study JAIA and meet eligibility requirements.



Abbreviations: QD = once daily; V = visit.

a Visit 1 of Study JAIM should be completed on the same calendar day as the Week-52 visit of the originating study.

b The follow-up visit should occur 4 weeks (±5 days) after the last dose of study drug.

Figure JAIM.5.1. Flow chart of study design for treatment period for Study JAIM.

5.2. Determination of Sample Size

It is expected that 80% of patients will complete Study JAHZ or JAIA and approximately 90% of patients who complete will enter Study JAIM; therefore, approximately 1100 patients will be enrolled from these originating studies to Study JAIM. Additional patients may enroll from addenda or other originating studies.

5.3. Method of Assignment to Treatment

Patients who meet all criteria for enrollment will be assigned to baricitinib treatment (2-mg once daily [QD] or 4-mg QD) according to the treatment they were receiving in the originating study. Patients randomized to active treatment during the originating study will be assigned to the same,

blinded dose of baricitinib in Study JAIM. Patients randomized to placebo during the originating studies will be randomized 1:1 to blinded baricitinib treatment (2-mg QD or 4-mg QD) in JAIM:

- Randomization will be stratified by corticosteroid dose on the day of randomization (<10 mg/day; ≥10 mg/day prednisone or equivalent), and region.
- Patients with renal impairment (as defined in protocol Section 7.2.2) who are randomized to baricitinib 4-mg QD will receive baricitinib 2-mg QD.

Table JAIM.5.1. Countries and Associated Geographical Regions from JAHZ and JAIA

Geographical Region for Patients from JAHZ & JAIA	Country or Countries
North America	United States
Central, South America and Mexico	Brazil, Mexico, Argentina, Chile, Colombia
Europe	Austria, Belgium, Croatia, Czech Republic, Germany, Greece, Hungary, Switzerland, The Netherlands, United Kingdom, France, Italy, Poland, Romania, Spain
Asia	China, Taiwan, Japan, Korea, Philippines
Rest of World	Australia, Israel, Russia, India, Serbia, South Africa

5.4. Dose Adjustment for Renal Impairment

The dose adjustment for renal impairment will be managed by an interactive web-response system (IWRS) to ensure maintenance of the treatment blind. For patients randomized to active treatment during the originating studies, the estimated glomerular filtration rate (eGFR) value from the originating screening visit will be used by an IWRS to assign the treatment doses accordingly. For patients randomized to placebo in the originating studies, the last available eGFR value from the originating study will used by an IWRS to assign the treatment doses accordingly.

- Patients with documented renal impairment (defined as eGFR ≥40 to <60 mL/min/1.73 m^2), who are assigned to active treatment (either to the baricitinib 4-mg arm or the baricitinib 2-mg arm) will receive a baricitinib 2-mg QD dose by the IWRS.
- No dose adjustment will be made for patients with eGFR ≥60 mL/min/1.73 m^2.

6. A Priori Statistical Methods

6.1. General Considerations

This Statistical Analysis Plan (SAP) describes analyses that will be utilized for the full clinical study report (CSR) for Study JAIM, at the time of JAIM primary outcome database lock. Analyses to be produced prior to the JAIM primary outcome database lock, for purposes of regulatory submissions, will be identified in the retained list of analyses. Additional safety analyses that integrate data from the originating studies will be specified in the integrated safety analysis plan (ISAP) for SLE. Any exploratory analyses will be described in a separate document. Statistical analysis of this study will be the responsibility of Eli Lilly and Company or its designee.

Not all displays described in this SAP will necessarily be included in the CSR. Not all displays will necessarily be created as a "static" display. Some may be incorporated into interactive display tools such as Spotfire 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.

Throughout this document, references to analyses of corticosteroids refer to analyses of systemic corticosteroids, unless otherwise specified.

Statistical analyses of categorical safety and efficacy endpoints will be descriptive in nature.

Frequency counts and percentages by treatment arm (see Table JAIM.6.1) will be provided for the safety and efficacy endpoints. For non-model-based efficacy endpoints, 95% confidence intervals (CIs) for single proportions using the Wilson Score (Wilson, 1927; Newcombe, 1998) method will be utilized.

Unless otherwise stated, inferential statistical analyses (p-values and 95% CIs) which compare treatment groups will not be conducted.

For continuous efficacy and health outcomes endpoints the significance of within-treatment group changes from baseline will be evaluated by testing whether the treatment group least squares mean (LS mean) changes from baseline are different from zero; the 95% CI and the standard error for the LS mean change will also be displayed.

Change from baseline (defined in Section 6.3) will be calculated as the postbaseline (at that visit) value minus the baseline value. If all baseline values are missing for a particular variable, the change from baseline and percent change from baseline will not be calculated.

Table JAIM.6.1 summarizes the treatment arms based on the treatment assigned in JAIM.

Table JAIM.6.1. Treatment Arms

JAIM Treatment Arms	Description for Patients from JAHZ/JAIA
Baricitinib 2-mg	All patients randomized in the originating studies to baricitinib 2-mg QD (They will remain on baricitinib 2-mg during JAIM).
Baricitinib 4-mg	All patients randomized in the originating studies to baricitinib 4-mg QD (They will remain on baricitinib 4-mg during JAIM.).
Placebo to baricitinib 2-mg	Patients randomized to Placebo at Week 0 of JAHZ/JAIA and subsequently randomized to baricitinib 2-mg at Week 0 (Visit 1) in JAIM.
Placebo to baricitinib 4-mg	Patients randomized to Placebo at Week 0 of JAHZ/JAIA and subsequently randomized to baricitinib 4-mg at Week 0 (Visit 1) in JAIM.
Additional JAIM Treatment Arms for Safety	
Pooled baricitinib	Patients treated with baricitinib 2-mg or 4-mg QD in Study JAIM.

Abbreviation: QD = once daily.

6.2. Definition of Populations

Analysis populations are defined in Table JAIM.6.2 along with their associated purpose. The treatment groups and inferential comparisons described in this table will be used unless otherwise specified. Also, unless otherwise specified, patients will be analyzed according to the treatment to which they were assigned. Patients who are randomized to 4 mg, but who have decreased renal function, identified by an eGFR measurement of <60 mL/min/1.73m^2, will receive a dose of 2 mg QD. These patients will be summarized within the 4-mg treatment arm for all summaries unless otherwise noted. A listing of treatment assignments showing the randomized treatment vs the treatment received for subjects with low eGFR (<60 mL/min/1.73m^2) will be provided. Maximized extended enrollment (ME2) patients in Study JAHZ will be enrolled in China to meet country-level regulatory requirements and these patients will be excluded from all analyses in Study JAIM, unless otherwise noted. Appendix 3 describes the analyses for all patients from China sites including the ME2 cohort.

Due to the confirmed misconduct at PPD that is likely to have significant impact on reliability and robustness of data at this particular site, the patients from this site will be excluded from the modified intent to treat (mITT) population and safety population.

Table JAIM.6.2. Analysis Populations

Population	Description
Entered Population	**Definition:** All patients who signed informed consent form in Study JAIM **Purpose:** Used for patient allocation summary **Treatment Groups:** As given in Table JAIM.6.1
Randomized Population	**Definition:** All patients who were randomized/assigned in Study JAIM (Note: patients randomized to active treatment during the originating study will be assigned to the same, blinded dose of baricitinib in Study JAIM; patients randomized to placebo during the originating studies will be randomized 1:1 to blinded baricitinib treatment of 2 mg once daily (QD) or 4 mg QD in Study JAIM) **Purpose:** Used for patient allocation summary **Treatment Groups:** As given in Table JAIM.6.1
Modified Intent-to-Treat (mITT) Population	**Definition:** All participating patients, excluding those from PPD, who receive at least 1 dose of study treatment in Study JAIM (regardless if the patient does not receive the correct treatment, or otherwise does not follow the protocol) **Purpose:** Used for efficacy and health outcomes related analyses **Treatment Groups:** As given in Table JAIM.6.1
Safety Population	**Definition:** For the purpose of analysis of Study JAIM alone, the safety population is defined as all patients, excluding those from PPD, who receive at least 1 dose of investigational product in Study JAIM and who did not discontinue from the study for the reason 'Lost to Follow-up' at the first postbaseline visit **Purpose:** Used for safety analyses **Treatment Groups:** Safety analyses will be analyzed according to 5 treatment groups as defined in Table JAIM.6.1

6.3. Definition of Baseline and Postbaseline Periods

For efficacy analyses, *baseline* for patients receiving active treatment (baricitinib 2-mg and 4-mg) in originating studies (JAHZ and JAIA) is defined as the last nonmissing measurement taken in the originating studies, prior to starting study medication. *Baseline* for patients receiving placebo in originating studies is defined as the last nonmissing assessment prior to the first dose of active study drug given in Study JAIM Visit 1 (Week 0). Any assessment collected after the JAIM Visit 1 dose through the date of Visit 11 or the early discontinuation visit is defined as *postbaseline* for the Treatment Period.

Protocol-defined visit windows are defined as ±4 days for Visit 2 (Week 4), ±7 days for Visit 3 (Week 12), ±10 days from Visit 4 (Week 24) until Visit 7 (Week 60), and ±14 days from Visit 8 (Week 84) until Visit 11 (Week 156). If there are data within the protocol-defined window, those data will be used in by-visit analyses. If a protocol-defined visit is missing, then the nearest unscheduled visit data that exist will be mapped to that visit and included in by-visit analyses.

For safety analyses, baseline is defined as the last nonmissing scheduled (planned) measurement prior to the first dose of baricitinib administration in Study JAIM for continuous measures by-

visit analyses and all nonmissing measurements prior to the first dose of the study drug administration (baricitinib) in Study JAIM for all other analyses.

Postbaseline measures for the safety analyses are defined as the nonmissing scheduled (planned) measurements after the first dose of baricitinib administration in Study JAIM for continuous measures by-visit analyses and all nonmissing measurements after the first dose of the study drug administration (baricitinib) in Study JAIM for all other analyses.

In general, when the baseline measurement is collected on the same day (where time is not collected) as the first dose date, this measurement will be used as the baseline value for efficacy and safety except adverse events (AEs).

6.4. Analysis Methods

Continuous efficacy and health outcome variables, except for Systemic Lupus International Collaborating Clinics (SLICC)/American College of Rheumatology (ACR) Damage Index will use a restricted maximum likelihood-based mixed model for repeated measures (MMRM) analysis. The response variable will be change from baseline across postbaseline visits. The model will include the following explanatory variables as fixed factors: treatment, baseline disease activity (total SLEDAI-2K <10; ≥10), baseline corticosteroid dose (<10 mg/day; ≥10 mg/day prednisone or equivalent), region (North America, Central/South America/Mexico, Europe, Asia and Rest of World), visit (as categorical variable), baseline value, treatment-by-visit interaction, and baseline value-by-visit interaction. Within-patient variance-covariance errors will be modeled using an unstructured covariance matrix. If this analysis fails to converge, the following structures will be tested in this pre-specified order: 1) heterogeneous Toeplitz (TOEPH), 2) heterogeneous autoregressive (ARH[1]), 3) heterogeneous compound symmetry (CSH), and 4) autoregressive (AR[1]). The Kenward-Roger method will be used to estimate the degrees of freedom. For each treatment group the least squares mean (LS mean), standard error, within-treatment p-value, and 95% CIs will be reported.

Systemic Lupus International Collaborating Clinics (SLICC)/ACR Damage Index change from baseline will be summarized with observed and modified last observation carried forward (mLOCF).

The binary efficacy measures will be summarized with observed case and non-responder imputation (NRI).

For both efficacy and safety evaluations, if study drug is temporarily interrupted (suspended) during a treatment period, the measurements taken during the temporary interruption will be included in the analysis.

Both observed and imputed analyses will be performed for efficacy and health outcomes measures.

All analyses will include the treatment arms given in Table JAIM.6.1.

6.4.1. Adjustments for Covariates

Randomization in the originating studies is stratified according to disease activity (total SLEDAI-2K at screening <10; ≥10), corticosteroid dose at baseline (<10 mg/day; ≥10 mg/day prednisone or equivalent), and region (North America, Central/South America/Mexico, Europe, Asia and Rest of World). Therefore, these factors will be adjusted for as described in Section 6.4.

6.4.2. Handling of Dropouts or Missing Data

Due to the long duration of JAIM, missing data that are imputed will be carried forward for the visit interval at Week 48, Week 108, and Week 156. Discontinued patients will not be included for evaluation in intervals subsequent to their discontinuation from the study. The following visit intervals are defined for NRI method and mLOCF method:

- Visit 6 (Week 48) [> Visit 1 (Week 0) and ≤ Visit 6 (Week 48)]
- Visit 9 (Week 108) [> Visit 6 (Week 48) and ≤ Visit 9 (Week 108)]
- Visit 11 (Week 156) [> Visit 9 (Week 108)]

6.4.2.1. As-Observed Analysis

The "as-observed" strategy is used in so-called "observed cases" or "completers" analysis, ubiquitous in the literature, but is not one of the recommended strategies in the ICH E9(R1). For this analysis, only data from completers at the visit are relevant, and therefore the analysis does not need to deal with missing data. This estimand is based on the subset of patients who would complete treatment through Visit X if assigned to it. Therefore, this estimand is conditional and targets the effect of treatment conditional on completion of treatment through the time point of interest. Because the estimand is defined for a subpopulation conditional on an intercurrent event, it is not causal. Summaries based on observed data at each postbaseline visit will be provided.

6.4.2.2. Non-responder Imputation (NRI)

The NRI method may be used when the estimand of interest uses the composite strategy (EMA 2020) for handling intercurrent events. In this strategy, patients with any intercurrent event that leads to the permanent discontinuation of assigned study treatment or to prohibited medication use are defined to have failed treatment for that particular visit interval but not for the subsequent visit intervals (as defined in Section 6.4.2). Therefore, using a composite estimand approach, no data would be missing and the function of the NRI is to estimate the estimand.

All patients who discontinue the study or permanently discontinue study treatment at any time for any reason or patients who have a prohibited medication in a particular visit interval will be defined as non-responders for that visit interval and excluded from the subsequent visit intervals. Enrolled patients without at least one postbaseline observation will be defined as non-responders for the first visit interval (Visit 6) and excluded from evaluation in subsequent visit intervals.

6.4.2.3. Mixed Model for Repeated Measures

For continuous variables, an MMRM analysis with a missing at random assumption for handling missing data will be performed. This likelihood-based analysis takes into account both missingness of data and the correlation of the repeated measurements.

The MMRM method may be justified when the estimand of interest uses the hypothetical strategy (EMA2020) for handling 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 as directed without dropping out of the study.

For patients who have prohibited medication use any time after baseline, any observed data after the date of the prohibited medication use will be excluded from the MMRM analysis. The earliest date of prohibited medications will be used for censoring.

6.4.2.4. Modified Last Observation Carried Forward (mLOCF)

For patients who receive prohibited medication, the last nonmissing observation at or before prohibited use will be carried forward to that particular visit interval but not for the subsequent visit intervals (as defined in Section 6.4.2). For all other patients discontinuing from the study or permanently discontinuing the study treatment for any reason, the last nonmissing postbaseline observation on or before discontinuation will be carried forward to that particular visit interval but not for the subsequent visit intervals (as defined in Section 6.4.2). If a patient does not have a nonmissing observed record (or one imputed by other means) for a postbaseline visit, the last postbaseline record prior to the missed visit will be used for this postbaseline visit. If a patient does not have any postbaseline record at a scheduled, unscheduled or early termination visit, they will be excluded from the analysis.

6.5. Multicenter Studies

This study will be conducted at multiple sites internationally for each originating study. When specified, statistical analyses will adjust for geographical region. Classification of geographical region is defined in Table JAIM.5.1.

6.6. Multiple Comparisons/Multiplicity

No adjustments for multiple comparisons will be utilized in the statistical analyses for this study.

6.7. Patient Disposition

Treatment and study disposition will be summarized for the modified intent-to-treat (mITT) population. Patient disposition will be summarized by treatment group with reasons for discontinuation. Included in the reasons for treatment or study discontinuation will be a category for discontinuations due to coronavirus disease 2019 (COVID-19)-related logistical issues (ie, quarantine or travel restrictions, non-illness related). For patients who discontinued the treatment or study due to COVID-19 infection, the reason for treatment discontinuation will be classified as “due to AEs” even when the discontinuation was due to the requirement of quarantine. Follow-up disposition will be summarized separately from the treatment period

summaries. A listing of patient disposition will be provided for all participating patients, with the extent of their participation in the study and the reason for discontinuation.

6.8. Patient Characteristics

Patient demographics and baseline characteristics will be summarized using the mITT population, by treatment group. No formal statistical comparisons will be made among treatment groups, unless otherwise stated. A listing of patient demographics will be provided.

6.9. Treatment Compliance

Patient compliance with investigational product will be assessed at each visit and at Early Termination during the treatment period by counting returned tablets. A patient will be considered significantly noncompliant if he or she misses more than 20% of the prescribed doses during the study, unless the patient's investigational product was withheld by the investigator for safety reasons. If a patient had his/her treatment temporarily interrupted by the investigator due to any reason, the number of days that drug was withheld or unavailable will be deducted from the total number of days in computing the expected total number of tablets used. Compliance for patients without available data due to early termination or lost-to follow-up will be based on last completed visit with available tablet count.

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

Patient treatment compliance will be summarized for the mITT population using the intervals defined in Section 6.4.2.

6.10. Concomitant Therapy

Concomitant medications will be coded and descriptively summarized by treatment group in terms of frequencies and percentages using the mITT population.

Medications taken during JAIM that started before the last date of the treatment period and are ongoing or ended during the treatment period will be classified as "concomitant" medication.

The baseline for concomitant medications is defined as the medication and dose taken on the day prior to the date of the first dose of study drug administration (baricitinib).

For Zoster and tuberculosis (TB) vaccine, since only start date is reported, if start date is missing, these medications will be considered to be prior therapy.

Patients may continue to receive the background standard therapies for SLE they were receiving at completion of the originating study. Changes in dose of background standard therapies for SLE, such as an increase in dose of antimalarials or immunosuppressants, are permitted during the study. In addition, changes in concomitant medication, such as an increase in corticosteroids to manage any increase in disease activity or corticosteroid tapering, are permitted during the study in accordance with long term management of disease activity and at the investigator's discretion. Any required corticosteroid bursts should be aligned with clinical practice.

Prohibited concomitant medications are defined as

- initiation of new antimalarials or immunosuppressants
- recurrent intravenous injection of any corticosteroids, or
- having any other biologics or protocol-prohibited concomitant medications, defined from the ATC codes of L01XC, L01XX, L04AB, L04AC, V98, or J06B (excluding J06BB).

Prohibited medications during the treatment period will be reviewed by the study team to assess clinical significance of change in medication. The final determination of clinical meaningfulness will be documented and will be incorporated into the analysis datasets.

Concomitant medication use during Study JAIM will be summarized by treatment, organized according to ATC code level 1 and preferred name, sorted by frequency in the placebo to baricitinib 4-mg group. The summaries of concomitant medications will be provided overall and separately for concomitant medications used to treat SLE and for statin concomitant medications. Note that a patient will only be counted once, regardless of how many times medication included under the same preferred name was taken.

A by-patient listing of all concomitant medications will be provided for the mITT patients.

For corticosteroids, all doses of steroids will be converted to an equivalent dose of prednisone (see Section 6.10.1). Using the prednisone-equivalent dose, a daily dose of corticosteroid will be derived for each study day from baseline to the treatment discontinuation day for patients who discontinue treatment early or to the day before for treatment completers.

6.10.1. Corticosteroid Dosing Conversion

To allow for assessments of changes in doses of various corticosteroids, all corticosteroid doses will be standardized to an equivalent prednisone dose. Appendix 1 provides a complete summary of conversion factors for each corticosteroid medication identified during the study, instructions for selecting corticosteroids, and the manual review process. An intramuscular (IM) injection is considered a single dose on the date IM injection is administered. This dose of corticosteroid will be referred to as "prednisone (or equivalent)" throughout this document.

Baseline prednisone (or equivalent) dose will be the total daily dose of all corticosteroids being taken by a patient on the day prior to the date of the first dose of study drug administration (baricitinib). This baseline dose will be used for baseline summaries and for comparisons to later visits. A daily dose of prednisone (or equivalent) will be calculated for each day between baseline and study discontinuation. The corticosteroid dose for a postbaseline visit will be the daily dose of prednisone (or equivalent) at each visit date minus 1.

6.11. Efficacy Analyses

6.11.1. General Considerations

Table JAIM.6.3 includes the description, derivation and definition of missingness of the efficacy and health outcomes measures and endpoints. Table JAIM.6.4 includes the description of the

analysis method, population and primary time point(s) of analysis for efficacy and health outcomes measures. Assessments collected at multiple visits will be analyzed at each visit, in addition to the primary time point. Details of each analysis will follow the general principles described under Analysis Methods in Section 6.4.

Table JAIM.6.3. Description and Derivation of Efficacy/Health Outcomes Measures and Endpoints

Measure	Description	Variable	Derivation/Comment	Definition of Missing
SRI-4	SRI-4 is a composite index to measure overall improvement in disease activity (SLEDAI-2K) while ensuring there is no worsening in other organ systems (BILAG and PGA)	SLE Responder Index 4	A decrease in SLEDAI-2K ≥4 (from baseline), and No new BILAG A and no more than 1 new BILAB B organ domain (both compared with baseline), and No worsening in PGA (defined as an increase of 0.3 points [10 mm] from baseline).	Missing, if any component (SLEDAI, BILAG, or PGA) remains missing after each instrument-level imputation rule is applied
SLEDAI-2K	The SLEDAI-2K is a global disease activity instrument that focuses on high-impact disease manifestations across 9 organ domains: constitutional, mucocutaneous, musculoskeletal, vascular, cardiorespiratory, central nervous system, immunologic, renal, and hematologic.	SLEDAI-2K Total Score	Calculated by summing the weighted organ manifestations. For the Renal domain, when the urinary casts, hematuria, and pyuria items are "not done" or "not accessed", they will be coded as "not present".	If any item scores or all items are missing (not done or not assessed), but the visit occurred, data for that item can be carried forward if obtained within the previous 36 days of visit.
		Individual Organ Domain Improvement defined by SLEDAI-2K	Among patients with SLEDAI-2K score >0 within the organ domain at baseline and able to decrease SLEDAI-2K (from baseline) within each organ domain score, separately.	Missing if any SLEDAI-2K item for that organ domain remains missing after instrument-level imputation rules are applied.
		Individual Organ Domain Worsening defined by SLEDAI-2K	Among patients with at least 1 SLEDAI-2K item = 0 in the organ domain score at baseline and able to increase SLEDAI-2K (from baseline) within each organ domain score, separately.	Missing if any SLEDAI-2K item for that organ domain remains missing after instrument-level imputation rules are applied.
		No worsening from baseline in SLEDAI-2K	No increment from baseline of >0 points in SLEDAI-2K.	Missing if any SLEDAI-2K item for that organ domain remains missing after instrument-level imputation rules are applied.
BILAG-2004	BILAG-2004 assesses 97 clinical signs, symptoms and laboratory	BILAG	A, B, C, D, or E score will be used in analyses for each of the 9 individual	Within each organ domain (except renal and hematological), any

Measure	Description	Variable	Derivation/Comment	Definition of Missing
	parameters across nine organ system domains: constitutional, mucocutaneous, neuropsychiatric, musculoskeletal, cardiorespiratory, gastrointestinal, ophthalmic, renal and hematological		organ systems.	missing data will be assumed to be 'Not present' if there is at least 1 nonmissing item in that organ. If all items in 1 organ domain are completely missing but the visit occurred, then the letter score of that organ from the previous visit will be pulled forward, provided data can be obtained within 36 days of visit; otherwise missing except when the letter score from the last nonmissing visit is E, then E score will be pulled forward. For Renal and Hematological domain, if any item is missing but there is at least 1 nonmissing item in the organ, the missing data will be assumed to be "No". If any vital/lab item with "Yes" but vital and lab value is missing, then vital/lab values can be pulled forward, provided data can be obtained within 36 days of visit; otherwise missing. After the value imputation, if the A/B/C score for the organ domain still cannot be derived, then the letter score of that organ from the previous visit will be pulled forward, provided data can be obtained within 36 days of visit; otherwise missing except when the letter score from the last nonmissing visit is E, then E score will be pulled forward.
		BILAG	Reduction of all baseline BILAG A to	Missing if any of 9 domains at

Measure	Description	Variable	Derivation/Comment	Definition of Missing
		improvement	B/C/D and baseline BILAG B to C/D and no BILAG worsening in other organ systems, where worsening is defined as ≥1 new BILAG A or ≥2 new BILAG B	baseline or at the visit remain missing after instrument-level imputation rules are applied, except when the missing is at the baseline (any one or all domains) and the value at the visit is not A or B, then the other nonmissing organ domains will be used to determine the response status.
		No BILAG worsening	No new BILAG A and no >1 new BILAG B disease activity score (both compared with baseline), where worsening is defined as ≥1 new BILAG A or ≥2 new BILAG B, both compared with the baseline. The baseline BILAG A improved to BILAG B at the visit will not be considered as the new BILAG B.	Missing if any of 9 domains at baseline or at the visit remains missing after instrument-level imputation rules are applied except when the baseline (any one or all domains) is BILAG A and the value at the visit is missing, then the other nonmissing organ domains will be used to determine the response status.
		Individual Organ Domain Improvement Defined by BILAG	Among patients with BILAG A or B at baseline and able to reduce the baseline BILAG A to B/C/D and BILAG B to C/D for each organ domain, separately.	Missing if baseline or value remains missing after instrument-level imputation rules are applied.
		Individual Organ Domain Worsening Defined by BILAG	Among patients without BILAG A at baseline, an increment of baseline BILAG B/C/D/E to A or baseline BILAG C/D/E to B for each organ domain, separately.	Missing if baseline or value remains missing after instrument-level imputation rules are applied.
Physician's Global Assessment of Disease Activity	The PGA is the physician's assessment of the patient's overall disease activity due to SLE. It is scored using a visual analog scale	PGA score PGA category	Permitted range of values is from 0 to 100 mm. PGA categories are defined as: None (0) = '0 mm', Mild (>0 and <1.5) = '>0 to <50 mm', Moderate (≥1.5 to ≤2.5)	If the visit occurred, data can be carried forward if obtained within 36 days of visit; otherwise missing.

Measure	Description	Variable	Derivation/Comment	Definition of Missing
(PGA)	where 0 mm indicates no disease activity and 100 mm indicates the most severe disease activity possible. There are benchmarks of 0 (0 mm), 1 (33 mm), 2 (67 mm), and 3 (100 mm) on the line corresponding to no, mild, moderate, and severe SLE disease activity, respectively.		= '≥50 mm to ≤83mm', Severe (>2.5) = '>83mm'.	
		No worsening in PGA	Worsening is defined as an increase of ≥0.3 points (10 mm) from baseline. Therefore, no worsening is defined as any decrease, no change, or <0.3 points (10 mm) increase from baseline.	Missing if baseline or value remains missing after instrument-level imputation rules are applied.
		PGA ≤1	PGA ≤33 mm	Missing if value remains missing at the visit after instrument-level imputation rules are applied.
BICLA	BICLA is a composite index to measure overall improvement in disease activity (BILAG) while ensuring there is no worsening in other organ systems (SLEDAI-2K and PGA)	BICLA	• Reduction of all baseline BILAG A to B/C/D and baseline BILAG B to C/D and no BILAG worsening in other organ systems, where worsening is defined as ≥1 new BILAG A or ≥2 • new BILAG B • No worsening from baseline in SLEDAI-2K, where worsening is defined as an increase from baseline of >0 points in SLEDAI-2K No worsening from baseline in participants' lupus disease activity, where worsening is defined by an increase ≥0.30 points on a 3-point PGA visual analogue scale (VAS).	Missing if any component (BILAG, SLEDAI or PGA) remains missing after each instrument level imputation rule is applied.
Cutaneous Lupus Erythematosus Disease Area and Severity Index (CLASI)	The CLASI is a validated instrument to assess cutaneous manifestations of SLE consisting of 2 scores: • Activity • Damage	Total Activity Score	Calculated within tablet system with edit checks; no additional derivation.	Within tablet system: if any item scores, but not all, are missing then impute a score of 0 for the missing item(s). For completely missing questionnaires, data can be carried forward if obtained within 36 days of visit; otherwise missing.

Measure	Description	Variable	Derivation/Comment	Definition of Missing
		Total Damage Score	Calculated within tablet system with edit checks; no additional derivation.	Within tablet system: if any item scores, but not all, are missing then impute a score of 0 for the missing item(s). For completely missing questionnaires, data can be carried forward if obtained within 36 days of visit; otherwise missing.
Lupus Low Disease Activity State (LLDAS)	The LLDAS is a composite measure designed to identify patients achieving a state of low disease activity.	LLDAS	• SLEDAI-2K ≤4with no activity in SLEDAI-2K major organ systems (CNS, Vascular, Renal, cardiorespiratory, and constitutional), where "no activity" is defined as all items of SLEDAI-2K within these major organ systems = 0. • No new features of Lupus disease activity in SLEDAI-2K compared to previous occurred visit, where the "new feature" is defined as any of SLEDAI-2K 24 items changed from 0 to >0. • PGA ≤1 • Current prednisone or equivalent dose ≤7.5 mg/day	Missing if any component remains missing after each instrument-level imputation rule is applied.
SFI flare	The SFI uses the SLEDAI score, disease activity scenarios, treatment changes, and PGA to define mild/moderate and severe flares.	SFI flare	Not derived; used as entered.	The absence of a flare record, or 'Not Applicable', both are indicative of no occurrence of flare.
		Annualized flare rate	Calculated as the number of SFI flares divided by the flare exposure time in days multiplied by 365.25.	No minimum data requirement.
Tender/Swollen Joint Count (28 Joints)	28 joints are assessed as tender only, swollen only, tender and swollen, not evaluable.	Tender Joint Count 28/ Swollen Joint Count 28	The number of tender and swollen joints will be calculated by summing all joints respectively.	For patients who have an incomplete set of joints evaluated, the joint count will be adjusted to a 28-joint count for tenderness and a 28-joint count

Measure	Description	Variable	Derivation/Comment	Definition of Missing
				for swelling by dividing the number of affected joints by the number of evaluated joints and multiplying by 28.
Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index (SLICC/ACR)	SLICC/ACR Damage Index assesses damage to 12 organ systems regardless of its cause.	SLICC/ACR Damage Index Score	Calculated as the sum of all available entries with a maximum score of 47.	"Missing" item is coded as 0.
Worst Pain NRS	This Worst Pain NRS is a single-item patient-administered 11-point horizontal scale anchored at 0 (no pain) and 10 (pain as bad as you can imagine)	Worst Pain NRS score	Permitted range of values is from 0 to 10.	Single item, missing if missing.
FACIT-Fatigue	The FACIT-Fatigue scale (Cella and Webster 1997) is a 13-item symptom-specific questionnaire that assesses self-reported severity of fatigue and its impact upon daily activities and functioning.	FACIT-Fatigue total score	The FACIT-F uses 0 ("not at all") to 4 ("very much") numeric rating scales to assess fatigue and its impact in the past 7 days. Scores range from 0 to 52 with higher scores indicating less fatigue. The FACIT-Fatigue Scoring Guidelines (Version 4) will be used to calculate the Total Score. Reversals are needed for all items except An5 and An7, as described in scoring manual. http://www.ser.es/wp-content/uploads/2015/03/FACIT-F_INDICE.pdf.	Missing items are acceptable as long as more than 50% of the items are answered (ie, a minimum of 7 out of 13 items), the sum of available items will be multiplied by 13 then divided by the number of items answered to obtain the total score. If less than 7 items are answered, the FACIT-Fatigue total score will be set to missing.
Steroid sparing	This measures the change in corticosteroid dose through week 156	Corticosteroid Dose	See Section 6.10.1 for detailed derivations.	No minimum data requirement.
Worst Joint Pain NRS	This Worst Joint Pain NRS is a single-item, patient-administered 11-point horizontal scale anchored at 0	Worst Joint Pain NRS Score	Permitted range of values is from 0 to 10.	Single item, missing if missing.

Measure	Description	Variable	Derivation/Comment	Definition of Missing
	(no joint pain) and 10 (joint pain as bad as you can imagine)			
Worst Fatigue NRS	This Worst Fatigue NRS is a patient-administered 11-point horizontal scale anchored at 0 (no fatigue) and 10 (fatigue as bad as you can imagine)	Worst Fatigue NRS Score	Permitted range of values is from 0 to 10.	Single item, missing if missing.
Patient Global Impression of Severity	The Patient's Global Impression of Severity is a single-item question asking the patient how they would rate their overall lupus symptoms over the last 24 hours.	Patient Global Impression of Severity Score	Possible values are: (0) No symptoms, (1) Very mild, (2) Mild (3) Moderate (4) Severe	Single item, missing if missing.
SF-36 v2 Acute	The SF-36 Version 2 acute (1-week recall) is a 36-item, patient-completed measure designed to be a short, multipurpose assessment of health (The SF Community – SF-36 Health Survey Update). The summary scores range from 0 to 100, with higher scores indicating better levels of function and/or better health. Items are answered on Likert scales of varying lengths. The SF-36 comprises 8 domain scores and 2 overarching component scores. SF-36 domain scores are: (1) Physical functioning, (2) Role-physical, (3) Role-emotional, (4) bodily pain, (5) vitality, (6) social functioning, (7) mental health and (8) general health. The component scores are: (1) the	SF-36 Domain scores and SF-36 Component Scores	Per copyright owner, the Quality Metric Health Outcomes™ Scoring Software will be used to derive SF-36 domain and component scores. SF-36V2 Normed-Based Scoring will be used for the domain and component scores and includes 2 steps: 1/ z-score standardization of SF-36v2 scales; and 2/ norm-based transformation of the SF-36v2 z scores. 2009 norms will be used. (Maruish 2011).	Missing data handling offered by SF-36 software will be used. Maximum Data Recovery will be selected for Missing Score Estimator in the software.

Measure	Description	Variable	Derivation/Comment	Definition of Missing
	Physical Component Summary (PCS) and (2) Mental Component Summary (MCS).			
EQ-5D-5L	The European Quality of Life–5 Dimensions–5 Level (EQ-5D-5L) is a standardized measure of health status used to provide a simple, generic measure of health for clinical and economic appraisal. The EQ-5D-5L consists of 2 components: a descriptive system of the respondent's health and a rating of his/her current health state using a 0- to 100-mm VAS. The descriptive system comprises the following 5 dimensions: Item 1: mobility Item 2: self-care Item 3: usual activities Item 4: pain/discomfort Item 5: anxiety/depression The respondent is asked to indicate his/her health state by ticking (or placing a cross) in the box associated with the most appropriate statement in each of the 5 dimensions. It should be noted that the numerals 1 to 5 have no arithmetic properties and should not be used as an ordinal score. The VAS records the respondent's self-rated health on a vertical VAS where the endpoints are	EQ-5D-5L Items	Five health profile dimensions, each dimension has 5 levels: 1 = no problems 2 = slight problems 3 = moderate problems 4 = severe problems 5 = extreme problems It should be noted that the numerals 1 to 5 have no arithmetic properties and should not be used as a primary score.	Each dimension is a single item, missing if missing. Note: a score of 9 is missing.
		EQ-5D-5L UK/US Population-based index score	Uses the concatenation of the value of each EQ-5D-5L dimension score in the order: Item 1, Item 2, Item3; Item 4; Item 5. Derive EQ-5D-5L UK/US Population-based index score according to the link by using the UK/US algorithm (Szende et al. 2006) to produce a patient-level index score between -0.59 and 1.0 (continuous variable): https://euroqol.org/eq-5d-instruments/eq-5d-5l-about/valuation-standard-value-sets/crosswalk-index-value-calculator/.	If any of the items is missing or equal to 9, the index score is missing
		EQ-5D VAS	Single item. Range from 0 = "worst imaginable health state" to 100 = "best imaginable health state". Note: higher value indicates better health state.	Single item, missing if missing

Measure	Description	Variable	Derivation/Comment	Definition of Missing
	labeled "best imaginable health state" and "worst imaginable health state." This information can be used as a quantitative measure of health outcome. The EQ 5D 5L health states, defined by the EQ 5D 5L descriptive system, may be converted into a single summary index by applying a formula that essentially attaches values (also called weights) to each of the levels in each dimension (Herdman et al. 2011; EuroQol Group 2015 [WWW]).			
WPAI-Lupus	The Work Productivity and Activity Impairment – Lupus (WPAI-Lupus) is a patient-reported instrument developed to measure the impact on work productivity and regular activities attributable to SLE. It contains 6 items that measure: 1) employment status, 2) hours missed from work due to the specific health problem, 3) hours missed from work for other reasons, 4) hours actually worked, 5) degree health affected productivity while working, and 6) degree health. Scores are calculated as impairment percentages (Reilly et al. 1993), with higher scores indicating greater productivity impairment.	Employment Status	Yes/No	Missing if answer is missing
		Absenteeism Score (%)	$\frac{Q2}{(Q2+Q4)} \times 100$	Missing if Q2 or Q4 are missing. Also missing if Employment Status is No.
		Presenteeism Score (%)	$\frac{Q5}{10} \times 100$	Missing if Q5 is missing. Also missing if Employment Status is No.
		Work productivity Loss Score (%)	$\left[\frac{Q2}{Q2+Q4} + \left(1 - \frac{Q2}{Q2+Q4}\right)\frac{Q5}{10}\right] \times 100$	Missing if Q2, Q4 or Q5 is missing. Also missing if Employment Status is No.
		Activity Impairment Score (%)	$\frac{Q6}{10} \times 100$	Missing if Q6 is missing. May still be present and nonmissing if patient is unemployed.

Abbreviations: BICLA = British Isles Lupus Assessment Group-based Composite Lupus Assessment; BILAG = British Isles Lupus Activity Group; CNS = Central Nervous System; FACIT = Functional Assessment of Chronic Illness Therapy; MCS= Mental Component Score; NRS = Numeric Rating Scale; PGA = Physician's Global Assessment; PCS= Physical Component Summary; SRI-4 = Systemic Lupus Erythematosus (SLE) Responder Index; SLEDAI-2K = Systemic Lupus Erythematosus Disease Activity Index 2000; SFI = SLEDAI Flare Index; SF-36 = Short Form 36; V = visit; VAS = Visual Analog Scale.

Table JAIM.6.4. Description of Efficacy/Health Outcomes Analyses

Measure	Variable	Analysis Method	Population	Key Time Point(s)[a]
SRI-4	Proportion of patients achieving SRI-4 response	Descriptive analysis with observed and NRI	mITT, patients with SLEDAI-2K score of ≥4 at baseline	Week 156 visit
SRI-x (where x is 5, 6, 7, 8)	Proportion of patients achieving SRI-x response	Descriptive analysis with observed and NRI	mITT, patients with SLEDAI-2K score of ≥x at baseline	Week 156 visit
SLEDAI-2K	Change from baseline in mean total SLEDAI-2K through Week 156	MMRM	mITT	Week 156 visit
	Proportion of patients achieving reduction of ≥4 points from baseline	Descriptive analysis with observed and NRI	mITT, patients with SLEDAI-2K score ≥4 at baseline	Week 156 visit
	Proportion of patients with improvement in each SLEDAI-2K individual organ systems vs baseline	Descriptive analysis with observed and NRI	mITT, patients with SLEDAI-2K score >0 within the organ domain at baseline	Week 156 visit
	Proportion of patients with worsening in each SLEDAI-2K individual organ system vs baseline	Descriptive analysis with observed and NRI	mITT, patients with at least 1 SLEDAI-2K item = 0 in the organ domain at baseline	Week 156 visit
BILAG	Proportion of patients with no new BILAG A and no more than 1 new BILAG B disease activity score (compared with baseline)	Descriptive analysis with observed and NRI	mITT	Week 156 visit
	Proportion of patients with improvement in each BILAG individual organ system vs baseline	Descriptive analysis with observed and NRI	mITT, patients with BILAG A or B in the organ domain at baseline	Week 156 visit
	Proportion of patients with worsening in each BILAG individual organ system vs baseline	Descriptive analysis with observed and NRI	mITT, patients without BILAG A in the organ domain at baseline	Week 156 visit
PGA	Proportion of patients with no worsening in PGA (compared with baseline)	Descriptive analysis with observed and NRI	mITT	Week 156 visit
	Change from baseline in PGA score over time through Week 156	MMRM	mITT	Week 156 visit
BICLA	Proportion of patients achieving BICLA response	Descriptive analysis with observed and NRI	mITT	Week 156 visit
28 Tender/Swollen Joint Count	Change from baseline in tender joint count	MMRM	mITT	Week 156 visit
	Change from baseline in swollen joint count	MMRM	mITT	Week 156 visit

Measure	Variable	Analysis Method	Population	Key Time Point(s)[a]
CLASI Activity Score	The proportion of patients with ≥ 50% reduction from baseline in total activity score	Descriptive analysis with observed and NRI	mITT – patients with CLASI total activity score ≥10 at baseline	Week 156 visit
SFI	Annualized severe flare rate	Descriptive analysis with observed rate	mITT	Week 0 to Week 156 visit
	Annualized severe flare rate	Descriptive analysis with observed rate	mITT	Week 0 to Week 156 visit
	Annualized severe flare rate	Descriptive analysis with observed rate	mITT	Week 0 to Week 156 visit
SLICC/ACR Damage Index Score	Change from baseline in SLICC/ACR Damage Index	Descriptive analysis with observed and mLOCF	mITT	Week 156 visit
LLDAS	Proportion of patients who achieve a LLDAS	Descriptive analysis with observed and NRI	mITT, patients with SLEDAI score of ≥4 at baseline	Week 156 visit
Corticosteroid Sparing	Change from baseline in prednisone dose	MMRM	mITT	Week 156 visit
Worst Pain NRS	Change from baseline in Worst Pain NRS	MMRM	mITT	Week 156 visit
FACIT-Fatigue	Change from baseline in FACIT-Fatigue total score	MMRM	mITT	Week 156 visit
Worst Fatigue NRS	Change from baseline in Worst Fatigue NRS	MMRM	mITT	Week 156 visit
Worst Joint Pain	Change from baseline in Worst Joint Pain NRS	MMRM	mITT	Week 156 visit
Patient Global Impression of Severity	Change from baseline in Patient Global Impression of Severity	MMRM	mITT	Week 156 visit
SF-36 V2 acute	Change from baseline in mental component score (MCS)	MMRM	mITT	Week 156 visit
	Change from baseline in physical component score (PCS)	MMRM	mITT	Week 156 visit
	Change from baseline in domain scores SF-36 health survey version 2 acute	MMRM	mITT	Week 156 visit
EQ-5D-5L	Change from baseline in EQ-5D-5L US/UK population-based index scores and VAS.	MMRM	mITT	Week 156 visit
	Proportion of patients with no problem in each EQ-5D-5L item score	Descriptive analysis with observed and NRI	mITT	Week 156 visit

Measure	Variable	Analysis Method	Population	Key Time Point(s)[a]
WPAI-Lupus	Change from baseline in WPAI-Lupus	MMRM	For absenteeism score, presenteeism score, and work productivity loss score, population is a subset of mITT that includes patients who had an employment status of Yes at baseline. For activity impairment score, use mITT population	Week 156 visit

Abbreviations: BICLA = British Isles Lupus Assessment Group-based Composite Lupus Assessment; BILAG = British Isles Lupus Activity Group; CNS = Central Nervous System; EQ-5D-5L = European Quality of Life5 Dimensions–5 Level; FACIT = Functional Assessment of Chronic Illness Therapy; mLOCF = modified last observation carried forward; mITT = modified intent-to-treat; MMRM = mixed model for repeated measures; NRI = non-responder imputation; NRS = Numeric Rating Scale; PCS = Physical Component Summary; SRI-4 = Systemic Lupus Erythematosus (SLE) Responder Index; SLEDAI-2K = Systemic Lupus Erythematosus Disease Activity Index 2000; SF-36 = Short Form 36; WPAI-Lupus = Work Productivity and Activity Impairment Questionnaire: Lupus V2.0.

[a] Assessments collected at multiple visits will be analyzed at each visit, in addition to the key time point listed below.

6.11.2. Primary Outcome and Methodology

Efficacy is not a primary outcome for this study.

6.11.3. Secondary Efficacy Outcomes & Methodology

6.11.3.1. Secondary Efficacy Analyses

The analysis of secondary endpoints is described in Table JAIM.6.3 and Table JAIM.6.4 and Section 6.4.

6.11.3.2. Exploratory Efficacy Endpoints

Analyses of exploratory efficacy endpoints will be described in a separate SAP.

6.12. Health Outcomes/Quality-of-Life Analyses

Analysis of the health outcomes/quality-of-life measures are described in Table JAIM.6.4 and Section 6.4. The descriptions and high level derivations are provided in Table JAIM.6.3. Additional analyses will be documented in a separate analysis plan.

6.13. Bioanalytical and Pharmacokinetic/Pharmacodynamic Methods

No pharmacokinetic or pharmacodynamic analyses are planned for this study.

6.14. Safety Analyses

Detailed analyses and discussion of Study JAIM safety data are more thoroughly assessed in the context of combining the safety data from the originating studies JAHZ and JAIA with the safety data from Study JAIM. The planned analyses are described in the iSAP.

The safety postbaseline period is defined in Table JAIM.6.5.

Table JAIM.6.5. Definition of Postbaseline Period

Analysis Population	Treatment Groups (Short Label)	Postbaseline Time Period
Safety Population	See Table JAIM.6.1	For all analyses except the "by-visit" continuous analysis: From immediately after first dose of baricitinib in Study JAIM and ending at the earliest of • 30 days after last dose date of study drug • the study disposition date, or • the cut-off date if subject is on treatment. For the "by-visit" continuous analyses: Includes planned/scheduled measurements after first dose of baricitinib up to last visit, excluding the follow up visit.

6.14.1. Adverse Events

For the purposes of the CSR for Study JAIM *alone*, the summaries given in Table JAIM.6.6 are planned.

Table JAIM.6.6. Summary Tables Related to Adverse Events

Analysis
Listing of adverse events leading to permanent discontinuation of study drug and discontinuation from the study
Listing of deaths
Listing of SAEs
Listing of Opportunistic Infections
Overview of AEs
Overview of Infections
The number and percentage of patients with TEAE using MedDRA Preferred Term nested within SOC
The number and percentage of patients who permanently discontinued from study drug due to an adverse event (including adverse events that led to death) using MedDRA Preferred Term nested within SOC
The number and percentage of patients reporting SAEs using MedDRA Preferred Term nested within SOC
The number and percentage of patients with at least one temporary interruption of study drug.

Abbreviations: AE = adverse event; PT = Preferred Term; SAE = serious adverse event; SOC = System Organ Class; TEAE = treatment-emergent adverse event.

A treatment-emergent AE (TEAE) is defined as an event that first occurred or worsened in severity after the first dose of study treatment in Study JAIM and on or prior to the last visit date during the analysis period. The analysis period is defined as the treatment period plus up to 30 days off-drug follow-up time.

6.14.2. Extent of Exposure

Exposure, including total patient-years of exposure, will be calculated for Study JAIM. Duration of exposure to study drug will be summarized using descriptive statistics (n, mean, standard deviation [SD], minimum, first quartile, median, third quartile, maximum). Cumulative

exposure and duration of exposure will be summarized in terms of frequency counts and percentages by category and treatment group.

Total patient-years of exposure (PYE) will be reported for each treatment group for overall duration of exposure. Exposure will also be classified by the frequency of patients falling into different exposure ranges. Exposure ranges will generally be reported in weeks using the following as a general guide:

- ≥4 weeks, ≥16 weeks, ≥24weeks, ≥52 weeks, ≥84 weeks, and ≥156 weeks.
- >0 to <4 weeks, ≥4 weeks to <16 weeks, ≥16 weeks to <24 weeks, ≥24weeks to <52 weeks, ≥52 weeks to <84 weeks, ≥84 weeks to <156 weeks, and ≥156 weeks.

Overall exposure in JAIM will be summarized in total PYE which is calculated according to the following formula:

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

Additionally, time spent in the study will be summarized with descriptive statistics (n, mean, SD, minimum, first quartile, median, third quartile and maximum).

6.14.3. Clinical Laboratory Evaluation

For the categorical analysis, the analysis period is defined as the treatment period plus up to 30 days off-drug follow-up time. The analysis period for the continuous by-visit analyses (eg, change from baseline by visit) is defined as the treatment period excluding off-drug follow-up time. Refer to the iSAP for more details. The following analyses are planned:

- clinical laboratory evaluation in terms of box plots and treatment-emergent summaries.

6.14.4. Vital Signs and Other Physical Findings

For the categorical analysis, the analysis period is defined as the treatment period plus up to 30 days off-drug follow-up time. The analysis period for the continuous by-visit analyses (eg, change from baseline by visit) is defined as the treatment period excluding off-drug follow-up time. Refer to the ISAP for more details. The following analyses are planned:

- vital signs in terms of box plots and treatment-emergent summaries.

6.14.5. Special Safety Topics, Including Adverse Events of Special Interest

Special safety topics will be analyzed in the context of combining the safety data from the originating Studies JAHZ and JAIA with the safety data from Study JAIM. The planned analyses are described in the ISAP.

For the purposes of the CSR for Study JAIM alone, the summaries of infections and type of infections will be provided.

The following summaries for each treatment group will be included in the overview of infections:

- patients with ≥1 treatment-emergent (TE) infection
- serious infections
- AEs that led to
    - permanent discontinuation from study drug, and
    - temporary interruption from study drug
- TE opportunistic infection
- TE herpes zoster
- TE herpes simplex
- TE tuberculosis, and
- TE viral hepatitis.

6.15. COVID-19 Trial Impact

Patients who experience an impact to their trial participation due to quarantine and/or travel restrictions related to COVID-19 will have their type of impact summarized. COVID-19-specific impacts for summarization include protocol deviations, including out of window visits, treatment interruptions, treatment and/or study discontinuations, and missed visits, regardless of whether or not the patient has a COVID-19 infection documented as an AE. The proportion of patients impacted in each category will be summarized. This summary will be provided for the overall mITT patients as well as by region. The number of patients who missed visits related to COVID-19 will be summarized by visit.

The following by-patient listings will be provided

- listing of study and treatment disposition related to COVID-19
- listing of important/nonimportant protocol deviations due to COVID-19
- listing of COVID-19 AEs (based on Standardized Medical Dictionary for Regulatory Activities Query [SMQ] 20000237 using the narrow-term classification)
- listing of temporary treatment interruptions related to COVID-19, and
- listing of missed visits related to COVID-19.

6.16. Protocol Violations

Protocol deviations will be identified throughout the study. Important protocol deviations (IPDs) are defined as those deviations from the protocol that would potentially compromise patient safety, data integrity, or study outcome.

A by-patient listing of IPDs will be provided.

A summary of the number and percentage of patients with an important protocol deviation by treatment group, overall, and by type of deviation will be provided.

6.17. Data Monitoring Committee

A data monitoring committee (DMC) will oversee the conduct of all the Phase 3 clinical trials evaluating baricitinib in patients with SLE. The DMC will consist of members external to Lilly.

This DMC will follow the rules defined in the DMC charter, focusing on assessing safety data for baricitinib. The same DMC will monitor safety-related data for ongoing studies of baricitinib in other indications.

Details of unblinding are given in Section 7 and the DMC charter.

6.18. Annual Report Analyses

Based on regulatory requirements for the Development Safety Update Report (DSUR), reports will be produced (if not already available from the CSR) for the reporting period covered by the DSUR.

6.19. 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 and 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 SAE 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 reported each event term
    - the number of events experienced.

Consistent with www.ClinicalTrials.gov requirements, 'Other' AEs that occur in fewer than 5% of patients/subjects in every treatment group may not be included if a 5% threshold is chosen (5% is the minimum threshold).

7. Unblinding Plan

Unblinding details are specified in a separated blinding/unblinding plan.

8. References

Cella D, Webster K. Linking outcomes management to quality-of-life measurement. *Oncology*. (Willeston Park). 1997;11(11A):232-235.

[EMA] European Medicines Agency. ICH E9 (R1) Addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials. Published February 18, 2020. Accessed July 16, 2021. https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e9-r1-addendum-estimands-sensitivity-analysis-clinical-trials-guideline-statistical-principles_en.pdf

Firth D. Bias reduction of maximum likelihood estimates. *Biometrika* 1993;80(1):27-38.

Maruish ME (Ed.). User's manual for the SF-36v2 Health Survey (3rd ed.). Lincoln, RI: QualityMetric Incorporated. 2011.

Newcombe RG. Interval estimation for the difference between independent proportions: comparison of eleven methods. *Stat Med*. 1998;17:873-890. Erratum in Stat Med. 1999;18:1293.

Szende A, Oppe M, Devlin N. EQ-5D value sets: Inventory, comparative review and user guide. EuroQol Group Monographs. Volume 2. Springer, 2006.

Wilson, E. B. Probable inference, the law of succession, and statistical inference. *JASA*. 1927;22:209-212.

Yee CS, Cressell L, Farewell V, Rahman A, Teh LS, Griffiths B, et al. (2010). Numerical scoring for the BILAG-2004 index. *Rheumatology* (Oxford). 2010;49:1665-1669.

9. Appendices

Appendix 1. Selection of Medications

The following categories of medication are required. For each category, instructions for selecting the correct medications are given.

Corticosteroids

The following ATC codes will be used to select all possible systemic corticosteroids:

- H02 Corticosteroids for systemic use (except H02AA Mineralocorticoids), specifically
    - H02AB Glucocorticoids
    - H02BX Corticosteroids for systemic use, combinations
    - H02CA Anticorticosteroids,

and additionally

- M01BA Anti-inflammatory/antirheumatic agents in combination with corticosteroids

All unique preferred names in the database falling under the above ATC codes will be reviewed by the Lilly Medical Group in order to determine which ones should be included.

For each patient who has recurrent, intravenous injection of any corticosteroids with a start date or stop date during the treatment period (ie, from the date of randomization to the day of completion/early termination), the medications for this patient will be reviewed and the patient will be classified as having "prohibited" or "not prohibited" use during the treatment period. Note: determination of prohibited use is based on a case-by-case assessment of recurrent intravenous corticosteroid use, considering if dose and/or duration of use are clinically relevant.

All corticosteroid doses need to be converted to prednisone equivalent doses. If additional conversion factors are required, these will be added to the table below in a SAP revision prior to database lock.

The following table should be used for converting nonprednisone medications to prednisone equivalent:

> Multiply the dose of the corticosteroid taken by the patient (in milligrams) in Column 1 by the conversion factor in Column 2 to get the equivalent dose of prednisone (in milligrams).
>
> *Example: Patient is taking 16 mg of methylprednisolone po daily. To convert to prednisone: 16 mg methylprednisolone × 1.25 = 20 mg prednisone. 16 mg of methylprednisolone po daily is equivalent to 20 mg of prednisone po daily.*

Column 1	Column 2
Corticosteroid Preferred Name	Conversion factor for converting to an equivalent prednisone dose
Prednisone	1
Prednisone acetate	1
Prednisolone	1
Prednisolone acetate	1
Prednisolone sodium phosphate	1
Methylprednisolone	1.25
Methylprednisolone acetate	1.25
Methylprednisolone sodium succinate	1.25
Triamcinolone	1.25
Triamcinolone acetonide	1.25
Triamcinolone hexacetonide	1.25
Cortisone	0.2
Cortisone acetate	0.2
Hydrocortisone	0.25
Hydrocortisone acetate	0.25
Hydrocortisone sodium succinate	0.25
Betamethasone	6.25
Betamethasone acetate	6.25
Betamethasone dipropionate	6.25
Betamethasone sodium phosphate	6.25
Dexamethasone	6.25
Dexamethasone acetate	6.25
Dexamethasone phosphate	6.25
Dexamethasone sodium phosphate	6.25
Paramethasone	2.5
Deflazacort	0.83
Celestona bifas	6.25
Depo-medrol med lidokain	1.25
Diprospan	6.25
Fluocortolone	1
Meprednisone	1.25

Antimalarials

The following ATC codes will be used to select all possible antimalarials:

- P01B Antimalarials, specifically
    - P01BA Aminoquinolines
    - P01BB Biguanides
    - P01BC Methanolquinolines
    - P01BD Diaminopyrimidines
    - P01BE Artemisinin and derivatives, plain
    - P01BF Artemisinin and derivatives, combinations
    - P01BX Other antimalarials
- M09AA Quinine and derivatives

All unique preferred names in the database falling under the above ATC code will be reviewed by the Lilly Medical Group in order to determine which ones should be included.

For patients who have at least 1 new antimalarial with a start date or stop date during the treatment period (ie, from the date of randomization to the day of completion/early termination), all antimalarial medications for this patient will be reviewed and the patient will be classified as having "prohibited" or "not prohibited" use during the treatment period. Note: determination of prohibited use is based on a case-by-case assessment of new antimalarial use, considering if dose and/or duration of use are clinically relevant (eg, switch from hydroxychloroquine to equivalent dose of chloroquine is not prohibited).

Immunosuppressants

The following ATC codes will be used to select all possible immunosuppressants:

- M01AX Other antiinflammatory and antirheumatic agents, non-steroids. Note: for M01AX, only if the medication preferred name is sulfasalazine
- M01C Specific Antirheumatic Agents, specifically
    - M01CA Quinolines
    - M01CB Gold preparations
    - M01CC Penicillamine and similar agents
    - M01CX Other specific antirheumatic agents
- L01AA Nitrogen mustard analogues
- L01BA Folic acid analogues
- L04AA: Selective immunosuppressants
- L04AD Calcineurin inhibitors

- L04AX Other immunosuppressants
- J04BA Drugs for Treatment of Lepra

All unique preferred names in the database falling under the above ATC codes will be reviewed by the Lilly Medical Group in order to determine which ones should be included.

For patients who have at least 1 new immunosuppressant medication with a start date or stop date during the treatment period (ie, from the date of randomization to the day of completion/early termination), all immunosuppressant medications for this patient will be reviewed and the patient will be classified as having "prohibited" or "not prohibited" during the treatment period. Note: determination of prohibited use is based on a case-by-case assessment of new immunosuppressant use, considering if dose and/or duration of use are clinically relevant (eg, switch from methotrexate oral daily dose to equivalent weekly injection is not prohibited).

Intravenous immunoglobulin

The following ATC code will be used to select all immunoglobulins:

- J06B Immunoglobulins (excluding J06BB as unlikely to be IV)

Biologics

The following ATC code will be used to select all biologics:

- L01XC Monoclonal antibodies
- L01XX Other antineoplastic agents
- L04AA Selective immunosuppressants
- L04AB Tumor necrosis factor alpha (tnf-) inhibitors
- L04AC Interleukin inhibitors
- M01CX Other specific antirheumatic agents
- M05BX Other drugs affecting bone structure and mineralization
- V98 Investigational drug

All unique preferred terms in the database falling under the above ATC code will be reviewed by the Lilly Medical Group in order to determine which ones should be included.

NSAIDs

The following ATC codes will be used to select all NSAIDs:

- M01 Anti-inflammatory and Antirheumatic Products, specifically
    - M01AA Butylpyrazolidines
    - M01AB Acetic acid derivatives and related substances
    - M01AC Oxicams

  - M01AE Propionic acid derivatives
  - M01AG Fenamates
  - M01AH Coxibs
  - M01BX Other antiinflammatory/antirheumatic agents in combination with other drugs
- M09AX Other drugs for disorders of the musculo-skeletal system

Live Vaccines

All medications falling under the following ATC code should be selected and reviewed by the Lilly medical group so that they can confirm which are live vaccines:

- J07 Vaccines

Appendix 2. Baseline Measures and Patient Characteristics

This appendix itemizes the specific baseline measures and patient characteristics to be presented and how they will be summarized. Changes to Table JAIM.APP.1, including the summary of additional patient characteristics, will not require an amendment to the SAP.

Table JAIM.APP.1 Summary Tables Related to Patient Characteristics

Variable	Continuous Summary	Categorical Summary[a]	Included in Interim 1 Analysis/Derivation
Age[b]	Yes	<65 years, ≥65 to <75 years, ≥75 to <85, ≥85; ≥65; ≥75; <40 years, ≥40 years;	Yes
Sex	No	Male, Female	Yes
Ethnicity (US only)	No	Hispanic/Latino, Non-Hispanic/Non-Latino	Yes
Race	No	American Indian/Alaska Native, Asian, Black/African American, Native Hawaiian or other Pacific Islander, White, or Multiple	Yes
Geographic Region	No	North America, Central/South America/Mexico, Europe, Asia, Rest of World	Yes
Country	No	By country	Yes
Height (cm)	Yes	None	Yes
Waist circumference (cm)	Yes	None	Yes
Weight (kg)	Yes	<80 kg, ≥80 kg	Yes
		<100 kg, ≥100 kg	Yes
BMI at Visit 2[c]	Yes	Underweight (<18.5 kg/m^2), Normal (≥18.5 and <25 kg/m^2), Overweight (≥25 and <30 kg/m^2), Obese (≥30 and <40 kg/m^2), Severely obese (≥40 kg/m^2)	Yes

Summary Tables Related to Patient Characteristics

Variable	Continuous Summary	Categorical Summary[a]	Included in Interim 1 Analysis/ Derivation
Baseline Disease Characteristics			
Time since diagnosis of lupus (years)[e]	Yes	<1 year, ≥1 to <3 years, ≥3 year to <7 years, ≥7 years	Yes/ Time since diagnosis of lupus (years) = (date of first dose in JAIM - date of diagnosis of lupus in original study + 1) / 365.25.
Age at diagnosis of SLE[f]	Yes	<6 years, ≥6 to <10 years, ≥10 to <17 years, ≥17 to <40 years, ≥40 years	Yes/ Age at diagnosis in years will be calculated as the time interval from the imputed date of birth (01 July in the year of birth collected in the eCRF) to the date of diagnosis in original study.
Any flare at baseline, (SELENA-SLEDAI Flare Index definition)	No	None, Mild/Moderate, Severe	No
SLEDAI-2K score	Yes	(<10 or ≥10)	No
Physician's Global Assessment of Disease Activity score	Yes	Mild (<1.5), Moderate (≥1.5 to ≤2.5), Severe (>2.5)	No
Total systemic Lupus International Collaborating Clinics/American College of Rheumatology (SLICC/ACR) Damage Index score	Yes	None	No
Patient's Global Impression of Severity (PGI-S)	Yes	5 categories: no symptoms, very mild, mild, moderate, severe	No
Number of tender joints (from 28-tender joint count)	Yes	None	No
Number of swollen joints (from 28-tender joint count)	Yes	None	No

Summary Tables Related to Patient Characteristics

Variable	Continuous Summary	Categorical Summary[a]	Included in Interim 1 Analysis/ Derivation
Cutaneous Lupus Erythematosus Disease Area and Severity Index (CLASI) Total Activity Score	Yes	CLASI (<10, ≥10)	No
CLASI Total Damage Score	Yes	None	No
Complement C3 levels	Yes	Low (<90mg/dL); Normal or higher (≥90mg/dL)	No
Complement C4 level	Yes	Low (<10mg/dL); Normal or higher (≥ 10mg/dL)	No
Complement C3 and C4 levels	No	Low C3 and/or low C4 levels; high C3 and C4 levels	No
Anti-dsDNA level (positive >15 U/mL)	Yes	Yes, No	No
Complement and Anti-dsDNA level	No	Low complement/anti-dsDNA positive (low C3 and/or C4 levels and anti-dsDNA positive); Others	No
Serum immunoglobulin IgA, IgG, and IgM concentrations	Yes	Less than the lower limit of normal [LLN] (Yes, No)	No
Estimated glomerular filtration rate (eGFR)	Yes	eGFR (<60, ≥60 mL/min/1.73m^2)	Yes

Summary Tables Related to Patient Characteristics

Variable	Continuous Summary	Categorical Summary[a]	Included in Interim 1 Analysis/ Derivation
British Isles Lupus Assessment Group (BILAG) organ system involvement at baseline (yes or no for each organ system domain). Involvement requires a baseline BILAG disease activity score of A or B.	No	Yes, No	No
BILAG A organ system involvement at baseline (yes or no for each organ system domain). Involvement requires a baseline BILAG disease activity score of A.	No	Yes, No	No
SLEDAI-2K organ system involvement at baseline (yes or no for each organ system domain)	No	Yes, No	No
Anti-Sm+ antinuclear antibodies (positive >10 U/mL)	Yes	Yes, No	No
Anti-RNP+ antinuclear antibodies (positive >10 U/mL)	Yes	Yes, No	No
Anti-Sjögren's-syndrome-related antigen A (SSA/Ro-52) antibodies (positive >10 U/mL)	Yes	Yes, No	No
Anti-Sjögren's-syndrome-related antigen A (SSA/Ro-60) antibodies (positive >10 U/mL)	Yes	Yes, No	No
Anti-Sjögren's-syndrome-related antigen B (also called anti-La [Anti-SSB/La+]) antibodies (positive >10 U/mL)	Yes	Yes, No	No
Anti-phospholipid antibody, overall (if any of cardiolipin IgA, IgG, IgM, beta-2-glycoprotein IgG, beta-2-glycoprotein IgM, or lupus anticoagulant are positive, then overall is positive, otherwise negative)	No	Positive, Negative	No
ANA positive (titer ≥1:80)	No	Yes, No	No

Summary Tables Related to Patient Characteristics

Variable	Continuous Summary	Categorical Summary[a]	Included in Interim 1 Analysis/ Derivation
Corticosteroid use[d]. See Section 6.10.1 for details of prednisone (or equivalent) baseline dose	Yes	Yes, No	Yes
		<10 mg/day or ≥10 mg/day	Yes
		≤7.5 mg/day or >7.5 mg/day	Yes
Immunosuppressant[d] use at baseline	No	Yes, No	Yes
Mycophenolate mofetil use at baseline	No	Yes, No	Yes
Azathioprine[d] use at baseline	No	Yes, No	Yes
Methotrexate[d] use at baseline	No	Yes, No	Yes
Antimalarial[d] use at baseline	No	Yes, No	Yes
Hydroxychloroquine[d] use at baseline	No	Yes, No	Yes
NSAID[d] use at baseline	No	Yes, No	Yes
Statin use at baseline	No	Yes, No	Yes
Other Baseline Measures			
Baseline SF-36 PCS, MCS	Yes	None	No
Baseline WPAI Lupus – Employment status	No	Yes, No	No
Baseline WPAI Lupus – Absenteeism score, Presenteeism Score, Work Productivity Loss Score, Activity Impairment Score	Yes[g]	None	No
EQ-5D-5L VAS and Item scores	Yes	None	No
FACIT-Fatigue Total score	Yes	None	No
Worst Pain NRS	Yes	None	No
Worst Joint Pain NRS	Yes	None	No
Worst Fatigue NRS	Yes	None	No

Summary Tables Related to Patient Characteristics

Abbreviations: ANA = antinuclear antibodies; APL = antiphospholipid antibody; BMI = body mass index; dsDNA = double stranded DNA; eCRF = electronic case report form; EQ-5D-5L = 5-level EQ-5D version; FACIT = Functional Assessment of Chronic Illness Therapy; IgA/G/M = immunoglobulin A/G/M; MCS = Mental Component Summary; NRS = Numeric Rating Scale; NSAID = Non-steroidal anti-inflammatory drug; PCS = Physical Component Summary; QIDS SR-16 = Quick Inventory of Depressive Symptomatology; SELENA-SLEDAI = Systemic Lupus Erythematosus Disease Activity Index; SF-36 = Short Form 36; SLE = Systemic Lupus Erythematosus; SLEDAI-2K = Systemic Lupus Erythematosus Disease Activity Index 2000; VAS = Visual Analog Scale; WPAI = Work Productivity and Activity Impairment.

[a] This column specifies the levels of the categorical variable to be summarized.

[b] Age in years will be calculated as length of the time interval from the imputed date of birth (July 1st in the year of birth collected in the electronic case report form [eCRF]) to the informed consent date. Formula below: age (in years) = (date of informed consent date – imputed date of birth)/365.25.

[c] Body Mass Index will be calculated as: $BMI\ (kg\ /\ m^2) = Weight\ (kg)\ /(Height\ (m))^2$.

[d] ATC codes for selecting medications are provided in Appendix 1.

[e] Time since diagnosis of lupus (years) = (date of first dose - date of diagnosis of lupus + 1) / 365.25.

[f] Age at diagnosis in years will be calculated as the time interval from the imputed date of birth (July 1st in the year of birth collected in the eCRF) to the date of diagnosis.

[g] Only applied for Employment Status at baseline is "Yes", except the Activity Impairment Score.

Appendix 3. Analysis Plan for China Patients

General Considerations

This analysis plan describes the efficacy and safety analyses for the overall China patients in Study I4V-MC-JAIM (JAIM) who were enrolled in the global cohort or under the maximized extended enrollment (ME2) addendum from the original study, I4V-MC-JAHZ (JAHZ). For the details of the protocol addendum including rational, design summary, and sample size determination, please refer to the JAHZ Statistical Analysis Plan (SAP) for ME2 Addendum Version 1.

Analysis Population

Patients who are screened in China with an informed consent date after 03 September 2020 in Study JAHZ will be considered in the China ME2 cohort (Cohort 2) in Study JAHZ.

Entered population - overall China patients set is defined as all patients from China who signed informed consent in Study JAIM and were originally enrolled in the global cohort (Cohort 1) or the ME2 cohort (Cohort 2) of Study JAHZ.

Unless otherwise specified, efficacy and health outcomes analyses will be conducted on the following populations even if the patient does not take the assigned treatment, does not receive the correct treatment, or otherwise does not follow the protocol:

- Modified intent-to-treat population - Overall China patients set, defined as all randomized patients from China in Study JAIM who were originally enrolled in the global cohort or the ME2 cohort of Study JAHZ and received at least 1 dose of study treatment in Study JAIM.

Safety analyses will be conducted on the following populations:

- Safety population - Overall China patients set, defined as all randomized patients from China in Study JAIM who were originally enrolled in the global cohort or the ME2 cohort of Study JAHZ and received at least 1 dose of study treatment in Study JAIM and did not discontinue from the study for the reason “lost to follow-up” at the first postbaseline visit.

Analysis Methods

The same planned efficacy, health outcomes, and safety analyses that are described in the main body of this SAP will be produced based on the overall China patients set. All analysis endpoint definitions, data handling, and statistical analysis methods will be the same as in the main body of this SAP unless otherwise specified.

Descriptive summary statistics, such as means and proportions, will be reported. Analysis considerations such as those presented in the main body of this SAP will be used. All analyses will be for descriptive purposes only.

The analyses of patient disposition, important protocol deviations, patient characteristics, and prior and concomitant therapy will be as described in the main body of this SAP.

For efficacy and health outcomes analyses, selected time points will be considered. The efficacy analyses will be as described in the main body of this SAP with the same analysis methods including the method for handling the missing data for both coronavirus disease 2019 (COVID-19) and non-COVID-19-related reasons, except region will be excluded from the analysis model.

The safety analyses will be the same as described in the main body of this SAP.

Leo Document ID = fb1c8cbb-3812-4999-a023-7c165213e5fe

Approver: PPD

Approval Date & Time: 30-Nov-2021 17:25:53 GMT

Signature meaning: Approved