I4V-MC-JAIX Statistical Analysis Plan Version 5

A Multicenter, Open-Label, Phase 3 Study to Evaluate the Efficacy and Safety of Baricitinib in Adult Patients with Moderate to Severe Atopic Dermatitis

NCT03559270

Approval Date: 24-Jun-2022

1. Statistical Analysis Plan: A Multicenter, Open-Label, Phase 3 Study to Evaluate the Efficacy and Safety of Baricitinib in Adult Patients with Moderate to Severe Atopic Dermatitis

BREEZE – AD6

Confidential Information

The information contained in this document is confidential and the information contained within it may not be reproduced or otherwise disseminated without the approval of Eli Lilly and Company or its subsidiaries.

Note to Regulatory Authorities: This document may contain protected personal data and/or commercially confidential information exempt from public disclosure. Eli Lilly and Company requests consultation regarding release/redaction prior to any public release. In the United States, this document is subject to Freedom of Information Act (FOIA) Exemption 4 and may not be reproduced or otherwise disseminated without the written approval of Eli Lilly and Company or its subsidiaries.

Baricitinib (LY3009104) Atopic Dermatitis

Study I4V-MC-JAIX is a Phase 3, multicenter, open-label, outpatient, 204-week study designed to evaluate the efficacy and safety of baricitinib 2-mg in patients with moderate to severe atopic dermatitis.

Eli Lilly and Company
Indianapolis, Indiana USA 46285

Protocol I4V-MC-JAIX
Phase 3

Statistical Analysis Plan V1 electronically signed and approved by Lilly on 21 June 2018
Statistical Analysis Plan V2 electronically signed and approved by Lilly on 24 January 2020
Statistical Analysis Plan V3 electronically signed and approved by Lilly on 22 May 2020
Statistical Analysis Plan V4 electronically signed and approved by Lilly on11 Feb 2021
Statistical Analysis Plan V5 electronically signed and approved by Lilly on date provided below:

2. Table of Contents

Section **Page**

Table of Contents

Table **Page**

3. Revision History

Statistical Analysis Plan (SAP) Version 1 was approved prior to first visit.

Statistical Analysis Plan Version 2 was approved based on Protocol I4V-MC-JAIX(c) (JAIX) and Program Safety Statistical Analysis Plan (PSAP) Version 6. The overall changes incorporated in Version 2 are as follows:

- updated primary and key secondary efficacy endpoints per JAIX Protocol Amendment (c)
- updated with the extension of the study
- updated with rolling over of responder from Study I4V-MC-JAIW (JAIW) to Study JAIX at Study JAIW Week 104
- updated baseline definition for the daily diary assessments (Itch Numeric Rating Score [NRS], Atopic Dermatitis Sleep Scale [ADSS], Skin Pain NRS, Patient Global Impression of Severity– Atopic Dermatitis [PGI-S-AD])
- updated daily diaries to use an interval mean weekly score
- updated derived data and baseline variables
- updated exploratory analyses
- updated safety analyses to only include duration of exposure, treatment-emergent adverse events (TEAEs), serious adverse events (SAEs), and adverse events (AEs) leading to permanent study drug discontinuation
- updated safety analysis period to be the treatment period plus up to 30 days off-drug follow-up time
- updated TEAE definition by using the Study JAIX baseline to derive TEAEs
- removed previous medication summary
- updated with interim analyses for the clinical study report (CSR)

Statistical Analysis Plan Version 3 was approved before the 4 month safety update database lock. The overall changes incorporated in Version 2 are as follows:

- updated safety analyses cohorts
- updated treatment compliance timeframe
- detailed efficacy summary analyses for 4 month safety update
- Section 4.3 – updated to include exploratory analyses accounting for duration of treatment exposure in Study JAIW
- Section 6.3.1 - added a section to include missing data imputation method for missingness due to COVID-19

Statistical Analysis Plan Version 4 was approved before Study JAIX (2) open-label addendum first patient visit:

- Section 4.3, Section 5.1, Section 5.2, and Section 6.1 - added Study JAIX (2) open-label addendum information
- Updated a treatment group for Study JAIX (2) open-label addendum in Table JAIX.6.1.
- Section 6.1.2 - updated baseline information for Study JAIX (2) open-label addendum

Statistical Analysis Plan Version 5 was approved before final data base lock. The overall changes incorporated in version 5 are summarized below:

- Section 4.1, Section 4.2, Section 4.3, Section 5.3, and Section 6.1 - added Study JAIX (2) open-label addendum information
- Section 6.1.1 – added population for Study JAIX (2) open-label addendum
- Section 6.1.2 - updated baseline information for Study JAIX (2) open-label addendum
- Section 6.6, Section 6.7, Section 6.7.1, Section 6.8, Section 6.10, and Section 6.10.2 - added Study JAIX (2) open-label addendum information

4. Study Objectives

4.1. Primary Objective

To describe the clinical response to baricitinib 2-mg once daily (QD) in patients with moderate to severe atopic dermatitis (AD) by prior treatment received (ie, baricitinib 2-mg, 1-mg, or placebo) and responder status achieved in Study JAIW as measured by proportion of patients achieving a 75% or greater improvement in the Eczema Area and Severity Index (EASI75) from baseline of originating study assessed at Week 16.

Study JAIX (2) open-label addendum

To describe the clinical response to baricitinib 2-mg QD in patients with moderate to severe AD as measured by proportion of patients achieving a 75% or greater improvement in the Eczema Area and Severity Index (EASI75) from baseline, assessed at Week 16.

4.2. Secondary Objectives

- To describe the clinical response to baricitinib 2-mg QD in AD as measured by improvement in signs and symptoms of AD by prior treatment received (ie, baricitinib 2-mg, 1-mg, or placebo) and responder status achieved in Study JAIW as measured by the following:
    - Proportion of patients achieving Validated Investigator's Global Assessment for Atopic Dermatitis (vIGA-AD) of 0 or 1 assessed at Week 16
    - Proportion of patients achieving IGA of 0 or 1 assessed at or before Week 16
    - Proportion of patients achieving EASI75 from baseline of originating study JAIW assessed at or before Week 16
    - Mean percent change from baseline in Eczema Area and Severity Index (EASI) score from baseline of originating Study JAIW assessed at Week 16
    - Proportion of patients achieving a body surface area (BSA) of ≤3% assessed at or before Week 16
- To describe the clinical response to baricitinib 2-mg QD in AD as assessed by patient-reported outcome measures by prior treatment received (ie, baricitinib 2-mg, 1-mg, or placebo) and responder status achieved in Study JAIW as measured by the following:
    - Proportions of patients achieving a ≥4-point improvement in Itch NRS from baseline of originating Study JAIW, assessed at Week 16
    - Proportions of patients achieving a ≥4-point improvement in Itch NRS from baseline of originating Study JAIW assessed at or before Week 16
    - Mean percent change from baseline of originating Study JAIW in Itch NRS assessed at Week 16

Study JAIX (2) open-label addendum

- To describe the clinical response to baricitinib 2-mg QD in AD as measured by improvement in signs and symptoms of AD as measured by the following:
    - Proportion of patients achieving Validated Investigator's Global Assessment for Atopic Dermatitis (vIGA-AD) of 0 or 1 assessed at Week 16

  - Proportion of patients achieving IGA of 0 or 1 assessed at or before Week 16
  - Proportion of patients achieving EASI75 from baseline, assessed at or before Week 16
  - Mean percent change from baseline in Eczema Area and Severity Index (EASI) score from baseline, assessed at Week 16
  - Proportion of patients achieving a body surface area (BSA) of ≤3% assessed at or before Week 16

- To describe the clinical response to baricitinib 2-mg QD in AD as assessed by patient-reported outcome measures as measured by the following:
  - Proportions of patients achieving a ≥4-point improvement in Itch NRS from baseline, assessed at Week 16
  - Proportions of patients achieving a ≥4-point improvement in Itch NRS from baseline, assessed at or before Week 16
  - Mean percent change from baseline in Itch NRS, assessed at Week 16

4.3. Exploratory Objectives

- Exploratory objectives may include evaluating the response to baricitinib treatment regimens on clinical measures and patient-reported outcomes. These endpoints may include dichotomous endpoints or change from baseline for the following measures: BSA, EASI, SCORing Atopic Dermatitis (SCORAD), Patient-Oriented Eczema Measure (POEM), Dermatology Life Quality Index (DLQI), Itch NRS, ADSS Item scores, and Skin Pain NRS.
  - Mean change from baseline of the originating study JAIW in Skin Pain NRS
  - Mean change from baseline of the originating study JAIW in ADSS1/ADSS2/ADSS3
  - Mean change from baseline of the originating study JAIW in BSA
  - Mean change and percentage from baseline of the originating study JAIW in SCORAD
  - Mean change and percentage change from baseline of the originating study JAIW in Itch NRS
  - Mean change from baseline of the originating study JAIW in total POEM
  - Mean change from baseline of the originating study JAIW in DLQI
  - Mean change from baseline of the originating study JAIW in EASI
  - Proportion of patients achieving a 90% or greater improvement in the Eczema Area and Severity Index (EASI90) or a 50% or greater improvement in the Eczema Area and Severity Index (EASI50) from baseline of the originating study
  - Proportions of patients achieving a Skin pain 4-point improvement for those with baseline skin pain ≥4 of the originating study JAIW
  - Proportions of patients achieving a DLQI score of 5 or less for those with a baseline DLQI score >5 of the originating study JAIW
  - Proportions of patients achieving a DLQI score of 0 or 1

Summaries of response rates defined based on IGA of 0 or 1, EASI75, or Itch >=4 point improvement will also be provided for previous responders at Week 16 of Study JAIW (IGA 0 or

1) with loss of response (defined as IGA ≥3 or requiring more than low-potency TCS) after Week 16, by duration of previous treatment (16-28, 28-36, >36 weeks).

Summaries or listings of efficacy and patient-reported outcome endpoints from Study JAIX (2) open-label addendum will be provided separately from the main protocol population, as applicable.

5. Study Design

5.1. Summary of Study Design

Study JAIX is a Phase 3, multicenter, open-label, outpatient study evaluating the efficacy and safety of baricitinib 2-mg QD in adult patients with moderate to severe AD. The study is divided into 2 periods: a 200-week Open-Label Treatment period and a 4-week Post-Treatment Follow-Up period. In addition, this study includes an open-label addendum, for Study JAIX (2) open-label addendum (I4V-MC-JAIX(2) Clinical Protocol Addendum). The following patients will be eligible to participate in this study:

- Partial Responders (IGA 2 at Week 16 of Study JAIW)
- Previous Responders at Week 16 of Study JAIW (IGA 0 or 1): Patients who were responders at Week 16 of Study JAIW, and thus remain in Study JAIW but later experience loss of response resulting in an IGA ≥3 (or requiring more than low-potency topical corticosteroids [TCS] after Week 16 to manage symptoms)
- Nonresponders (IGA 3 or 4 at Week 16 of Study JAIW or rescued with topical [eg, TCS, topical calcineurin inhibitor (TCNI)] or systemic therapies prior to Week 16)
- Responders who completed Study JAIW through Week 104 (Visit 15)
- Up to approximately 30 patients are planned to be enrolled in Study JAIX (2) open-label addendum

5.2. Determination of Sample Size

The study is descriptive in nature and the sample size is not based on any statistical power calculations. All of the approximate 450 patients participating in the originating study can be assessed for eligibility to participate in Study JAIX. In addition, there are approximately 30 patients planned to be enrolled in Study JAIX (2) open-label addendum.

5.3. Method of Assignment to Treatment

This is an open-label trial where all patients will receive baricitinib 2-mg QD in the study JAIX and in Study JAIX (2) open-label addendum.

6. A Priori Statistical Methods

Statistical analyses of this study will be the responsibility of Eli Lilly and Company (Lilly) or its designee.

Any change to the data analysis methods described in the protocol will require an amendment ONLY if it changes a principal feature of the protocol. Any other change to the data analysis methods described in the protocol and the justification for making the change will be described in the CSR. Additional exploratory analyses of the data will be conducted as deemed appropriate for the Study JAIX and JAIX (2) open-label addendum.

6.1. General Considerations

As this study is open-label and all patients are receiving baricitinib 2-mg, no hypothesis tests for differences in treatment will be performed. Unless otherwise stated, data will be summarized both overall and by previous therapy received in the originating Study JAIW.

Analyses for Study JAIX (2) open-label addendum will be separate and will be analyzed descriptively, as deemed appropriate.

In general, summaries of efficacy and safety data will follow the following reporting rules:

- Cumulative summaries incorporating data from the originating Study JAIW baseline values to evaluate long-term results within Study JAIX.
- Summaries of Study JAIX alone to evaluate responses for those patients who are either up-titrating or receiving baricitinib for the first time.

All discrete efficacy and health outcome variables in the study JAIX and the study JAIX (2) open-label addendum, will be summarized using frequencies and percentages. Normal approximation to the binomial distribution will be used to construct a confidence interval for the proportion.

Continuous efficacy and health outcome variables will be summarized displaying number of patients (n), mean, standard deviation (SD), median, minimum, and maximum. Within-group least-squares (LS) mean changes and percent changes will be analyzed within the framework of mixed models repeated measures (MMRM) analysis models with baseline disease severity (IGA) in Study JAIW, effects for therapy in the originating Study JAIW, visit and therapy (in originating study)-by-visit interaction as fixed effects and baseline and baseline-by-visit as fixed continuous effects. An unstructured (co)variance structure will be used to model the between- and within-patient errors. If this analysis fails to converge, the heterogeneous autoregressive [ARH(1)], followed by the heterogeneous compound symmetry (CSH), followed by the heterogeneous Toeplitz (TOEPH), followed by autoregressive [AR(1)], followed by compound symmetry (CS) will be used. The p-values associated with the null hypothesis tests of LS mean=0 will be produced.

Frequencies and percentages will be computed for TEAE, SAEs, and discontinuations due to AEs or death.

All the efficacy endpoints will be summarized for Week 0 through Week 16 and can go to Week 104 for the study JAIX and to Week 52 for study JAIX (2) open-label addendum.

6.1.1. Analysis Populations

Modified Intent-to-Treat (mITT) population: The mITT population is defined as all patients who received at least 1 dose of investigational product in Study JAIX.

For the US submission safety lock in January 2020, efficacy analyses will not be performed since it is expected only part of the subjects would have reached Week 16.

For the 4-month safety update, efficacy summary will include following using observed values:

- Proportion of patients achieving EASI75 from baseline of originating Study JAIW, assessed at Weeks 0 to 16
- Proportion of patients achieving vIGA-AD of 0 or 1 assessed at Weeks 0 to 16
- Proportions of patients achieving a ≥4-point improvement in Itch NRS from baseline of originating Study JAIW, assessed at Weeks 0 to 16

Open-label Addendum Modified Intent-to-Treat (mITT) Population: The open-label addendum mITT population is defined as all patients who received at least 1 dose of investigational product in Study JAIX (2) open-label addendum.

Safety population: The safety population is defined as all patients who received at least 1 dose of investigational product in Study JAIX or in the study JAIX (2) open-label addendum, and who did not discontinue from the study for the reason 'Lost to Follow-up' at the first postbaseline visit in Study JAIX or in the study JAIX (2) open-label addendum. This definition excludes patients with no safety assessments postbaseline so that incidence rates are not underestimated.

Unless otherwise specified, the efficacy and health outcome summaries will be conducted on the mITT population. Safety analyses will be performed on the safety population.

Table JAIX.6.1. Efficacy and Safety Analyses Cohort

Cohort	Definition
Placebo/BARI 2-mg	PBO in Study JAIW switching to BARI 2-mg at entry to Study JAIX up to end of the study
BARI 1-mg/BARI 2-mg	BARI 1-mg in Study JAIW switching to BARI 2-mg at entry to Study JAIX up to end of the study
BARI 2-mg/BARI 2-mg	BARI 2-mg in Study JAIW staying at the same dose in Study JAIX up to end of the study
BARI 2-mg	BARI 2-mg in Study JAIX up to end of the study, all groups including BARI 2-mg in study JAIX (2) Open-label addendum
BARI 2-mg open-label addendum	BARI 2-mg for Study JAIX (2) open-label addendum

Abbreviations: BARI = baricitinib; PBO = placebo.

6.1.2. Definition of Baseline and Post-Baseline Measures

Baseline

The baseline value for the efficacy and health outcomes for the treatment period is defined as the last non-missing measurement on or prior to the date of first study drug administration in Study JAIW.

For the purpose of Study JAIX alone, the safety analyses or relevant analyses for the Study JAIX (2) open-label addendum will be conducted using the Study JAIX baseline, which is defined as the last non-missing measurement on or prior to the date of first study drug administration in Study JAIX.

The baseline value for the daily diary assessments (Itch NRS, ADSS, Skin Pain NRS, PGI-S-AD) is defined as the mean assessment of the non-missing assessments in the 7 days prior to the date of first study drug administration of the originating Study JAIW. If there are less than 4 non-missing assessments in the baseline diary window, the interval lower bound can be extended up to 7 additional days, 1 day at a time, to obtain the most recent 4 non-missing values. If there are not at least 4 non-missing assessments in the baseline period, the baseline mean is missing.

Post-Baseline

Post-baseline measurements are collected after study drug administration for electronic patient-reported outcome (ePRO), Itch NRS, Skin Pain NRS, ADSS, and PGI-S-AD up to Week 16 (Visit 4).

Non-missing efficacy data collected at scheduled visits (eg, electronic version of Clinical Outcome Assessment [eCOA], clinician-reported outcome [ClinRO]) will be used for analyses. If an assessment is missing at a scheduled visit, an unscheduled post-baseline assessment can be used, provided it falls within a ±4 day window of the scheduled visit date Visit 2 (Week 4) and Visit 3 (Week 8) and a ±7 day window to Visit 4 (Week 16) and all other subsequent visits. If there is more than 1 unscheduled visit within the defined visit window and no scheduled visit assessment is available, the unscheduled visit closest to the scheduled visit date will be used. If 2 unscheduled visits of equal distance are available, then the latter of the 2 will be used. If there is no non-missing measure collected at the scheduled visit, or an unscheduled visit falling within the visit window, the assessment is missing for that scheduled visit.

Postbaseline daily diary endpoints will be the mean of weekly visit windows (diary windows) anchored on day of first dose (Day 1) and day of Week 16 scheduled visit. These are defined as follows:

Week	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Day	1-7	8-14	15-21	22-28	29-35	36-42	43-49	50-56	57-63	64-70	71-77	78-84	85-91	92-98

Week 16 Daily Diary Window Construction

The following sequential steps will be used to determine the Week 16 diary window. The general goal is to anchor on the scheduled Week 16 visit (or a proximal unscheduled visit) if such a visit exists, or to use an interval based on days in study for cases where a scheduled Week 16 or a proximal surrogate does not exist.

Step 1: If the Week 16 scheduled visit exists, the Week 16 diary interval is the 7 days prior to the Week 16 date provided that window has at least 4 non-missing observations. If there are less than 4 non-missing observations, the diary window's lower bound will be extended 1 day at a time (up to day 99) to a maximum of 14 days prior to the Week 16 date until 4 non-missing observations are obtained. If, after extending this diary window's lower bound to 14 days, there are less than 4 non-missing observations, then go to Step 3.

Step 2: If the Week 16 scheduled visit does not exist, the 7 days prior to the last visit (scheduled or unscheduled) occurring after Day 105 and before Visit 8 will constitute the Week 16 diary window provided that window contains at least 4 non-missing observations. If there are less than 4 non-missing observations, the diary window's lower bound will be extended 1 day at a time (up to Day 99) to a maximum of 14 days prior to the unscheduled visit date until 4 non-missing observations are obtained. If, after extending this diary window's lower bound to 14 days, there are less than 4 non-missing observations, then go to Step 3.

Step 3: If neither a Week 16 scheduled visit is available or an unscheduled visit to act as a surrogate for the Week 16 diary window, then the Week 16 window will be Day 106 to Day 112. If there are less than 4 non-missing observations, the diary window's lower bound will be extended 1 day at a time to Day 99 until 4 non-missing observations are obtained. If the steps above do not detect a window with at least 4 non-missing observations, then the Week 16 window is 7 days from either the Week 16 visit, the surrogate visit, or Day 106 to Day 112 and the mean is missing and subject to imputation rules.

Week 15 Daily Diary Window Construction

The lower boundary of the Week 15 diary window is defined as Day 99. The upper bound of the Week 15 diary window is the minimum of either Day 105 or the lower bound of the Week 16 diary window -1. Consequently, Week 15 may be less than 4 days if the Week 16 scheduled visit is before Day 112. Moreover, as the Week 15 diary window cannot exceed 7 days, there could be daily assessments between Weeks 15 and 16 diary windows that do not fall into a diary window. If, after constructing the diary windows, there are fewer than 4 non-missing values, the mean for Week 15 is missing and subject to imputation rules.

Handling of Duplicate Diary Records

If there is more than 1 diary record on a particular date, the first record on that particular date will be used in the analysis.

6.1.3. Derived Data

- Age (year), derived using first dose date as the reference start date and July 1 of birth year, and truncated to a whole-year (integer) age. Patients whose derived age is less than 18 will have the required minimum age of 18 at informed consent confirmed. Reporting for age, age groups, and laboratory ranges, however will be based on their derived age.
- Age group (<65, ≥65 years old)
- Age group (<65, ≥65 to <75, ≥75 to <85, ≥85 years old)
- Body mass index (BMI) (kg/m^2) = Weight (kg)/((Height (cm)/100)^2)
- BMI category (<25 kg/m^2, ≥25 to <30 kg/m^2, ≥30 kg/m^2)
- The duration of AD from diagnosis (year) = [(Date of informed consent – Date of AD diagnosis)+1]/365.25.
  If year of onset is missing, duration of AD will be set as missing. Otherwise, unknown month will be taken as January, and unknown day will be taken as 01. The duration of AD will be rounded to 1 decimal place.
- Duration of AD (years) category (0 to <2 years, ≥2 to <5 years, ≥5 to <10 years, ≥10 to <20 years, ≥20 years)
- Diagnosis age (number of months between date of AD diagnosis and July 1 of birth year) /12, and truncated to a whole-integer age
- Diagnosis age group (<18, ≥18 and <50, ≥50 years old)
- Change from baseline = post-baseline measurement at Visit x – baseline measurement.
  If a baseline value is missing, it will not be imputed and the change from baseline will not be calculated.
- Percent change from baseline at Visit x:

  ((Post-baseline measurement at Visit x - Baseline measurement)/Baseline measurement)*100.

  If a baseline value is missing, it will not be imputed and percent change from baseline will not be calculated.

- Weight (kg) = weight (lbs)*0.454.
- Weight category (<60 kg, ≥60 to <100 kg, ≥100 kg)
- Height (cm) = height (in)*2.54.
- Cyclosporine inadequate efficacy response(yes, no)
  - Set **yes** if the reason for discontinuation is inadequate response
- Cyclosporine intolerance (yes, no)
  - Set **yes** if the reasons for discontinuation are intolerance to medication or contraindication (physician-indicated cyclosporine was used and a contraindication was noted)
- Cyclosporine contraindication [ineligible] (yes, no)
  - Set to **yes** if cyclosporine never used because of a contraindication
- Cyclosporine inadvisable (yes, no)
  - Set to **yes** if the following reasons were selected for either not using the medication or discontinuing the medication:

        - Reason for not using medication: contraindication, unfavorable benefit/risk, or physician decision
        - Reason for discontinuation: inadequate response, intolerance to medication, or contraindication.
- TCNI Topical calcineurin inhibitor inadequate efficacy response (yes, no)
    - Set **yes** if the reason for discontinuation is inadequate response
- TCNI Topical calcineurin inhibitor intolerance (yes, no)
    - Set **yes** if the reasons for discontinuation are intolerance to medication or contraindication (physician-indicated TCNI was used and a contraindication was noted)
- TCNI Topical calcineurin inhibitor contraindication / [ineligible](yes, no)
    - Set to **yes** if TCNI never used because of a contraindication
- TCNI Topical calcineurin inhibitor inadvisable (yes, no)
    - Set to **yes** if the following reasons were selected for either not using the medication or discontinuing the medication:
        - Reason for not using medication: Physician decision, concern about side effects, unfavorable benefit risk, contraindication
        - Reasons for discontinuation: inadequate response, intolerance to medication, or contraindication

6.2. Adjustments for Covariates

Disease severity (IGA), baseline value, and baseline-by-visit interactions will be included as covariates when MMRM analysis is performed.

6.3. Handling of Dropouts or Missing Data

Patients who discontinue the study or the study treatment at any time for any reason will be defined as nonresponders for the nonresponder imputation (NRI) summary of categorical efficacy variables such as IGA 0/1 or EASI50/75/90 after discontinuation and onward. No imputation methods will be applied to continuous data. Missing value will be imputed by MMRM for continues efficacy variables.

6.3.1. Missingness Due to COVID-19

For missingness that COVID-19 is identified as the cause by the study team, the following imputation methods would be applied.

- For continuous measures, if missing due to COVID-19,
    - When they are analyzed using analysis of covariance (ANCOVA) for a single timepoint, modified last observation carried forward (mLOCF) method will be used.
- For categorical variables, if missing due to COVID-19,
    - mLOCF method will be used. This applies to categorical variables that are collected categorically, as well as categorical variables that are derived from measures that are collected as continuous, such as EASI50, EASI75 and EASI90.

- In addition, for primary and key efficacy measures including IGA 0 or 1 and EASI75, multiple imputation (MI) method will be used. Seed = 123456 will be used for MI. Results from MIs will be aggregated to generate the final estimates for statistical inferences. This will be the main imputation method for these primary and key efficacy measures for visits with missing data due to COVID-19.

All measures collected through ePRO diary device should not be impacted by COVID-19.

For a specific database lock, if COVID-19 is still ongoing at the time of the database lock, the list of visits with missing data due to COVID-19 identified by the study for that lock may not be complete and may be updated in future locks. The imputation for missingness due to COVID-19 for each database lock is based on the list we receive for that specific lock.

For the lock to occur on 02 June 2020, as there are very limited missing data due to COVID-19, the imputation method specified in this section (Section 6.3.1) will not be applied to this lock.

6.4. Multicenter Studies

This study will be conducted by multiple investigators at multiple sites in North America. As the summaries of these data are descriptive in nature, no adjustments for site differences will be made.

6.5. Multiple Comparisons/Multiplicity

No multiplicity control measures will be used.

6.6. Patient Disposition

Patient disposition will be summarized for the mITT population for JAIX study. Frequency counts and percentages of patients who complete the study treatment visits or discontinue early from the study will be summarized by prior treatment group and overall.

Patient disposition will be summarized for the mITT population for JAIX (2) open-label addendum study.

A listing of patient disposition will be provided for all patients, with the extent of their participation in the study JAIX and the study JAIX (2) open-label addendum and the reason for discontinuation.

6.7. Patient Characteristics

Patient characteristics including demographics and baseline characteristics will be summarized descriptively by previous treatment group and overall for the mITT population for JAIX study. Additionally, historical illnesses and pre-existing conditions will be summarized. Descriptive statistics including n, mean, SD, minimum, 1st quartile, median, 3rd quartile, and maximum will be provided for continuous measures, and frequency counts and percentages will be tabulated for categorical measures.

Patient characteristics will also be summarized for the mITT population for JAIX (2) open-label addendum study.

6.7.1. Patient Demographics

Patient demographics will be summarized as described above for both in the study JAIX and the study JAIX (2) open-label addendum. The following demographic information will be included:

- Age
- Age group (<65 vs. ≥65)
- Age group (<65, ≥65 to <75, ≥75 to <85, ≥85)
- Sex (male, female)
- Race (American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, White, Multiple)
- Weight (kg)
- Weight category (<60 kg, ≥60 to <100 kg, ≥100 kg)
- Height (cm)
- BMI (kg/m^2)
- BMI category (<25 kg/m^2, ≥25 to <30 kg/m^2, ≥30 kg/m^2)

A listing of patient demographics will also be provided for the enrolled population.

6.7.2. Baseline Disease Characteristics

The following baseline disease information (but not limited to only these) will be categorized and presented for baseline AD clinical characteristics, baseline health outcome measures, and other baseline demographic and disease characteristics as described above:

- Duration since AD diagnosis (years)
- Duration since AD diagnosis category (0 to <2 years, 2 to <5 years, 5 to <10 years, 10 to <20 years, ≥20 years)
- Age at diagnosis (years)
- Age group at diagnosis (<18 years, ≥18 to <50 years, ≥50 years)
- Habits (Alcohol: Never, Current, Former; Tobacco: Never, Current, Former)
- Skin infections treated with a pharmacological agent within past year (yes, no, unknown; number if yes)
- Atopic Dermatitis flares within past year (yes, no, unknown; number if yes)
- Validated Investigator's Global Assessment for AD score
- Eczema Area and Severity Index (EASI) score
- SCORing Atopic Dermatitis (SCORAD)
- Body Surface Area (BSA) affected by AD
- Hospital Anxiety Depression Scale (HADS) subscales
- Patient-Oriented Eczema Measure (POEM)
- Itch Numerical Rating Scale (NRS)
- Atopic Dermatitis Sleep Scale (ADSS) Item 2
- Dermatology Life Quality Index (DLQI)
- Skin Pain NRS

- Patient Global Impression of Severity (PGI-S-AD)
- prior therapy (topical therapy only, systemic therapy)
- prior use of Cyclosporine (yes, no)
- Cyclosporine inadequate response (yes, no)
- Cyclosporine intolerance (yes, no)
- Cyclosporine contraindication [ineligible] (yes, no)
- Cyclosporine inadvisable (yes, no)
- Prior use of TCNI (yes, no)
- TCNI inadequate response (yes, no)
- TCNI intolerance (yes, no)
- TCNI contraindication [ineligible] (yes, no)
- TCNI inadvisable (yes, no)
- Vaccine:
  - Zoster vaccine (Yes, No)
  - Tuberculosis vaccine (Yes, No)
- Baseline renal function status: impaired (estimated glomerular filtration rate [eGFR] <60 mL/min/1.73 m^2) or not impaired (eGFR ≥60 mL/min/1.73 m^2)
- Immunoglobulin E: intrinsic (<200 kU/I) or extrinsic (≥200 kU/I)

6.7.3. Historical Illnesses and Pre-existing Conditions

Historical illnesses are recorded in the originating Study JAIW and not in the Study JAIX electronic case report form (eCRF).

Pre-existing conditions are defined as those conditions recorded in the Pre-existing Conditions and Medical History eCRF, the Prespecified Medical History: Comorbidities eCRF, or the Adverse Events eCRF with a start date prior to the date of the Study JAIX informed consent and an end date after informed consent or missing. For events occurring on the day of the first dose of previous study, the date and time of the onset of the event will both be used to determine if the event was pre-existing. Conditions with a partial or missing start date (or time if needed) will be assumed to be 'not pre-existing' unless there is evidence, through comparison of partial dates, to suggest otherwise. Pre-existing conditions will be categorized using the Medical Dictionary for Regulatory Activities (MedDRA) Standardized MedDRA Queries (SMQs) or similar predefined lists of Preferred Terms (PTs) of interest. Frequency counts and percentages of patients with selected pre-existing conditions will be summarized for the mITT population.

6.8. Treatment Compliance

Patient compliance with study medication will be assessed using the mITT population.

A patient is considered noncompliant if he or she misses >20% of the prescribed doses during the study, unless the patient's study drug is withheld by the investigator. 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 calculation.

Compliance in the period of interest up to Visit *x* will be calculated as follows:

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

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

Descriptive statistics for percent compliance and non-compliance rate will be summarized for the mITT population for Week 0 through Week 16 and can go to Week 200, as deemed appropriate. Sub-intervals of interest, such as compliance between visits, may also be presented. The number of expected doses, tablets dispensed, tablets returned, and percent compliance will be listed by patient for both in the study JAIX and the study JAIX (2) open-label addendum.

6.9. Previous and Concomitant Therapy

Previous medications are described in the originating Study JAIW.

Summaries of concomitant medications will be based on the mITT population.

Concomitant therapy for the treatment period is defined as therapy that starts before or during the treatment period of the previous study and ends during the Study JAIX treatment period or is ongoing (has no end date or ends after the treatment period). Should there be insufficient data to make this comparison (eg, the concomitant therapy stop year is the same as the treatment start year, but the concomitant therapy stop month and day are missing), the medication will be considered as concomitant for the treatment period. Medications started during Study JAIX are also considered as concomitant medications.

Summaries of concomitant medications will be provided for the following category:

- Summary of concomitant medications excluding background therapy
- Summary of background therapy

6.10. Efficacy Analyses

The general methods used to summarize efficacy variables is given in Section 6.1.
Table JAIX.6.2 includes the descriptions and derivations of the primary, secondary, and exploratory efficacy outcomes in the study JAIX and the study JAIX (2) open-label addendum.

Table JAIX.6.2. Description and Derivation of Primary, Secondary, and Exploratory Efficacy Outcomes

Measure	Description	Variable	Derivation/Comment	Imputation Approach if Missing Components
Eczema Area and Severity Index (EASI)	The EASI assesses objective physician estimates of 2 dimensions of AD, disease extent and clinical signs (Hanifin et al. 2001), by scoring the extent of disease (percentage of skin affected: 0 = 0%; 1 = 1-9%; 2 = 10-29%; 3 = 30-49%; 4 = 50-69%; 5 = 70-89%; 6 = 90-100%) and the severity of 4 clinical signs (erythema, edema/papulation, excoriation, and lichenification), each on a scale of 0-3 (0 = none, absent; 1 = mild; 2 = moderate; 3 = severe) at 4 body sites (head and neck, trunk, upper limbs, and lower limbs). Half scores are allowed between severities 1, 2 and 3. Each body site will have a score that ranges from 0-72, and the final EASI score will be obtained by weight averaging these 4 scores. Hence, the final EASI score will range from 0-72 for each time point.	▪ EASI score	Derive EASI region score for each of head and neck, trunk, upper limbs, and lower limbs as follows: $EASI_{region}$ = (Erythema + edema/papulation + Excoriation + Lichenification)*(value from percentage involvement), where erythema, edema/papulation, excoriation, and lichenification are evaluated on a scale of 0-3 and value from percentage involvement is on a scale of 0-6. Then total EASI score is as follows: $EASI = 0.1*EASI_{head\ and\ neck} + 0.3*EASI_{trunk} + 0.2*EASI_{upper\ limbs} + 0.4*EASI_{lower\ limbs}$	N/A – partial assessments cannot be saved.
		▪ EASI50	% improvement in EASI score from baseline ≥50%: % change from baseline ≤-50	Missing if baseline or observed value is missing
		▪ EASI75	% improvement in EASI score from baseline ≥75%: % change from baseline ≤-75	Missing if baseline or observed value is missing
		▪ EASI90	% improvement in EASI score from baseline ≥90%: % change from baseline ≤-90	Missing if baseline or observed value is missing
		▪ Change from baseline in EASI score ▪ Percent change from baseline EASI score	Change from baseline: observed EASI score – baseline EASI score % change from baseline: $100 \times \frac{Observed\ score - Baseline}{Baseline}$	Missing if baseline or observed value is missing

Description and Derivation of Primary, Secondary, and Exploratory Efficacy Outcomes

Measure	Description	Variable	Derivation/Comment	Imputation Approach if Missing Components
Body Surface Area (BSA) Affected by AD	BSA affected by AD will be assessed for 4 separate body regions and is collected as part of the EASI assessment: head and neck, trunk (including genital region), upper extremities, and lower extremities (including the buttocks). Each body region will be assessed for disease extent ranging from 0% to 100% involvement. The overall total percentage will be reported based off of all 4 body regions combined, after applying specific multipliers to the different body regions to account for the percent of the total BSA represented by each of the 4 regions.	▪ BSA score	Use the percentage of skin affected for each region (0 to 100%) in EASI as follows: BSA Total = 0.1*BSAhead and neck + 0.3*BSAtrunk + 0.2* BSAupper limbs + 0.4*BSAlower limbs	N/A – partial assessments cannot be saved.
		Change from baseline in BSA score	Change from baseline: observed BSA score – baseline BSA score	Missing if baseline or observed value is missing.
		▪ Proportion of patients achieving a BSA of ≤3%	Observed score of ≤3% for patients with BSA >3% or missing at baseline	Missing if observed value is missing.
Validated Investigator's Global Assessment for AD (IGA)	The validated IGA based on a static, numeric 5-point scale from 0 (clear) to 4 (severe). The score is based on an overall assessment of the degree of erythema, papulation/induration, oozing/crusting, and lichenification.	▪ IGA score	Single item. Range: 0 to 4 0 represents "clear" 4 represents "severe"	Single item, missing if missing.
		▪ IGA[0,1] with ≥2-point improvement	Observed score of 0 or 1 and change from baseline ≤-2	Missing if baseline or observed value is missing.

Description and Derivation of Primary, Secondary, and Exploratory Efficacy Outcomes

Measure	Description	Variable	Derivation/Comment	Imputation Approach if Missing Components
SCORing Atopic Dermatitis (SCORAD)	The SCORAD index uses the rule of nines to assess disease extent (head and neck 9%; upper limbs 9% each; lower limbs 18% each; anterior trunk 18%; back 18%; and genitals 1%). It evaluates 6 clinical characteristics to determine disease severity: (1) erythema, (2) edema/papulation, (3) oozing/crusts, (4) excoriation, (5) lichenification, and (6) dryness on a scale of 0 to 3 (0=absence, 1=mild, 2=moderate, 3=severe). The SCORAD index also assesses subjective symptoms of pruritus and sleep loss in the last 72 hours on visual analog scales of 0 to 10, where 0 is no itch or sleep loss and 10 is worst imaginable itch or sleep loss. The 3 aspects of extent of disease, disease severity, and subjective symptoms combine to give a maximum possible score of 103 (Stalder et al. 1993; Kunz et al. 1997; Schram et al. 2012).	▪ SCORAD score	SCORAD = A/5 + 7B/2 + C, where A is extent of disease, range 0-100 B is disease severity, range 0-18 C is subjective symptoms, range 0-20	Missing if components A and B are missing or if component C is missing. Partial assessments performed by physician cannot be saved and partial assessments performed by subject cannot be saved.
		▪ Change from baseline in SCORAD score	Change from baseline: observed SCORAD score – baseline SCORAD score	Missing if baseline or observed value is missing

Abbreviations: AD = atopic dermatitis; N/A = not applicable.

6.10.1. Primary Outcome and Methodology

The primary analysis of the study is to summarize the baricitinib 2-mg patients in the proportion of patients achieving EASI75 from baseline at Week 16 for the mITT population. Missing data will be imputed using the NRI method described in Section 6.3.

6.10.2. Secondary and Exploratory Efficacy Analyses

Secondary efficacy analyses are detailed in Table JAIX.6.2. Secondary and exploratory endpoints will be summarized as described in Section 6.1. Missing EASI50/75/90 and responses will be imputed using the NRI method described in Section 6.3. Exploratory, health outcomes and quality-of-life measures endpoint may be analyzed, as deemed appropriate.

6.11. Health Outcomes/Quality-of-Life Analyses

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

Health outcomes and quality-of-life measures will generally be analyzed according to the formats discussed in Section 6.10.2.

Table JAIX.6.3 includes the descriptions and derivations of the health outcomes and quality-of-life measures.

Table JAIX.6.3. Description and Derivation of Health Outcomes and Quality-of-Life Measures

Measure	Description	Variable	Derivation/Comment	Imputation Approach if with Missing Components
Itch Numeric Rating Scale (NRS)	The Itch NRS is a patient-administered, 11-point horizontal scale anchored at 0 and 10, with 0 representing "no itch" and 10 representing "worst itch imaginable." Overall severity of a patient's itching is indicated by selecting the number that best describes the worst level of itching in the past 24 hours (Naegeli et al. 2015; Kimball et al. 2016). Refer to Section 6.1.2 for details on how to calculate the weekly score that will be used in the continuous analysis.	Itch NRS score	Single item; range 0-10. Refer to Section 6.1.2 for how to derive the visit score.	Refer to Section 6.1.2 for how to derive the visit score.
		▪ Change from baseline in Itch NRS ▪ Percent change from baseline in Itch NRS	Change from baseline: observed itch score – baseline itch score % change from baseline: $100 \times \frac{Observed\ score - Baseline}{Baseline}$	Missing if baseline or observed value is missing
		▪ 4-point itch improvement in subgroup of patients with baseline Itch NRS ≥4	Change from baseline ≤-4 and baseline ≥4	Missing if baseline is missing or <4 or observed value is missing
Skin Pain NRS	The Skin Pain NRS is a patient-administered, 11-point horizontal scale anchored at 0 and 10, with 0 representing "no pain" and 10 representing "worst pain imaginable." Overall severity of a patient's skin pain is indicated by selecting the number that best describes the worst level of skin pain in the past 24 hours Refer to Section 6.1.2 for details on how to calculate the weekly score that will be used in the continuous analysis.	Skin Pain NRS score	Single item; range 0-10. Refer to Section 6.1.2 for how to derive the visit score.	Refer to Section 6.1.2 for how to derive the visit score.
		▪ Change from baseline in Skin Pain NRS	Change from baseline: observed skin pain score – baseline skin pain score	Missing if baseline or observed value is missing
		▪ 4-point Skin Pain improvement in subgroup of patients with baseline Skin Pain NRS ≥4	Change from baseline ≤-4 and baseline ≥4	Missing if baseline is missing or <4 or observed value is missing

Description and Derivation of Health Outcomes and Quality-of-Life Measures

Measure	Description	Variable	Derivation/Comment	Imputation Approach if with Missing Components
Atopic Dermatitis Sleep Scale (ADSS)	The ADSS is a 3-item, patient-administered questionnaire developed to assess the impact of itch on sleep including difficulty falling asleep, frequency of waking, and difficulty getting back to sleep last night. Patients rate their difficulty falling asleep and difficulty getting back to sleep, items 1 and 3, respectively, using a 5-point Likert-type scale with response options ranging from 0 "not at all" to 4 "very difficult." Patients report their frequency of waking last night, item 2, by selecting the number of times they woke up each night, ranging from 0 to 29 times. The ADSS is designed to be completed each day with respondents thinking about sleep "last night." Each item is scored individually. Refer to Section 6.1.2 for details on how to calculate the weekly score which will be used in the continuous analysis.	▪ Item 1 score of ADSS ▪ Item 2 score of ADSS ▪ Item 3 score of ADSS	Single items: Item 1, range 0 to 4; Item 2, range 0 to 29; Item 3, range 0 to 4. Refer to Section 6.1.2 for how to derive the visit score.	Refer to Section 6.1.2 for how to derive the visit score.
		▪ Change from baseline in score of Item 1 of ADSS ▪ Change from baseline in score of Item 2 of ADSS ▪ Change from baseline in score of Item 3 of ADSS	Change from baseline: observed ADSS item score – baseline ADSS item score	Missing if baseline or observed value is missing.

Description and Derivation of Health Outcomes and Quality-of-Life Measures

Measure	Description	Variable	Derivation/Comment	Imputation Approach if with Missing Components
Patient-Oriented Eczema Measure (POEM)	The POEM is a simple, 7-item, patient-administered scale that assesses disease severity in children and adults. Patients respond to questions about the frequency of 7 symptoms (itching, sleep disturbance, bleeding, weeping/oozing, cracking, flaking, and dryness/roughness) over the last week. Response categories include "No days," "1-2 days," "3-4 days," "5-6 days," and "Every day" with corresponding scores of 0, 1, 2, 3, and 4, respectively. Scores range from 0-28 with higher total scores indicating greater disease severity (Charman et al. 2004).	POEM score	POEM total score: sum of questions 1 to 7, Range 0 to 28.	If a single question is left unanswered, then that question is scored as 0. If more than 1 question is unanswered, then the tool is not scored. If more than 1 response is selected, then the response with the highest score is used.
		▪ Change from baseline in POEM score	Change from baseline: observed POEM score – baseline POEM score	Missing if baseline or observed value is missing.

Description and Derivation of Health Outcomes and Quality-of-Life Measures

Measure	Description	Variable	Derivation/Comment	Imputation Approach if with Missing Components
Dermatology Life Quality Index (DLQI)	The DLQI is a simple, patient-administered, 10-item, validated, quality-of-life questionnaire that covers 6 domains including symptoms and feelings, daily activities, leisure, work and school, personal relationships, and treatment. The recall period of this scale is over the "last week." Response categories include "a little," "a lot," and "very much," with corresponding scores of 1, 2, and 3, respectively, and "not at all," or unanswered ("not relevant") responses scored as 0. Scores range from 0-30 with higher scores indicating greater impairment of quality of life. A DLQI total score of 0 to 1 is considered as having no effect on a patient's health-related QoL (Hongbo et al. 2005) and a 4-point change from baseline is considered as the minimal clinically important difference threshold (Khilji et al. 2002; Basra et al. 2015).	Symptoms and feelings domain	Sum of questions 1 and 2, range 0 to 6.	N/A – partial assessments cannot be saved.
		Daily activities domain	Sum of questions 3 and 4, range 0 to 6.	N/A – partial assessments cannot be saved.
		Leisure domain	Sum of questions 5 and 6, range 0 to 6.	N/A – partial assessments cannot be saved.
		Work and school domain	Sum of questions 7 and 7B (if it is answered), range 0 to 3. Responses of "yes" and "no" on Question 7 are given scores of 3 and 0 respectively. If Question 7 is answered "no" then Question 7b is answered with "a lot", "a little", "not at all" getting scores of 2, 1, 0 respectively.	N/A – partial assessments cannot be saved.
		Personal relationships domain	Sum of questions 8 and 9, range 0 to 6.	N/A – partial assessments cannot be saved.
		Treatment domain	Question 10, range 0 to 3.	N/A – partial assessments cannot be saved.
		DLQI total score	DLQI total score: sum of all six DLQI domain scores, range 0 to 30.	N/A – partial assessments cannot be saved.
		Change from baseline in DLQI	Change from baseline: observed DLQI score – baseline DLQI score	Missing if baseline or observed value is missing.

Description and Derivation of Health Outcomes and Quality-of-Life Measures

Measure	Description	Variable	Derivation/Comment	Imputation Approach if with Missing Components
Dermatology Life Quality Index (DLQI) (continued)		Proportions of patients achieving a DLQI score of 5 or less for those with baseline DLQI >5 of originating study	Post-baseline DLQI total score ≤5 with baseline total score >5	Missing if baseline is missing or ≤5 or observed value is missing
		DLQI 0 or 1	Post-baseline DLQI total score in (0,1)	N/A – partial assessments cannot be saved.

Abbreviations: N/A = not applicable; QoL = quality of life.

6.12. Safety Analyses

Detailed analyses and discussion of Study JAIX safety data are more thoroughly assessed in the context of combining the safety data from the originating Study JAIW with the safety data from Study JAIX, and will be presented in an integrated submission document and not in the JAIX CSR. Those planned analyses are described in the PSAP.

For the purpose of Study JAIX and JAIX (2) open-label addendum, the following are planned to be analyzed:

- TEAEs by PT nested within System Organ Class (SOC)
- SAEs by PT nested within SOC, and listing of SAEs including deaths, if any
- AEs leading to permanent study drug discontinuation and temporarily interruption to study drug by PT nested within SOC, and listing of AEs leading to permanent study drug discontinuation

The general methods used to summarize safety data, including the definition of baseline value, are described in Section 6.1.

For selected safety assessments, descriptive statistics may be presented for the last measure observed during post-treatment follow-up (up to 30 days after the last dose of treatment including rescue, regardless of study period).

For all the safety analysis, data from the study JAIX and JAIX (2) open-label addendum study will be presented together.

6.12.1. Extent of Exposure

Duration of exposure (in days) to study drug will be summarized for the safety population using descriptive statistics (n, mean, SD, minimum, 1st quartile, median, 3rd quartile, maximum). 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:

- Duration of exposure to investigational product: date of last dose of treatment– date of first dose of study drug + 1.

Last dose of treatment is calculated as last date on study drug.

Total patient-years (PY) of exposure will be reported using descriptive statistics (n, mean, SD, minimum, 1st quartile, median, 3rd quartile, and maximum) will be provided as patient-days of exposure and the frequency of patients falling into different exposure ranges will be summarized. Exposure ranges will generally be reported in weeks using the following as a general guide:

- ≥28 days, ≥56 days, ≥112 days, ≥168 days, ≥252 days, ≥364 days, ≥448 days, 532 days, ≥616 days, ≥728 days, ≥840 days, ≥952 days, ≥1064 days, ≥1176 days, ≥1288 days, and ≥1400 days
- >0 to <28 days, ≥28 to <56 days, ≥56 to <112 days, ≥112 to <168 days, ≥168 to <252 days, ≥252 to <364 days, ≥364 to <448 days, ≥448 to <532 days, ≥532 to <616 days,

≥616 to <728 days, ≥728 to <840 days, ≥840 to <952 days, ≥952 to <1064 days, ≥1064 to <1176 days, ≥1176 to <1288 days, ≥1288 to <1400 days, and ≥1400 days

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

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

6.12.2. Adverse Events

Adverse events are recorded in the eCRFs. Each AE will be coded to SOC and PT using the MedDRA version that is current at the time of database lock. Severity of AEs is recorded as mild, moderate, or severe.

A TEAE is defined as an event that either first occurred or worsened in severity after the first dose of study treatment in Study JAIX and JAIX (2) open-label addendum study 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.

Adverse events are classified based on the MedDRA PT. The MedDRA Lowest Level Term (LLT) will be used in defining which events are treatment-emergent. The maximum severity for each LLT during the baseline period up to first dose of the study medication in Study JAIX will be used as baseline. If an event with missing severity is preexisting during the baseline period, and persists during the treatment period, then the baseline severity will be considered mild for determining TEAE (that is, the event is treatment-emergent if the severity is coded moderate or severe postbaseline and not treatment-emergent if the severity is coded mild postbaseline). If an event occurring postbaseline has a missing severity rating, then the event is considered treatment-emergent unless the baseline is severe, in which case the event is not treatment-emergent. The day and time for events where onset is on the day of the first dose of study treatment will both be used to distinguish between pretreatment and posttreatment in order to derive treatment-emergence. Should there be insufficient data for AE start date to make this comparison (eg, the AE start year is the same as the treatment start year, but the AE start month and day are missing), the AE will be considered treatment-emergent.

In general, summaries will include the number of patients in the safety population (N), frequency of patients experiencing the event (n), and percentage n/N*100. For any events that are gender-specific based on the displayed PT, the denominator used to compute the percentage will only include patients from the given gender. The number and percentage of patients with TEAEs will be summarized by MedDRA PT nested within SOC with decreasing frequency in SOC, and events ordered within each SOC by decreasing frequency in the baricitinib 2-mg group.

6.12.3. Common Adverse Events

Common TEAEs are defined as TEAEs that occurred in ≥2% (before rounding) of patients. Common TEAEs will be analyzed in the context of combining the safety data from the originating Study JAIW with the safety data from Study JAIX. The planned analyses are described in the PSAP and will not be analyzed for Study JAIX alone.

6.12.4. Serious Adverse Events

Consistent with the International Conference on Harmonisation (ICH) E2A guideline (1994) and 21 Code of Federal Regulations (CFR) 312.32 (a) (2010), a SAE is any AE that results in any one of the following outcomes:

- Death
- Initial or prolonged inpatient hospitalization
- A life-threatening experience (that is, immediate risk of dying)
- Persistent or significant disability/incapacity
- Congenital anomaly/birth defect

Important medical events that may not be immediately life-threatening or result in death or hospitalization but may jeopardize the patient or may require intervention to prevent one of the other outcomes listed in the definition above.

The number and percentage of patients who experienced any SAE will be summarized using MedDRA PT nested within SOC.

An individual listing of all SAEs will be provided. A listing of deaths, regardless of when they occurred during the study, will also be provided.

6.12.5. Other Significant Adverse Events

Other significant AEs to be summarized will provide the number and percentage of patients who permanently discontinued study drug because of an AE or death; by treatment using MedDRA PT nested within SOC. Events will be ordered by decreasing frequency in the SOC.

A listing of all AEs leading to permanent discontinuation from the study drug or from the study will be provided.

6.13. 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 number and percentage of patients having important protocol deviations (IPDs) will be summarized within category and subcategory of deviation using the mITT population. Individual patient listings of IPDs will be provided.

6.14. Planned Exploratory Analyses

Exploratory objectives are listed in Section 4.3 and their associated analyses are described in Sections 6.10 and 6.11.

6.15. Interim Analyses

The first interim database lock is a safety lock to support US Food and Drug Administration (FDA) submission. It is anticipated that at the time of data cutoff, only some patients will be

able to reach the end of the Week 16 (the primary endpoint); thus, efficacy analyses will not be included. The summary on safety analyses will be up to the date cut for all safety population.

Additional interim analyses may be performed to support global regulatory submissions and reviews.

6.16. Annual Report Analyses

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

6.17. 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 a 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 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 experienced each event term
  - the number of events experienced.
- Consistent with 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).
- AE reporting is consistent with other document disclosures (eg, CSR, manuscripts).

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

7. References

Basra MK, Salek MS, Camilleri L, Sturkey R, Finlay AY. Determining the minimal clinically important difference and responsiveness of the Dermatology Life Quality Index (DLQI): further data. *Dermatology*. 2015;230(1):27-33.

[CFR] US National Archives and Records Administration. Code of Federal Regulations (CFR). Investigational New Drug Application (IND) safety reporting. 2010;Title 21:Section 312.32. Available at: https://www.ecfr.gov/cgi-bin/text-idx?SID=fd87d5f6fb00b95672dbe924a6d3b2f0&mc=true&node=se21.5.312_132&rgn=div8. Accessed July 21, 2017.

Charman CR, Venn AJ, Williams, HC. The Patient-Oriented Eczema Measure: development and initial validation of a new tool for measuring atopic eczema severity from the patients' perspective [published correction appears in *Arch Dermatol*. 2005;141(3):381]. *Arch Dermatol.* 2004;140(12):1513-1519.

Hanifin JM, Thurston M, Omoto M, Cherill R, Tofte SJ, Graeber M. The eczema area and severity index (EASI): assessment of reliability in atopic dermatitis. EASI Evaluator Group. *Exp Dermatol*. 2001;10(1):11-18.

Hongbo Y, Thomas CL, Harrison MA, Salek MS, Finlay AY. Translating the science of quality of life into practice: What do dermatology life quality index scores mean? *J Invest Dermatol*. 2005;125(4):659-664.

[ICH] International Conference on Harmonisation. Harmonised Tripartite Guideline: Clinical safety data management: definitions and standards for expedited reporting. E2A. 1994;Step 4. Available at: http://www.ich.org/products/guidelines/efficacy/article/efficacy-guidelines.html. Accessed July 21, 2017.

Khilji FA, Gonzalez M, Finlay AY. Clinical meaning of change in Dermatology Life Quality Index scores. *Br J Dermatol*. 2002;147(suppl 62):25-54.

Kimball AB, Naegeli AN, Edson-Heredia E, Lin CY, Gaich C, Nikai E, Yosipovitch G. Psychometric properties of the Itch Numeric Rating Scale in patients with moderate-to-severe plaque psoriasis. *Br J Dermatol.* 2016;175(1):157-162.

Kunz B, Oranje A, Labrèze, Stadler JF, Ring J, Taïeb A. Clinical validation and guidelines for the SCORAD index: consensus report of the European Task Force Atopic Dermatitis. *Dermatology*. 1997;195(1):10-19.

Naegeli AN, Flood E, Tucker J, Devlen J, Edson-Heredia E. The Worst Itch Numeric Rating Scale for patients with moderate to severe plaque psoriasis or psoriatic arthritis. *Int J Dermatol.* 2015;54(6):715-722.

Schram ME, Spuls PI, Leeflang MMG, Lindeboom R, Bos JD, Schmitt J. EASI, (objective) SCORAD and POEM for atopic eczema: responsiveness and minimal clinically important difference. *Allergy.* 2012;67(1):99-106.

Stalder J, Taieb A, Atherton D, Bieber P, Bonifazi E, Broberg A, Calza A, Coleman R, De Prost Y, Stalder J, Gelmetti C, Cuannetti A, Harper J, Kunz B, Lachapelle J, Langeland T, Lever R, Oranje A, Oueille-Roussel C, Revuz J, Ring J, Roujeau J, Saurat J, Song M, Tennstedt D, Van Neste D, Vieluf D, Poncet M. Severity scoring of atopic dermatitis: the SCORAD index. Consensus Report of the European Task Force on Atopic Dermatitis. *Dermatology*. 1993;186(1):23-31.