Statistical Analysis Plan Version 8 I4V-MC-JAHN

A Phase 3 Multicenter, Double-Blind Study to Evaluate the Long-Term Safety and Efficacy of Baricitinib in Adult Patients with Atopic Dermatitis

NCT03334435

Approval Date: 25-Aug-2023

1. Statistical Analysis Plan: I4V-MC-JAHN: A Phase 3 Multicenter Double-Blind Study to Evaluate the Long-Term Safety and Efficacy of Baricitinib in Adult Patients with Atopic Dermatitis

BREEZE-AD3

Baricitinib (LY3009104) Atopic Dermatitis

Study I4V-MC-JAHN (NCT03334435) is a 4-year Phase 3 multicenter, double-blind long-term extension study to evaluate the safety and efficacy of placebo, baricitinib 1 mg, baricitinib 2 mg, and baricitinib 4 mg in adult patients with atopic dermatitis, including a blinded randomized treatment withdrawal and randomized down-titration.

Eli Lilly and Company
Indianapolis, Indiana USA 46285

Protocol I4V-MC-JAHN
Phase 3

2. Table of Contents

Section **Page**

Table of Contents

Table **Page**

Table of Contents

Figure **Page**

Table of Contents

Appendix **Page**

3. Revision History

Statistical Analysis Plan (SAP) Version 1 is based on Protocol I4V-MC-JAHN(a) (JAHN[a]) and was approved 12 April 2019 prior to the first unblinding. The analyses for Treatment Period 1 of Addendum I4V-MC-JAHN(7) (JAHN[7]) are described in Appendix 1. The analyses for the Withdrawal and Down-titration Substudy of I4V-MC-JAHN (JAHN), for both the main protocol JAHN(a) and addendum JAHN(7), will be described in subsequent versions of this SAP. SAP Version 2 was approved 03 May 2019 prior to the Week 16 database lock (DBL) and includes the following changes:

- *Section 5.1* – separated "responders" into 2 groups: "responders" and "partial responders"
- *Section 6.2.1* – added verbiage to distinguish patients from originating studies: monotherapy studies (I4V-MC-JAHL [JAHL]/I4V-MC-JAHM [JAHM]) and combo therapy study (I4V-MC-JAIY [JAIY])
- *Section 6.2* – updated model to include region; added methods for time to 4-point itch reduction and time to partial response, and aligned Cyclosporine wording to that of originating studies.
- *Section 6.4* – updated censoring rule, updated sensitivity analyses
- *Section 6.8.2* – Baseline Disease Characteristics: Cyclosporine wording updated to "…prior to entry into originating study" and added topical calcineurin inhibitor (TCNI) and vaccine
- *Section 6.9.1* – Background Therapy: added 2 new analyses "Summary of AD Background Therapy" and "Summary of time to initial use of AD background therapy post baseline" (for responders and partial responders)

SAP Version 3 was approved 29 July 2019 prior to the DBL that supports the EU and Japan submissions and includes the following changes:

- *Section 6.2.1* – a Week 36 efficacy evaluable set was defined for Week 36 efficacy and health outcome analyses.
- *Section 6.2.2* – clarified that for step 2 of the Week 16 daily diary window construction, the visit anchor date will be on or before a scheduled visit occurring after Week 16.
- *Sections 6.2.3 and 6.11.1; Tables JAHN.6.6 and JAHN.6.8* – clarified that efficacy and health outcome analyses for the EU and Japan DBL will be done for Weeks 0 through 24, and for Week 36 only descriptive statistics will be included.
- *Section 6.2.3* – analyses for safety, for the EU and Japan DBL, will include data for Weeks 0 through 52 entered up to the database cutoff date.
- *Table JAHN.6.2* – added modified baseline observation carried forward (mBOCF) imputation method for nonresponder modified intent-to-treat (mITT) continuous sensitivity analyses.
- *Section 6.7* – clarified that a by-patient listing of disposition and of treatment assignments will be provided for the intent-to-treat (ITT) population.
- *Section 6.9.1* was edited and reorganized for clarity.
- *Section 6.11.3* – added sensitivity analyses for categorical outcomes using JAHN baseline instead of the baseline from the respective originating study.

- *Tables JAHN.6.6 and JAHN.6.8* – added the following analyses:
  - proportion of patients with a response of Investigator's Global Assessment (IGA) [0,1], analysis using JAHN as the baseline value
  - proportion of patients with a 4-point improvement from JAHN baseline in Itch Numeric Rating Scale (NRS) in subgroup of patients who had baseline Itch NRS ≥4
  - proportion of patients with a 4-point improvement from baseline of originating study in Skin Pain in subgroup of patients who had baseline Skin Pain NRS ≥4
  - proportion of patients with a 1-point improvement from baseline of originating study in a subgroup of patients with baseline item score of ≥1 for the following Atopic Dermatitis Sleep Scale (ADSS) item scores:
    - Item 1 score
    - Item 2 score
    - Item 3 score
  - proportion of patients with a 2-point improvement from baseline of originating study in item 2 score of ADSS in a subgroup of patients with baseline item 2 score of ≥2
  - proportion of patients with a 4-point improvement in Patient-Oriented Eczema Measure (POEM) score from baseline of originating study in a subgroup of patients with baseline POEM ≥4
  - change from baseline in Hospital Anxiety Depression Scale (HADS) total score
  - proportion of patients with HADS domain (anxiety and depression) score of <8 in a subgroup of patients with baseline HADS domain (anxiety and depression) score of ≥8
  - proportion of patient with HADS anxiety or depression domain score of <8 in a subgroup of patients with baseline HADS anxiety or depression domain score of ≥8
  - proportion of patients with Dermatology Life Quality Index (DLQI) (0,1); and
  - proportion of patients with a 4-point improvement from baseline in originating study in a subgroup of patients with baseline DLQI score ≥4
- *Section 6.14.1* – defined exposure ranges for the EU and Japan DBL.
- *Section APP.1.2.5; Tables APP.1.3 and APP.1.4* – clarified that for the EU and Japan DBL the efficacy and health outcome analyses will include descriptive statistics for Weeks 0 through 36 only.
- *Table APP.1.2* – removed the mixed -model repeated measures (MMRM)/mBOCF imputation methods for continuous endpoints.
- *Section APP.1.2.8* – clarified that the by-patient listing for treatment assignments will be provided for the ITT population.
- *Section APP.1.2.15* clarified that the open-label safety analyses will be pooled with the baricitinib 2-mg treatment group from Study JAHN.

SAP Version 4 was approved 23 January 2020 prior to the DBL that supports the US submission. This version of the SAP pertains exclusively to the US submissions DBL and does not contain information related to any other prior or upcoming DBLs. This version is based on Protocol I4V-MC-JAHN(c) (JAHN[c]) and was approved prior to the first unblinding. The analyses for treatment Period 1 of Addendum I4V-MC-JAHN(7.1) (JAHN[7.1]) are described in Appendix 1. This includes the following changes:

- *Section 5.1* – changed the duration of Treatment Period 2 in accordance with Protocol amendment JAHN(c).
- *Figure JAHN.5.1.* – changed to reflect the change in Study Period 2 timeline.
- *Section 6.1* – updated sample size for the study.
- *Section 6.2.1* – added the following populations:
  - Week 24 efficacy evaluable set for patients originating from study JAIY
  - Weeks 56, 60, 64, and 68 efficacy evaluable set for patients entering downtitration and withdrawal substudy (Period 2)
  - safety population for Period 2
- *Table JAHN.6.1* – updated to reflect separate treatment groups for Period 1 and Period 2. Added treatment groups for Treatment Period 2.
- *Section 6.2.2* – the following changes were incorporated:
  - Baseline and postbaseline measures are separately defined for Period 1 and Period 2.
  - Week 52 baseline is defined.
  - Baseline for safety is defined.
  - Daily diary window calculation for Period 2 is added.
- *Section 6.2.3* – the following changes were incorporated:
  - Treatment comparisons for patients originating from combination therapy study and for patients entering Period 2 are added.
  - Analysis of covariance (ANCOVA) is added as the primary analysis method for continuous endpoints for Period 2.
  - Deleted analysis using Cox proportional hazards model.
- *Section 6.3* – added week 52 baseline values for Period 2 covariate adjustments.
- *Section 6.4* – the following changes were incorporated:
  - added retreatment after downtitration or early withdrawal as an intercurrent event
  - added Period 2 censoring rule
  - deleted section on mBOCF method for missing data
- *Table JAHN.6.2* was updated as follows:
  - deleted placebo multiple imputation (pmi) method for missing data handling for IGA 0,1 endpoint
  - modified last observation carried forward (mLOCF) added as an imputation method
  - deleted mBOCF imputation method
  - added imputation methods used in Period 2
- *Section 6.7* – added patient disposition calculation for Period 2
- *Section 6.8* – added patient compliance calculation for Period 2
- *Section 6.9.1* – added background therapy calculation for Period 2
- *Section 6.7* – added patient disposition calculation for Period 2
- *Table JAHN.6.3* – added treatment groups for Period 2
- *Table JAHN.6.4* – added variables relapse and rebound
- *Table JAHN.6.5* – added analysis for the following:
  - patients originating from combination therapy study
  - patients entering Period 2
  - changed time point analyzed for monotherapy studies in accordance with the US DBL

  - added analysis for endpoints rebound and relapse
- *Tables JAHN.6.7* – added analysis corresponding to patients originating from combination therapy study
- *Section 6.8* – added safety analysis for patients entering Period 2
- *Appendix 1* – aligned with protocol amendment JAHN(7.1)
- *Section APP.1.2.5; Table APP.1.3 and APP.1.4* – clarified that for the EU and Japan DBLs the efficacy and health outcome analyses will include descriptive statistics for Weeks 0 through 36 only
- *Figure APP.1.2.* – changed to reflect the change in Study Period 2 timeline.
- *Section APP.1.2.8* clarified that the by-patient listing for treatment assignments will be provided for the ITT population
- *Section APP.1.2.3.2.1* – defined Week 24 efficacy evaluable set for open label patients
- *Table APP.1.3* – changed analysis time period to be Week 0 to 16 and Week 24
- *Table APP.1.4* – added analysis for the following endpoints:
  - 4-point improvement in Skin Pain NRS
  - 1.5-point improvement in ADSS Item 2 score
  - 4-point improvement in POEM score
  - DLQI (0,1) response
  - 4-point improvement in DLQI score

SAP Version 5 was approved 19 May 2020 prior to the 4-month safety update (20 May 2020). The overall changes incorporated in Version 5 are as follows:

- *Section 6.4.4* – section added to include missing data imputation method for missingness due to COVID-19.
- *Appendix 2* – appendix 2 added to include the list of efficacy analyses planned for the 4-month safety update DBL.

SAP Version 6 was approved18 January 2021 prior to the DBL (February 2021), when all subjects completed or discontinued by Week 52; In addition, all sub study eligible patients will have reached Week 68 (16 weeks Follow-up post re-randomization into the sub study), and most of the substudy-eligible patients will reach Week 76. The changes incorporated in Version 6 are summarized as follows:

- *Section 4.3* – added the summary for pooled response analysis.
- *Section 6.2.2* – updated the Period 2 baseline usage per recent study team's discussion.
- *Section 6.4* –updated Period 1and Period 2 censoring rule to make it more specific and clear.
- *Section 6.4.4* – added a paragraph about not implementing a special missing imputation method if the missing data due to Covid-19 does not exceed 5%.
- *Table 6.5* and *6.7* – updated to include summary and analysis endpoints for DBL in February 2021.
- *Section 6.14* – added the summary of Safety for DBL in February 2021.

SAP Version 7 was approved 14 December 2021 prior to the DBL (December 2021), when all subjects completed or discontinued by Week 104; This version of the SAP pertains to the analysis for Period 2 for all substudy eligible patients, patients not entered into substudy and long-term baricitinib 4 mg. The changes incorporated in Version 7 are summarized as follows:

- *Section 4.3* - added the summary for pooled response analysis from Week52 to 104; added the exploratory analyses for nonsubstudy.
- *Section 6.2.1* - added new analysis population.
- *Section 6.*8 - added baseline characteristics and disposition summary for new analysis population.
- *Table 6.*3 - added treatment group for newly defined analysis population.
- *Table 6.5* and *6.7* - updated to include summary and analysis endpoints for DBL in December 2021.

SAP Version 8 was approved prior to unblinding as per baricitinib roll-off program. Baricitinib is approved for moderate-to-severe atopic dermatitis in over 50 countries. Based on widespread approval, Lilly has made the decision to begin closure of sites in countries where there is regulatory approval and reimbursement. This is being rolled out in a "phased" approach based on country approval and accessibility. Patients will be rolled off by completion of their early termination visit following confirmation that baricitinib is available locally. Based on specific study design elements and required safety monitoring per labeling, unblinding investigators to patients' treatment assignment will occur after completion and documentation of final treatment visits to allow for informed clinical decision-making on the best, optimized treatment plan post-clinical trial completion. All data up to Week 104, including all primary and secondary endpoints, have been locked, and any blinded data required for regulatory submissions has been submitted, thus, it is considered that unblinding investigators at this time has no impact on overall validity of study data and outcomes.

This version of the SAP pertains to the analysis at final DBL. The changes incorporated in Version 8 are summarized as follows:

- *Section 4.3* - updated the summary for analysis beyond Week 104
- *Section 6.2.1* – updated for final DBL.
- *Section 6.*8 – updated for Period 2 baseline characteristics and disposition summary
- *Table 6.1* – extended period 2 from Week 104 to Week 200, and
- *Table 6.5* and *Table 6.7* – added footnote for analysis of endpoints beyond Week 104

4. Study Objectives

4.1. Primary Objective

The primary objective of this study is to estimate the effect of long-term therapy with baricitinib on responders and partial responders at entry of Study I4V-MC-JAHN (JAHN) as assessed by the proportion of patients with a response of Investigator's Global Assessment (IGA) 0 or 1 at Weeks 16, 36, and 52.

In particular, the associated estimand for this objective is to measure the effect of long-term therapy with baricitinib on responders and partial responders at entry of Study JAHN as assessed by the proportion of patients with a response of IGA 0 or 1 at Weeks 16, 36, and 52 assuming that treatment response disappears after patients discontinue from study or treatment. See Sections 6.4.1 and 6.11.1 on how this estimand handles outcomes after the occurrence of any intercurrent event through nonresponder imputation (NRI).

4.2. Secondary Objectives

The secondary objectives of this study are:

Objectives	Endpoints
Weeks 0-52	
Baricitinib Patients at Entry to Study JAHN To evaluate the effect of increasing or maintaining baricitinib dose on clinical measures and patient-reported outcomes.	• Proportion of patients with a response of Investigator's Global Assessment (IGA) 0, 1, or 2 assessed at Weeks 16, 36, and 52 • Proportion of patients with a response of IGA 0 or 1 assessed at Weeks 16, 36, and 52 (nonresponders) • Proportion of patients achieving response of ≥75% improvement from baseline in Eczema Area and Severity Index score (EASI75) from baseline of originating study assessed at Weeks 16, 36, and 52 • Proportion of patients with a 4-point improvement from baseline of originating study in Itch Numeric Rating Scale (NRS) at Week 16
Placebo Nonresponders at Entry to Study JAHN To evaluate the effect of starting baricitinib 2-mg versus 4-mg on clinical measures and patient-reported outcomes.	• Proportion of patients with a response of IGA 0, 1, or 2 assessed at Weeks 4, 16, 24, 52 • Proportion of patients with a response of IGA 0 or 1 assessed at Weeks 4, 16, 24, 52 • Proportion of patients achieving response of EASI75 from baseline of originating study assessed at Weeks 4, 16, 24, 52 • Proportion of patients with a 4-point improvement from baseline of originating study in Itch NRS at Week 16

Objectives	Endpoints
Weeks 52-104	
All Patients Entering the Substudy To evaluate the change in clinical response after treatment withdrawal or downtitration. **Patients Entering the Substudy with IGA 0 or 1** To evaluate the change in clinical response after treatment withdrawal or downtitration.	• Proportion of patients with a response of IGA 0, 1, or 2 assessed at 16 weeks after rerandomization (Week 68) and Week 104 • Proportion of patients with a response of IGA 0 or 1 assessed at 16 weeks after rerandomization (Week 68) and Week 104 • Proportion of patients with a response of EASI75 from baseline of originating study assessed at 16 weeks after rerandomization (Week 68) and Week 104 • Time to retreatment (time to IGA ≥3)
Patients Retreated during Substudy To evaluate the ability to recapture efficacy based on clinical measures after experiencing a loss of treatment benefit:	• Proportion of patients with a response of IGA 0, 1, or 2 within 16 weeks of retreatment • Proportion of patients with a response of IGA 0 or 1 within 16 weeks of retreatment • Proportion of patients with a response of EASI75 from baseline of originating study within 16 weeks of retreatment
Patients Not Entered into the Substudy To evaluate the effect of maintaining baricitinib dose on clinical measures.	• Proportion of patients with a response of IGA 0, 1, or 2 assessed at Week 104 • Proportion of patients with a response of IGA 0 or 1 assessed at Week 104 • Proportion of patients with a response of EASI75 from baseline of originating study assessed at Week 104

4.3. Exploratory Objectives

The exploratory objectives may include evaluating the response to baricitinib treatment regimens on clinical measures and patient-reported outcomes (PROs).

These endpoints may include dichotomous endpoints or change from baseline for the following measures: IGA, Eczema Area and Severity Index (EASI), SCORing Atopic Dermatitis (SCORAD), Patient-Oriented Eczema Measure (POEM), Dermatology Life Quality Index (DLQI), Work Productivity and Activity Impairment (WPAI-AD), European Quality of Life–5 Dimensions–5 Levels (EQ-5D-5L), Itch Numeric Rating Scale (NRS), Atopic Dermatitis Sleep Scale (ADSS) Item 2 score, Skin Pain NRS, and Patient Global Impression of Severity (PGI-S-AD).

Additionally, assessment of background topical corticosteroids (TCS), such as the weight of sponsor-provided TCS and number of days not using TCS, will be analyzed. Patients continuing on placebo as responders will be assessed during the long-term extension for relevant efficacy endpoints.

Time to 4-point itch reduction during the first 14 days after initiation of treatment will be assessed.

Summary of IGA response (0,1) at Week 16 for the proportion of nonresponders with originating study treatment assignment in placebo, baricitinib 1 mg, or baricitinib 2 mg with IGA response (3,4) at the start of study JAHN (Week 0) will be provided.

For IGA (0,1), EASI75, and Itch NRS ≥4-point improvement, a pooled response descriptive summary for baricitinib 2 mg to baricitinib 2 mg and baricitinib 4 mg to baricitinib 4 mg will be done for Week 0 through Week 52 based on I4V-MC-JAHL (JAHL)/I4V-MC-JAHM (JAHM), I4V-MC-JAIY (JAIY), and open label, respectively.

For all the randomized patients who have not entered into the substudy and received at least 1 dose in Period 2, the exploratory analyses will be assessed on the following endpoints: EASI change percent change from baseline, DLQI, SCORAD, POEM, Hospital Anxiety and Depression Scale (HADS). Open-label patients who were not entered into substudy and received at least 1 dose in Period 2 will be analyzed separately in a similar way.

In addition, for the Long-term Baricitinib 4-mg population, the following endpoints, IGA response (0,1), EASI, DLQI, SCORAD, and subjective characteristics of SCORAD (pruritus and sleep loss), will be analyzed in a descriptive way for Week 52 through 104 based on pooled patients, as well as by monotherapy and combination therapy study respectively.

For the final DBL, efficacy endpoints beyond Week 104 will be exploratory in nature and will be analyzed in a descriptive way using observed data, unless otherwise stated. For endpoint collected beyond Week 104, refer to protocol schedule of activities.

5. Study Design

5.1. Summary of Study Design

Study JAHN is a Phase 3, multicenter, double-blind study to evaluate the long-term safety and efficacy of daily baricitinib 1 mg, 2 mg, and 4 mg in patients with atopic dermatitis (AD) for approximately 2 years. The study will consist of 3 study periods and 1 substudy: randomized treatment withdrawal and downtitration. Patients entering Study JAHN will be classified as "responders," "partial responders" or "nonresponders" based on the following criteria:

- **Responders (IGA of 0 or 1):** patients entering Study JAHN who have an IGA score of 0 or 1 and were not rescued in the originating study
- **Partial Responders (IGA of 2):** patients entering Study JAHN who have an IGA score of 2 and were not rescued in the originating study
- **Nonresponders:** any patient who does not meet the "responder" or "partial responder" definitions

Figure JAHN.5.1 illustrates the study design. The full visit schedule is outlined in the Study Schedule of Activities in the Protocol. Patients who completed originating studies, such as Studies JAHL, JAHM or JAIY may be eligible for enrollment into Study JAHN. This study also includes an open-label addendum for which patients do not have to complete an originating study.

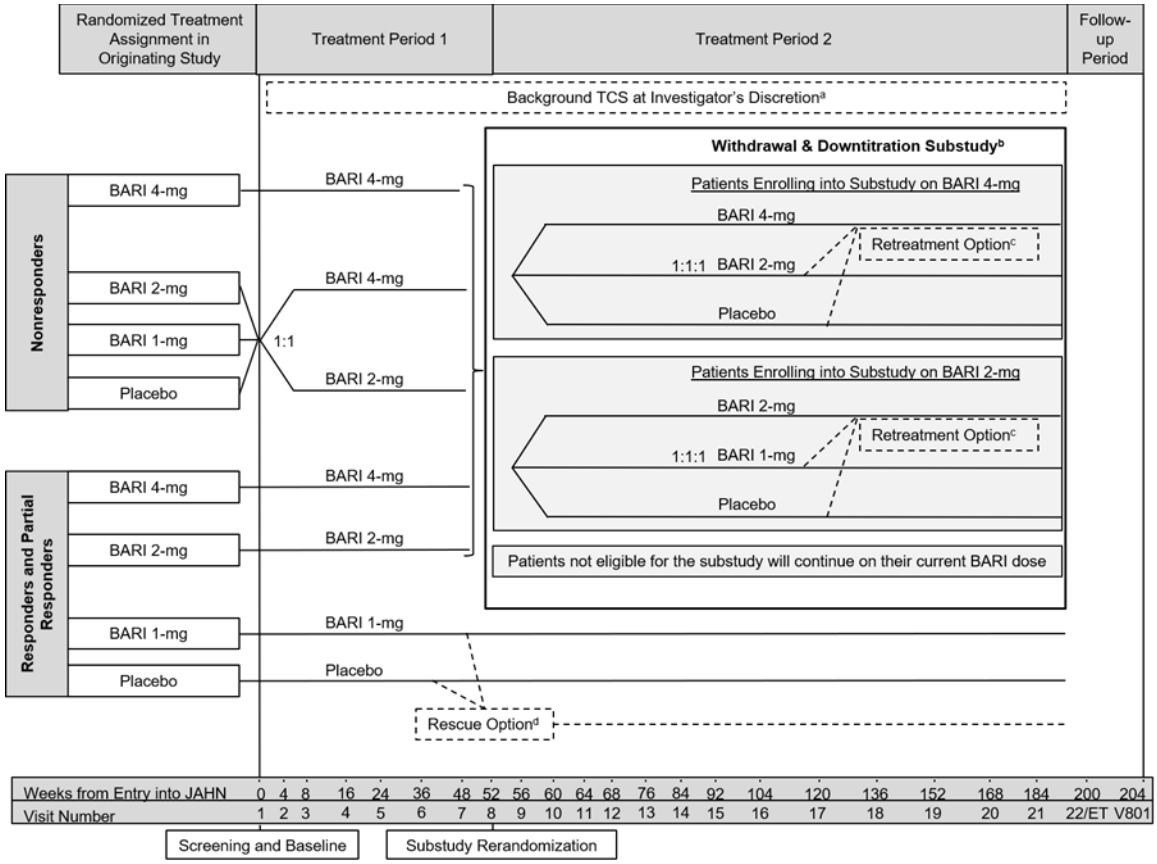
Patients completing Studies JAHL, JAHM or JAIY will have had approximately 16 weeks of treatment with investigational product (IP) (baricitinib or placebo) and will enter Study JAHN at Visit 1; the final visit of Study JAHL, JAHM or JAIY will be the first visit of Study JAHN unless washout of a concomitant medication was required in which case a new Visit 1 is required. If oral systemic AD treatments were administered for rescue in the originating study, then a minimum washout of 4 weeks is required prior to study drug initiation in Study JAHN (Visit 1). Patients will have been using emollients daily in their originating study and thus will continue during Study JAHN participation.

Daily diary dispensation/collection will continue through the first 4 months of treatment in JAHN and then dispensation will resume at Week 48 for an additional 5 months of collection to capture PROs prior to and during the withdrawal and downtitration substudy (except for patients known not to be in the substudy).

Study JAHN will consist of 4 periods:

- Screening and baseline period: Screening should occur during the last visit of the originating study unless the patient received oral systemic AD treatment as rescue therapy during the originating study and needs additional time to complete the required 4-week washout period or the sponsor gave approval.
- Treatment Period 1: Treatment period from Week 0 to Week 52. Patients will be classified as responders, partial responders or nonresponders at the entry of Study JAHN based on the criteria given in Section 5.1.

- Treatment Period 2: Treatment period from Week 52 to 200 and will contain a randomized withdrawal and downtitration substudy. At Week 52, all patients will be evaluated for substudy eligibility, and the criteria are given in the Protocol. Patients not entered in the substudy will continue on their current treatment.
- Posttreatment follow-up period: Period from Week 200 (Visit 22) or Early Termination Visit (ETV) to approximately 28 days after the last dose of IP.



Abbreviations: BARI = baricitinib; IGA = Investigator's Global Assessment; TCS = topical corticosteroids.

[a] Background TCS may be initiated or reinitiated at any time during the study, following the guidelines in the Protocol (Section 7.7.1) and will be provided as part of rescue or retreatment any time a patient's IGA score becomes ≥3 as described in the Protocol (Sections 5.1.2.1 and 5.1.3.1).

[b] Eligible patients will be rerandomized in the withdrawal and downtitration substudy as described in the Protocol (Section 5.1.3.1). Patients who do not enroll in the substudy will remain on their treatment.

[c] Patients enrolled in the substudy will automatically be retreated if their IGA score becomes ≥3 as described in the Protocol (Section 5.1.3.1).

[d] Rescue is available as described in the Protocol (Section 5.1.2.1).

Figure JAHN.5.1. Illustration of study design for Clinical Protocol I4V-MC-JAHN.

5.2. Method of Assignment to Treatment

At entry into Study JAHN, patients who meet all criteria for enrollment will be randomized or assigned treatment by a computer-generated random sequence using an interactive web-response system (IWRS). Patients originally assigned to placebo, 1-mg baricitinib, or 2-mg baricitinib in Studies JAHL or JAHM, or patients originally assigned to placebo or 2-mg baricitinib in Study JAIY and classified as nonresponders, will be randomized in a 1:1 ratio to 2-mg baricitinib or 4-mg baricitinib and will be stratified by disease severity at baseline of JAHN (IGA 0, 1, 2 versus IGA 3 versus IGA 4). All other patients will be assigned to treatment in Study JAHN matching their prior assignment from Studies JAHL, JAHM, or JAIY.

Additional addendum patients will be permitted to enroll directly into Study JAHN on baricitinib 2 mg. These patients will not be blinded to treatment.

At Week 52, patients eligible for the withdrawal and downtitration substudy will be assigned to treatment by a computer-generated random sequence using an IWRS. Rerandomization will follow a 1:1:1 ratio allocation and will be stratified by disease severity (IGA 0, 1 versus IGA 2). Patients entering the substudy on baricitinib 4 mg will be rerandomized to baricitinib 4 mg, baricitinib 2 mg, and placebo; patients entering the substudy on baricitinib 2 mg will be rerandomized to baricitinib 2 mg, baricitinib 1 mg, and placebo.

The IWRS will be used to assign blister packs, each containing double-blind IP tablets, to each patient, starting at Visit 1 (Week 0), and at each visit up to and including Visit 21 (Week 184). Site personnel will confirm that they have located the correct blister packs by entering a confirmation number found on the blister packs into the IWRS.

This study will be conducted internationally in multiple sites. Table JAHN.5.1 describes how regions were defined for stratification in the originating studies.

Table JAHN.5.1. Geographic Regions for Stratification

Region	Country
Europe	Austria, Czech Republic, Denmark, France, Germany, Hungary, Italy, Poland, Spain, Switzerland
Japan	Japan
Rest of World	Argentina, Australia, India, Israel, Korea, Mexico, Russia, Taiwan

6. A Priori Statistical Methods

6.1. Determination of Sample Size

It is anticipated that 95% of enrolled patients will complete Studies JAHL, JAHM, and JAIY and roll over into Study JAHN. Therefore, planned enrollment into Study JAHN from the originating Studies JAHL, JAHM, and JAIY will be approximately 1425 patients. Patients who are nonresponders at entry into Study JAHN and were not randomized to baricitinib 4 mg will be randomized 1:1 to either baricitinib 4 mg or baricitinib 2 mg and will be stratified by disease severity at baseline of JAHN (IGA 0, 1, and 2 versus IGA 3 versus IGA 4). This study is intended to evaluate patients' long-term response of baricitinib, and the sample sizes are not determined to detect differences between baricitinib and an appropriate comparator in a statistically powered manner. Additional patients may enroll from addenda or other studies.

Patients at Week 52 will be stratified by responder status (IGA 0 or 1 versus IGA 2) when entering the randomized withdrawal and downtitration substudy. It is estimated that there will be approximately 600 patients entering the randomized withdrawal and downtitration substudy. The substudy is meant to evaluate the change in clinical response after treatment withdrawal or downtitration and does not account for whether the sample size is sufficient to detect a difference between baricitinib and placebo. Maintenance of treatment benefit is defined as response of IGA 0, 1, or 2.

6.2. General Considerations

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

Statistical analysis of this study will be the responsibility of Eli Lilly and Company (Lilly). The statistical analyses will be performed using SAS® Version 9.4 or higher.

Not all displays described in this statistical analysis plan (SAP) will necessarily be included in the clinical study report (CSR). Not all displays will necessarily be created as a "static" display. Some may be incorporated into interactive display tools instead of or in addition to a static display. Any display described in this SAP and not included in the CSR would be available upon request.

Statistical tests of treatment effects and confidence intervals (CIs) will be performed at a 2-sided significance level of 0.05, unless otherwise stated.

Data collected at early termination visits will be mapped to the closest scheduled visit if it falls within the visit window as discussed in Section 6.2.2. For by-visit summaries, only visits in which a measure was scheduled to be collected will be summarized. Any unscheduled visit data will be included at the patient-level listings. However, the data will still be used in other analyses, including categorical analyses for safety.

6.2.1. Analysis Populations

Intent-to-treat (ITT) population: The ITT population analysis set is defined as all enrolled patients in Study JAHN.

Modified intent-to-treat (mITT) population: The mITT population analysis set is defined as all randomized patients who received at least one dose of the IP in Study JAHN.

The efficacy and health outcome analyses will be conducted on subsets of the mITT population, which is dependent on the objectives. For all analyses of patients from originating studies, responders, partial responders and nonresponders will be distinguished by their originating studies at entry of Study JAHN – monotherapy (JAHL/JAHM) and combination therapy (JAIY). Patients will be analyzed according to the treatment to which they were randomized.

The analysis of the primary objective will be conducted on the mITT population who are responders and partial responders (Section 5.1). Additionally, these endpoints will also be evaluated by parsing this group into responder mITT and partial responder mITT groups.

The populations for the secondary objectives (nonresponders who were on a baricitinib dose in originating study and nonresponders who were on placebo in originating study) will also be comprised of patients meeting the mITT definition. Exploratory objectives will also be conducted on the same populations.

For the database lock (DBL) that supports the United States (US) submission (13DEC2019) (subsequently referred to as the US DBL), the population originating from combination therapy study JAIY, Week 24 efficacy and health outcome analyses will be conducted on the **Week 24 efficacy evaluable set**. This subset of the mITT Population is anchored on the database cut-off date for the US DBL. Specifically, a patient will be included in the Week *x* efficacy evaluable set if their Week *x* visit has occurred or their expected Week *x* visit date plus a 15-day buffer is on or prior to the database cutoff date. The expected Week *x* visit date will be calculated as follows: date of first dose date + (*x* weeks * 7 days) + 15 days.

For the US DBL, for Treatment Period 2 (downtitration and withdrawal), Week 56, Week 60, Week 64, and Week 68 efficacy analyses will be conducted on the Week 56, Week 60, Week 64, and Week 68 efficacy evaluable sets respectively. The efficacy evaluable sets are a subset of the mITT population and are anchored on the database cut-off date for the US DBL.

For the DBL in December 2021 and final DBL, the following analysis populations are defined for efficacy and health outcome analyses.

Randomized Downtitration Withdrawal Substudy Population: All patients (including open-label patients) who are rerandomized at Week 52, enter the substudy, and receive at least 1 dose of the IP in Period 2.

Re-Treatment Substudy Population: A subset of Randomized Downtitration Withdrawal Substudy Population who experience IGA ≥3 at any time in Period 2 and receive at least 1 dose of retreatment of the original dose.

Period 2 Nonsubstudy Population: All patients randomized at Week 0 of main Study JAHN (excluding open-label patients) who are not eligible to enter the substudy at Week 52 and receive at least 1 dose of the IP in Period 2.

Open-Label Period 2 Nonsubstudy Population: All patients enrolled in the open-label addendum who are not eligible to enter the substudy at Week 52 and receive at least 1 dose of the IP in Period 2.

Long-Term Baricitinib 4 mg Population: Patients initially randomized to baricitinib 4 mg in the originating studies and remain on 4-mg up to Week 52 of Study JAHN. At Week 52 of Study JAHN, patients also need to continue 4-mg and receive at least one dose in Period 2 regardless of eligibility of entering sub-study.

Safety population: For the purpose of Study JAHN alone, the safety population is defined as all patients who received at least 1 dose of IP in Study JAHN and who did not discontinue from the study for the reason 'Lost to Follow-up' at the first postbaseline visit in Study JAHN. For the safety analysis of Treatment Period 2, patients with at least 1 dose of IP in Treatment Period 2 will be included in the safety analysis of Treatment Period 2.

Safety analyses will be performed using the safety population and will be analyzed according to the following treatment groups (or cohorts) defined in Table JAHN.6.1. Patients will be analyzed according to the dosing regimen to which they were assigned.

Table JAHN.6.1. Safety Population Treatment Groups

Treatment Groups for Treatment Period 1 (Week 0-52)	**Definition**
PBO	Placebo at entry to Study JAHN followed to end of Treatment Period 1 or censored at treatment change
BARI 1 mg	BARI 1 mg followed to end of Treatment Period 1 or censored at dose or treatment change
PBO/BARI 1 mg to BARI 2 mg	PBO/BARI 1 mg switching to BARI 2 mg at entry to Study JAHN followed to end of Treatment Period 1 or censored at dose or treatment change
BARI 2 mg	BARI 2 mg followed to end of Treatment Period 1 or censored at dose or treatment change
PBO/BARI 1 mg/BARI 2 mg to BARI 4 mg	PBO/BARI 1 mg/BARI 2 mg switching to BARI 4 mg at entry to Study JAHN followed to end of Treatment Period 1 or censored at dose or treatment change
BARI 4 mg	BARI 4 mg followed to end of Treatment Period 1 or censored at dose or treatment change
Pooled BARI	BARI 1 mg/BARI 2 mg/BARI 4 mg at entry to Study JAHN followed to end of Treatment Period 1 or censored at dose or treatment change
Treatment Groups for Treatment Period 2 (Week 52-200)	**Definition**
BARI 2 mg to PBO	Randomized to PBO in Treatment Period 2 (previously treated with BARI 2 mg) followed to end of study or censored at retreatment with BARI 2 mg

Treatment Groups for Treatment Period 2 (Week 52-200)	Definition
BARI 2 mg to BARI 1 mg	Randomized to BARI 1 mg in Treatment Period 2 (previously treated with BARI 2 mg) followed to end of study or censored at retreatment with BARI 2 mg
BARI 2 mg to BARI 2 mg	Randomized to stay on BARI 2 mg in Treatment Period 2 followed to end of study
BARI 4 mg to PBO	Randomized to PBO in Treatment Period 2 (previously treated with BARI 4 mg) followed to end of study or censored at retreatment with BARI 4 mg
BARI 4 mg to BARI 2 mg	Randomized to BARI 2 mg in Treatment Period 2 (previously treated with BARI 4 mg) followed to end of study or censored at retreatment with BARI 4 mg
BARI 4 mg to BARI 4 mg	Randomized to stay on BARI 4 mg in Treatment Period 2 followed to end of study
BARI 2 mg	Not part of the substudy, stayed on BARI 2 mg in Treatment Period 2 followed to end of study (this treatment group includes the open-label BARI 2 mg addendum patients)
BARI 4 mg	Not part of substudy, stayed on BARI 4 mg in Treatment Period 2 followed to end of study
Treatment Group for Treatment Periods 1 and 2 (Week 0-104) (for Box Plots Only)	**Definition**
PBO	Placebo at entry to Study JAHN followed to end of study or censored at treatment change
BARI 1 mg	BARI 1 mg followed to end of study or censored at dose or treatment change
BARI 2 mg to PBO	BARI 2 mg in Treatment Period 1, rerandomized to PBO in Treatment Period 2 followed to end of study or censored at retreatment with BARI 2 mg
BARI 2 mg to BARI 1 mg	BARI 2-mg in Treatment Period 1, rerandomized to BARI 1 mg in Treatment Period 2 followed to end of study or censored at retreatment with BARI 2 mg
BARI 2 mg to BARI 2 mg	BARI 2 mg in Treatment Period 1, rerandomized to stay on BARI 2 mg in Treatment Period 2 followed to end of study
BARI 4 mg to PBO	BARI 4 mg in Treatment Period 1, rerandomized to PBO in Treatment Period 2 followed to end of study or censored at retreatment with BARI 4 mg
BARI 4 mg to BARI 2 mg	BARI 4 mg in Treatment Period 1, rerandomized to BARI 2 mg in Treatment Period 2 followed to end of study or censored at retreatment with BARI 4 mg
BARI 4 mg to BARI 4 mg	BARI 4 mg in Treatment Period 1, rerandomized to stay on BARI 4 mg in Treatment Period 2 followed to end of study
BARI 2 mg	BARI 2 mg in Treatment Period 1, did not enter the substudy and stayed on BARI 2 mg in Treatment Period 2 followed to end of study (this treatment group includes the open label BARI 2 mg addendum patients)
BARI 4 mg	BARI 4 mg in Treatment Period 1, did not enter the substudy and stayed on BARI 4mg in Treatment Period 2 followed to end of study

Abbreviations: BARI = baricitinib; PBO = placebo.

An all-baricitinib exposure, defined as patients valid for the safety population with at least 1 dose of baricitinib anytime during Study JAHN, will be analyzed for the final analysis (final datalock).

Follow-up population: The follow-up population is defined as patients who entered the follow-up period and completed the follow-up visit.

For the February 2021 DBL, adverse event (AE) Period 1 tables will be summarized by treatment in Period 1, which include treatment group of (PBO, BARI 1-mg, BARI-2mg, BARI 4-mg, Pooled BARI).

For final DBL, AE tables will be summarized by treatment in Period 2 as specified in Table JAHN.6.1.

6.2.2. Definition of Baseline and Postbaseline Measures

Baseline

Period 1 (Weeks 0 to 52)

The baseline value for efficacy variables measured at scheduled visits is defined as the last nonmissing measurement on or prior to the date of first study drug administration in the originating study (expected at Week 0, Visit 2 of the originating study). For patients who directly enroll into Study JAHN without having completed an originating study, baseline value for efficacy variables measured at scheduled visits is defined as the last nonmissing measurement on or prior to the date of first study drug administration in Study JAHN.

The baseline value for the daily diary assessments is the mean of the nonmissing assessments in the 7 days prior to the date of first study drug administration in the originating study (expected at Week 0, Visit 2).

If there are <4 nonmissing assessments in the baseline diary window, the interval lower bound can be extended up to 7 additional days, one day at a time, to obtain the most recent 4 nonmissing values. If there are not at least 4 nonmissing assessments in the baseline period, the baseline mean is missing.

Additionally, the Study JAHN baseline may be derived for some efficacy analyses (eg, patients enrolled directly into Study JAHN without having completed an originating study). In this case, the Study JAHN baseline is defined as the last nonmissing measurement on or prior to the date of first study drug administration in Study JAHN. For daily diary data, the baseline is the mean of the nonmissing assessments in the 7 days prior to the date of first study drug administration in the JAHN study. If there are <4 nonmissing assessments in the baseline diary window, the interval lower bound can be extended up to 7 additional days.

For patients who are enrolled from an originating study and do not require washout, the baseline will be the last visit of the originating study. However, patients who received oral systemic AD treatment as rescue therapy during the originating study require additional time to compete the

required 4-week washout and have an additional screening visit in Study JAHN (Visit 1 [Week 0]). Thus, for these patients the last nonmissing measurement on or after last visit of originating study and on or prior to first dose of study drug in Study JAHN will serve as the baseline.

For patients who do not require washout, baseline will be the last visit of the originating study. Patients who received oral systemic AD treatment as rescue therapy during the originating study require a 4-week washout and will have an additional screening visit in Study JAHN (Visit 1 [Week 0]). The assessments collected at this visit will provide the Study JAHN baseline for these patients. If labs were not collected at this visit, then the most recent lab value collected at the last visit of the originating study will be used as baseline.

Period 2

For analysis of endpoints assessed at scheduled visits in Period 2, baseline from originating study will be used for categorical endpoints, such as EASI75, Itch NRS ≥4-point improvement response, etc. Baseline from originating study will also be used for EASI-related, continuous endpoints, such as EASI change or percent change from baseline. For the other continuous endpoints such as Itch NRS and DLQI change from baseline, the analysis will be done using the Week 52 baseline. The methodology for defining Week 52 as the baseline is the same as the methodology for defining Week 0 of the originating study described above, anchoring on the first study drug administration of Period 2.

Safety Analyses

Safety analyses for treatment Period 1, and for Treatment Periods 1 and 2 combined (box plots of laboratory analytes and vital signs) will be conducted using the Study JAHN baseline, which is defined as the last nonmissing scheduled (planned) measurement on or prior to the date of first study drug administration for continuous measures by-visit analyses and non-missing measurements on or prior to the date of first study drug administration in Study JAHN for all other analyses. Safety analyses for Treatment Period 2 will be conducted for AEs only. Baseline for Treatment Period 2 is defined as nonmissing measurements on or prior to the date of first study drug administration in Treatment Period 2 of Study JAHN.

Postbaseline

Period 1(Weeks 0 to 52)

Nonmissing postbaseline efficacy data collected at scheduled visits will be used for analyses. If an assessment is missing at a scheduled visit, an unscheduled postbaseline assessment can be used provided it falls within a ±4-day window of the scheduled visit date. 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 nonmissing measure collected at the scheduled visit or an unscheduled visit falling within the visit window, the assessment is missing for that scheduled visit.

Diary endpoints were collected up to Week 16 in Period 1 of Study JAHN. Postbaseline daily diary endpoints will be the mean of weekly visit windows (diary windows) anchored on day of first dose (Day 1) for Weeks 1 through 14 as follows:

Week	Days
1	1-7
2	8-14
3	15-21
4	22-28
5	29-35
6	36-42
7	43-49
8	50-56
9	57-63
10	64-70
11	71-77
12	78-84
13	85-91
14	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 nonmissing observations. If there are <4 nonmissing 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 nonmissing observations are obtained. If after extending this diary window's lower bound to 14 days there are <4 nonmissing observations, then go to Step 2.

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 on or before a scheduled visit occurring after Week 16 will constitute the Week 16 diary window provided that window contains at least 4 nonmissing observations. If there are <4 nonmissing 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 nonmissing observations are obtained. If after extending this diary window's lower bound to 14 days there are less than 4 nonmissing observations then go to Step 3.

Step 3: If neither a Week 16 scheduled visit or an unscheduled visit to act as a surrogate for the Week 16 diary window is available, then the Week 16 window will be Day 106 to Day 112. If there are <4 nonmissing observations, the diary window's lower bound will be extended 1 day at a time to Day 99 until 4 nonmissing observations are obtained.

If the steps above do not detect a window with at least 4 nonmissing observations, then the Week 16 window is 7 days from either the Week 16 visit, the surrogate visit, or Day 106 through 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 <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 the 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 nonmissing values, the mean for Week 15 is missing and subject to imputation rules.

PostBaseline

Period 2 (Study Weeks 52-68)

Diary endpoints were collected up to Week 16 in Period 2 of JAHN. For Treatment Period 2, postbaseline daily diary endpoints will be the mean of weekly visit windows (diary windows) anchored on day of the Week 52 visit for Weeks 53 through 66 as follows:

Week of Period 2	Days in Period 2
53	1-7
54	8-14
55	15-21
56	22-28
57	29-35
58	36-42
59	43-49
60	50-56
61	57-63

62	64-70
63	71-77
64	78-84
65	85-91
66	92-98

Week 68 Daily Diary Window Construction

Week 68 daily diary window construction for Period 2 will follow the exact same steps as the window construction for Week 16 in Period 1.

Week 67 Daily Diary Window Construction

Week 67 daily diary window construction for Period 2 will follow the exact same steps as the window construction for Week 15 in Period 1.

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.

6.2.3. Analysis Methods

As few placebo patients are expected to be responders, no formal comparisons between treatment groups and placebo will be conducted for the responder and partial responder mITT populations. For the nonresponder mITT population, the following comparisons will be made for Period 1:

- for patients from originating monotherapy studies:
  - baricitinib 2 mg to baricitinib 4 mg versus baricitinib 2 mg to baricitinib 2 mg
  - baricitinib 1 mg to baricitinib 4 mg versus baricitinib 1 mg to baricitinib 2 mg
  - placebo to baricitinib 4 mg versus placebo to baricitinib 2 mg
- for patients from originating combination therapy study:
  - baricitinib 2 mg to baricitinib 4 mg versus baricitinib 2 mg to baricitinib 2 mg
  - placebo to baricitinib 4 mg versus placebo to baricitinib 2 mg

For the Period 2 substudy, the following treatment comparisons will be performed:

- for patients on baricitinib 4 mg in Period 1:
  - placebo versus baricitinib 4 mg
  - baricitinib 2 mg versus baricitinib 4 mg
- for patients on baricitinib 2 mg in Period 1:
  - placebo versus baricitinib 2 mg
  - baricitinib 1 mg versus baricitinib 2 mg

For Period 1, the main analysis method of categorical efficacy variables will use a logistic regression analysis with region, baseline disease severity from originating study (IGA 3 or IGA 4), baseline value from originating study, and treatment group in the model. For Period 2, the main analysis method of categorical efficacy variables will use a logistic regression analysis but with region, baseline disease severity at Week 52, baseline value at Week 52, and treatment

group in the model. In this case, the baseline disease severity at Week 52 can take values in 1 of 2 levels: 1) IGA of 0 or 1, and 2) IGA of 2. Firth's correction will be used in order to accommodate (potential) sparse response rates. The p-value for the odds ratio from the logistic regression model will be used for statistical inference, unless Firth's correction still results in quasi-separation. In that case, a Fisher's exact test p-value will be used for statistical inference. The difference in percentages and 100(1-alpha)% CI of the difference in percentages using the Newcombe-Wilson method without continuity correction will be reported.

For nonresponders, the main analysis method for all continuous efficacy variables will use MMRM analysis. The MMRM model will use a restricted maximum likelihood estimation. The model will include treatment, region, baseline disease severity (IGA), visit, and treatment-by-visit interaction as fixed categorical effects, and baseline and baseline-by-visit interaction as fixed continuous effects. For daily diary assessment endpoints, the model term 'visit' will include each of 16 weeks. 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, followed by autoregressive (AR[1]), followed by compound symmetry will be used. The Kenward-Roger method will be used to estimate the degrees of freedom. Treatment least square means (LSM) will be estimated within the framework of the MMRM using type 3 sums of squares. Differences in LSM between each dose of baricitinib and an appropriate comparator (and associated p-values, standard errors and 95% CI) will be used for statistical inference. The LSM difference, standard error, p-value, and 95% CI will be reported. For responders and partial responders, descriptive statistics will be provided for all continuous efficacy and health outcomes variables.

Treatment comparisons for select continuous efficacy and health outcomes variables may also be made using ANCOVA for primary and secondary objectives. For Period 1 analysis, when an ANCOVA model is used, the model includes region, baseline disease severity, treatment group, and baseline value. Treatment LSM will be estimated within the framework of the ANCOVA using type 3 sums of squares. Reported differences in LSM and associated p-values, standard errors, and 95% CI will be used for statistical inference. Analysis of covariance will be the primary method of analysis for continuous efficacy and health outcome variables for Period 2. In this Period, ANCOVA model includes region, baseline disease severity at Week 52, treatment group, and baseline value at Week 52. Baseline disease severity is defined previously.

Patients on baricitinib 2 mg or baricitinib 4 mg who are considered nonresponders (IGA >2) at entry into Study JAHN may respond with continued baricitinib treatment. Time to first response (IGA of ≤ 1) will be assessed using cumulative incidence functions for patients remaining on the same dose from the originating study. This will be used to determine when continued baricitinib treatment is futile. Missing IGA data will be replaced using the NRI rule.

For the US submission DBL (data cut-off 13Dec2019), all patients originating from Studies JAHL or JAHM are expected to have either reached the Week 52 visit or have discontinued from the study. Fisher's exact test will be used to test for differences between each baricitinib dose (including those that are randomized to a different dose) and placebo in proportions of patients

experiencing AEs) discontinuation from study drug, and for other categorical safety data. Continuous vital signs, body weight, and other continuous safety variables, including laboratory variables will be analyzed by an ANCOVA with treatment group and baseline value in the model. The significance of within-treatment-group changes from baseline will be evaluated by testing whether or not the treatment group LSM changes from baseline are different from zero; the standard error for the LSM change will also be displayed. Differences in LSM will be displayed, with the p-value associated with the LSM comparison to placebo or appropriate comparator and a 95% CI on the LSM difference also provided. In addition to the LSMs for each group, the within-group p-value for the change from baseline will be displayed.

6.2.4. Derived Data

- age (year), derived using first dose date as the reference start date and July 1st of birth year and truncated to a whole-year (integer) age. Patients whose derived age is <18 will have the required minimum age of 18 at informed consent confirmed; reporting for age, age groups, and lab 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 (years) = ([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 (year), derived using diagnosis date as the reference start date and July 1st of birth year and truncated to a whole-integer age
- diagnosis age group (<18, ≥18 to <50, ≥50 years old)
- change from baseline = postbaseline measurement at Visit x – baseline measurement from originating study
  - 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:
  ([postbaseline 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 prior to entry into originating study (yes, no)
  - Set to **yes** if the reason for discontinuation is inadequate response.

- cyclosporine intolerance prior to entry into originating study (yes, no)
  - Set to **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) prior to entry into originating study (yes, no)
  - Set to **yes** if cyclosporine was never used because of a contraindication.
- cyclosporine inadvisable prior to entry into originating study (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.
- TCNI inadequate efficacy prior to entry into originating study (yes, no)
  - Set to **yes** if the reason for discontinuation is inadequate response.
- TCNI intolerance prior to entry into originating study (yes, no)
  - Set to **yes** if the reasons for discontinuation are: intolerance to medication or contraindication (physician indicated TCNI was used and a contraindication was noted).
- TCNI contraindication (ineligible) prior to entry into originating study (yes, no)
  - Set to **yes** if TCNI was never used because of a contraindication
- TCNI inadvisable prior to entry into originating study (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.

6.3. Adjustments for Covariates

The randomization to treatment groups in the originating studies is stratified by disease severity (IGA) and geographic region. For Period 1, the covariates used in the logistic model for categorical data will include the baseline value from the originating study. When an MMRM analysis is performed, baseline value from originating study and baseline-by-visit interactions will be included as covariates. Unless otherwise specified, the statistical analysis models will adjust for these stratification variables. For Period 2, the covariates used in the logistic model for categorical data will include the baseline value from Week 52.

The baseline value from Study JAHN will be used as a covariate in the ANCOVA model for continuous safety data.

6.4. Handling of Dropouts or Missing Data

Intercurrent events (ICH E9 R1) are events which occur after randomization such that subsequent data (collected after the intercurrent event) are difficult to interpret.

Depending on the estimand being addressed, different methods will be used to handle missing data as a result of intercurrent events. Intercurrent events can occur through the following:

- application of the censoring rule (including after permanent study drug discontinuation or for responders and partial responders [on Placebo or baricitinib 1-mg] who are rescued to a higher dose [baricitinib 2 mg or baricitinib 4 mg])
- retreatment after downtitration or early withdrawal
- discontinuation
- missing an intermediate visit prior to discontinuation or rescue
- lost to follow-up

Noncensor intercurrent events are events that are not due to the application of any censoring rule (ie, the last 3 items in the list above).

Note that as efficacy and health outcome data can accrue after a patient permanently discontinues study drug or is rescued to either baricitinib 2 mg or baricitinib 4 mg (for responders and partial responders on placebo or baricitinib 1 mg), a censoring rule will be applied to all efficacy and health outcome observations subsequent to these events depending on the estimand being addressed.

The *Period 1 censoring rule* will censor efficacy after permanent study drug discontinuation or for responders and partial responders (on placebo or baricitinib 1 mg) who are rescued to a higher dose (baricitinib 2 mg or baricitinib 4 mg). This censoring rule will be applied to all efficacy endpoints conducted for the responder, partial responder, and nonresponder mITT populations (defined in Section 6.2.1). For responders and partial responders, the censoring rule will censor efficacy data after permanent study drug discontinuation or rescue date, whichever occurs first. For nonresponders, the censoring rule will censor efficacy data after permanent study drug discontinuation.

The *Period 2 censoring rule* will censor efficacy data after permanent study drug discontinuation or after retreatment, whichever occurs first. This censoring rule will be applied to efficacy endpoints and safety conducted for the Period 2 population corresponding to the withdrawal and downtitration substudy.

For patients treated in Period 2 but not eligible that enter the withdrawal and downtitration substudy, the censoring rule will censor efficacy data after permanent study drug discontinuation or rescue date, whichever occurs first.

Table JAHN.6.2 summarizes the various imputation techniques being used for select efficacy analyses. Sections 6.4.1 through 6.4.3 summarize the imputation methods.

Table JAHN.6.2. Imputation Techniques for Various Efficacy Variables

Population	Endpoints[a]	Imputation Method[b]
Responder and partial responder mITT Period 1	IGA(0,1), IGA(0,1,2), EASI75, 4-point Itch NRS improvement	NRI mLOCF [c]
	Exploratory endpoints	NRI (categorical endpoints) mLOCF [c]
Nonresponder mITT Period 1	IGA(0,1), IGA(0,1,2), EASI75, 4-point Itch NRS improvement	NRI mLOCF [c]
	Exploratory endpoints	NRI (categorical endpoints), MMRM (continuous endpoints) mLOCF [c]
Randomized Downtitration Withdrawal Substudy Population Period 2 nonsubstudy population Open label Period 2 non-substudy population	IGA(0,1), IGA(0,1,2), EASI75, 4-point Itch NRS improvement	NRI mLOCF [c]
	Exploratory endpoints	NRI (categorical endpoints) mLOCF [c]

Abbreviations: EASI = Eczema Area and Severity Index; IGA = Investigator's Global Assessment for atopic dermatitis; mITT = modified intent-to-treat; mLOCF = modified last observation carried forward; MMRM = mixed-model repeated measures; NRI = nonresponder imputation; NRS = Numeric Rating Scale.

[a] Refer to Table JAHN.6.6 and Table JAHN.6.8.

[b] All analysis with censoring.

[c] mLOCF will be used as Primary analysis in/after Feb2021 DBL.

6.4.1. Nonresponder Imputation

An NRI method imputes missing values as nonresponses and can be justified based on the composite strategy for handling intercurrent events (ICH E9 R1). This imputation procedure assumes the effects of treatment disappear after an occurrence of the intercurrent event (described in Section 6.4). For analyses which utilize censoring, randomized patients without at least 1 postbaseline observation will be defined as nonresponders for all visits. As well, patients who are missing a value prior to discontinuation or were rescued to a higher dose (if censoring on the latter) (ie, the patient is missing an intermediate visit) will be imputed as nonresponders on that visit only.

6.4.2. Mixed-Model for Repeated Measures

Mixed-model for repeated measures analyses will be performed on continuous endpoints to mitigate the impact of missing data. This approach assumes missing observations are missing at random (missingness is related to observed data) and borrows information from patients in the same treatment arm taking into account both the missingness of data through the correlation of the repeated measurements.

Essentially MMRM estimates the treatment effects had all patients remained on their initial treatment throughout the study. For this reason, the MMRM imputation implies a different estimand (hypothetical strategy [ICH E9 R1]) than the one used for NRI on categorical outcomes.

MMRM is executed after application of any censoring rules.

6.4.3. Modified Last Observation Carried Forward

An mLOCF imputation technique replaces missing data with the most recent nonmissing postbaseline assessment. The specific modification to the LOCF is data after an intercurrent event will not be carried forward thus the mLOCF is applied after the specified censoring rule is implemented. The mLOCF assumes the effect of treatment remain the same after the event that caused missing data as it was just prior to the missing data event. Analyses using mLOCF require a nonmissing baseline and at least 1 postbaseline measure otherwise the data is missing for analyses purposes. Analyses using mLOCF help ensure the number of randomized patients who were assessed postbaseline is maximized and is reasonable for this data as data directly prior to an intercurrent event (such as initiation of rescue therapy or drop out) is likely a nonefficacious response.

mLOCF will be used as the Primary analysis in/after February 2021 DBL, as NRI may underestimate potential treatment effect in a long-term study where patients drop out for multiple reasons. For period 1 and period 2 summary and analysis, mLOCF will be used as the primary analysis for all categorical endpoints (including PROs), and NRI will be used as sensitivity analysis.

6.4.4. Missingness Due to COVID-19

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

- For continuous measures, if missing due to Covid-19,
    - when they are analyzed using ANCOVA for a single timepoint, mLOCF method will be used.
    - when they are analyzed using MMRM with repeated measures, no imputation will be done.
- 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 EASI 90.
    - for primary and key efficacy measures including IGA 0 or 1 and EASI75, multiple imputation method will be used. Seed = 123456 will be used for multiple imputation. Results from multiple imputations 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 DBL, if Covid-19 is still ongoing at the time of the DBL, 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 DBL is based on the list we receive for that specific lock.

For the lock to occur on 20 May 2020, as there are very limited missing data due to Covid-19, the imputation method specified in this section (Section 6.4.4) will not be applied to this lock.

For the planned locks to occur in February 2021, if the missing data due to Covid-19 does not exceed a prespecified 5% for each of the analysis population set, then there would no special imputation method described in Section 6.4.4 applied.

For the lock in December 2021, there would be no special imputation method described in Section 6.4.4 applied.

6.5. Multicenter Studies

This study will be conducted by multiple investigators at multiple sites internationally. The countries will be categorized into geographic regions, as described in Section 5.2.

6.6. Multiple Comparisons/Multiplicity

As this study is designed to assess the long-term efficacy and safety of baricitinib in patients with AD, no adjustments for multiple comparisons will be utilized in the statistical analyses for this study.

6.7. Patient Disposition

Period 1:

An overview of patient populations will be summarized by treatment group. Frequency counts and percentages of patients excluded prior to randomization by primary reason for exclusion will be provided for patients who failed to meet study entry requirements during screening.

Patient disposition through Weeks 16, 36, and 52 will be summarized using the mITT population. Frequency counts and percentages of patients who complete the study treatment visits or discontinue early from the study along with whether they completed follow-up or did not complete follow-up will be summarized separately by treatment group for patients who are responder mITT, partial responder mITT, and for patients who are nonresponder mITT, along with their reason for study discontinuation. Rescue status for the responder mITT and partial responder mITT population will also be summarized. Frequency counts and percentages of patients who complete the treatment or discontinue treatment early will also be summarized separately by treatment group for patients who are responder mITT, partial responder mITT, and for patients who are nonresponder mITT, along with their reason for treatment discontinuation and their rescue status.

A listing of patient disposition will be provided for the ITT population, with the extent of their participation in the study and the reason for discontinuation. A listing of the ITT population with their treatment assignments for both treatment periods will also be provided.

Period 2:

For Period 2, study and study treatment disposition for Randomized Downtitration Withdrawal Substudy Population will be reported similarly as Period 1 above. Disposition for Period 2 Nonsubstudy Population and Open Label Period 2 Nonsubstudy Population will also be reported.

6.8. Patient Characteristics

Patient characteristics including demographics, baseline characteristics and pre-existing conditions will be summarized descriptively by treatment group. Period 1 summaries will be presented for the responder mITT, partial responder mITT, and nonresponder mITT populations.

Period 2 demographics and baseline characteristics will be presented by each treatment group for Randomized Downtitration Withdrawal Population, for non substudy population, for open label Period 2 non substudy population. No formal statistical comparisons will be made among treatment groups unless otherwise stated.

6.8.1. Demographics

Patient demographics will be based on entry into originating study and summarized as described above. The following demographic information will be included:

- age
- age group (<65 versus ≥65)
- age group (<65, ≥65 to <75, ≥75 to <85, ≥85)
- gender (male, female)
- race (American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, White, Multiple)
- region (as defined in Table JAHN.5.1)
- country
- 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 mITT population.

6.8.2. Baseline Disease Characteristics

Baseline disease characteristics will be reported where the baseline is from the originating study. Open-label population uses baseline from JAHN. 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)
- validated IGA for AD score
- EASI score
- SCORAD
- body surface area affected by AD
- HADS subscales
- POEM
- Itch NRS
- ADSS Item 2
- DLQI
- Skin Pain NRS
- PGI-S-AD
- prior therapy using originating study baseline (topical therapy only; systemic therapy only; topical and systemic therapy)
- use of cyclosporine prior to entry into originating study (yes, no)
- cyclosporine inadequate response prior to entry into originating study (yes, no)
- cyclosporine intolerance prior to entry into originating study (yes, no)
- cyclosporine contraindication (ineligible) prior to entry into originating study (yes, no)
- cyclosporine inadvisable prior to entry into originating study (yes, no)
- use of TCNI prior to entry into originating study (yes, no)
- TCNI inadequate response prior to entry into originating study (yes, no)
- TCNI intolerance prior to entry into originating study (yes, no)
- TCNI contraindication (ineligible) prior to entry into originating study (yes, no)
- TCNI inadvisable prior to entry into originating study (yes, no)
- vaccine prior to entry into originating study (yes, no)
- originating study 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)

6.8.3. Historical Illness and Preexisting Conditions

Historical illnesses are described in the originating studies and no additional summaries will be created for Study JAHN.

Preexisting conditions are defined as those conditions recorded in the Preexisting Conditions and Medical History electronic case report form (eCRF), or the Adverse Events eCRF with a start date prior to the first dose of the study drug and stop dates that are at or after the informed consent date or have no stop date (ie, are ongoing). For events occurring on the day of the first dose of study drug, the date and time of the onset of the event will both be used to determine if the event was preexisting. Conditions with a partial or missing start date (or time if needed) will be assumed to be 'not preexisting' unless there is evidence, through comparison of partial dates,

to suggest otherwise. Preexisting conditions will be categorized using the Medical Dictionary for Regulatory Activities (MedDRA®, most current available version) algorithmic standardized MedDRA queries or similar predefined lists of Preferred Terms (PTs) of interest. Frequency counts and percentages of patients with selected preexisting conditions will be summarized by treatment group. Analyses will be presented for the responder mITT, partial responder mITT and nonresponder mITT populations.

6.9. Treatment Compliance

Period 1 patient compliance with study medication will be assessed from Week 0 (Visit 2) to Week 52 (Visit 8) (or Early Termination if it occurs prior to Week 52), using the mITT population. Period 2 compliance will be summarized by treatment group up to Week 16 for the US filing.

All patients are expected to take 3 tablets daily from a blister pack as described in Section 5.2. Each blister pack contains 27 tablets. 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 noncompliance rate will be summarized for the mITT population by treatment group for both study periods. Subintervals 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 for all mITT patient for Week 0 to 52, and Week 52 to 200 for patients entering the randomized downtitration and withdrawal substudies. A listing of overall compliance for the period Week 0 to 200 will also be provided.

6.9.1. Background Therapy

Background TCS therapy with low-potency and/or moderate-potency TCS (eg, hydrocortisone 2.5% ointment and/or triamcinolone 0.1% cream) may be used on active lesions. Such TCS

therapy may be provided by the sponsor or the investigative sites. For more detail, see Section 7.7.1 in the protocol.

A summary of the reasons initial background therapy is used will be produced, as well as a summary of the proportion of patients with sponsor-provided TCS at each study visit and a summary of other AD therapy for the responder and partial responder mITT population. Also, a summary of time to initial use of AD background therapy postbaseline will be produced for the responder and partial responder mITT population.

The number of days TCS therapy is used for each patient is also collected on the diary device. The diary device is only dispensed up until Week 16 and is redispensed for patients who are eligible for the randomized withdrawal and downtitration substudy for an additional 16 weeks of the substudy. The proportion of time that the patients did not use TCS therapy will be summarized for these visit intervals (ie, from Week 0 to Week 16 for the responder mITT, partial responder mITT, and nonresponder mITT patients provided with low or moderate TCS). For this analysis, the date of the first entry on the diary device will be used to signify the first day of TCS therapy use.

Descriptive statistics for drug accountability of sponsor-provided TCS (topical low and moderate corticosteroids) will also be provided, including the amount utilized (dispensed tube with cap weight minus the weight of the returned tube with cap) throughout the treatment period from Week 0 to Weeks 16, 36, and 52. The total amount utilized in grams for low and moderate potency will be summarized between visits (Week 0 through Week 4, Week 4 through Week 8, Week 8 through Week 16, and so on) as well as throughout the treatment period from Week 0 through Week 52 (Period 1) and Week 52 through 68 (Period 2). The total amount utilized will also be presented for the aforementioned visit intervals, irrespective of potency. Analyses will be summarized for the responder mITT, partial responder mITT, and nonresponder mITT populations for Period 1. Additionally a summary of total amount used will be provided for the first 16 Weeks of Period 2 by treatment group for those that entered the randomized withdrawal and downtitration studies.

The dispensed weight of sponsor-provided TCS full tubes used for background therapy for the 2 different potencies (low and moderate) varies between countries due to different supply regions. Average weights of full tubes were used to determine the dispensed weights for each region. Returned tubes were weighed with cap (without the carton) to determine the amount of TCS in grams (g) used between visits. For low and moderate potency TCS, the dispensed tube weight with cap (without the carton) in Japan is 13.5 g. For countries supplied by European distributors (Austria, Czech Republic, Denmark, France, Germany, Hungary, Italy, Poland, Spain, and Switzerland), the dispensed weight of low potency TCS is 21 g and moderate potency TCS is 38 g. The remaining countries, supplied by US distributors (Argentina, Australia, India, Israel, Korea, Mexico, Russia, and Taiwan), the weight of low and moderate potency TCS is 40 g.

If a returned tube is not weighed in grams or the tube is not returned then the tube can be classified as partially used, fully used, unused, or unknown. Partially used TCS medication tubes will be considered to be 50% used whereas fully used and unused tubes will be considered as

100% and 0% used, respectively. When drug accountability is not performed for a particular tube of TCS medication or an answer of unknown is given for a tube which is not returned, that particular tube will not be included in the analysis.

6.10. Previous and Concomitant Therapy

Previous medications are described in the originating studies.

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

At each visit, concomitant therapy will be recorded for each patient. Concomitant therapy for the treatment period is defined as therapy that starts before or during the treatment period and ends during the 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.

A summary of concomitant therapy including concomitant therapies of special interest will be provided for the mITT population.

6.11. Efficacy Analyses

The general methods used to summarize efficacy data, including the definition of baseline value for assessments are described in Section 6.2.

Efficacy analyses will generally be analyzed according to the groups in Table JAHN.6.3 and patients will be analyzed according to the IP being analyzed in the objective.

Table JAHN.6.3. Efficacy Analyses: Comparisons or Treatment Paradigms

Treatment Period	Treatment Groups: Originating Study Treatment to JAHN Treatment	Comparison or Treatment Paradigms	Analysis
Treatment Period 1 (main JAHN protocol)	**Responders and Partial Responders:** PBO to PBO (PBO), BARI 1 mg to BARI 1 mg (BARI 1 mg), BARI 2 mg to BARI 2 mg (BARI 2 mg), BARI 4 mg to BARI 4 mg (BARI 4 mg)	PBO, BARI 1 mg, BARI 2 mg, BARI 4 mg; descriptive	Primary, Secondary, Exploratory
	Nonresponders: PBO to BARI 2 mg, PBO to BARI 4 mg, BARI 1 mg to BARI 2 mg, BARI 1 mg to BARI 4 mg, BARI 2 mg to BARI 2 mg, BARI 2 mg to BARI 4 mg, BARI 4 mg to BARI 4 mg	PBO to BARI 4 mg vs PBO to BARI 2 mg; BARI 1 mg to BARI 4 mg vs BARI 1 mg to BARI 2 mg; BARI 2 mg to BARI 4 mg vs BARI 2 mg to BARI 2 mg	Secondary, Exploratory
Treatment Period 2 (Randomized Withdrawal and Downtitration SubStudy	**Treatment Groups:** Bari 4 mg at Week 52 Bari 4 mg to Bari 2 mg Bari 4 mg to PBO Bari 2 mg at Week 52	 Bari 4 mg to Bari 2 mg Bari 4 mg to PBO	Secondary

Treatment Period	Treatment Groups: Originating Study Treatment to JAHN Treatment	Comparison or Treatment Paradigms	Analysis
	Bari 2 mg to Bari 1 mg Bari 2 mg to PBO	Bari 2 mg to Bari 1 mg Bari 2 mg to PBO	
Treatment Period 2 (Non substudy) (main JAHN population)	**Treatment Groups:** Bari 4 mg Bari 2 mg Bari 1 mg PBO	No comparison	Secondary, Exploratory
Treatment Period 2 (Non substudy) (open label Bari 2mg population)	**Treatment Groups:** Bari 2mg	No comparison	Secondary, Exploratory

Abbreviations: BARI = baricitinib; PBO = placebo.

Table JAHN.6.4 includes the descriptions and derivations of the primary, secondary, and exploratory efficacy outcomes.

Table JAHN.6.5 provides the detailed analyses including analysis type, method and imputation, population, time point, and comparisons for efficacy analyses.

Table JAHN.6.4. Description and Derivation of Primary, Secondary, and Exploratory Efficacy Outcomes

Measure	Description	Variable	Derivation / Comment	Imputation Approach if Missing Components
Validated Investigator's Global Assessment for AD (IGA)	The validated IGA of the patient's overall severity of their AD, 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: 0-4 0 represents "clear" 4 represents "severe"	Single item, missing if missing.
		▪ IGA (0,1) ▪ IGA (0,1,2)	▪ Observed score of 0 or 1 ▪ Observed score of 0, 1, or 2	Single item, missing if missing.
		Relapse	If IGA score at any visit after Week 52 is observed to be >2 then relapse = 'Y'	Missing if all post-Week 52 IGA assessments are missing
Eczema Area and Severity Index (EASI)	The EASI assesses objective physician estimates of 2 dimensions of atopic dermatitis – 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. 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
		▪ Change from baseline in EASI score ▪ Percent change from baseline in 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.
		EASI50	% Improvement in baseline ≥50%: % change from baseline ≤ -50	Missing if baseline or observed value is missing.
		EASI75	% Improvement in baseline ≥75%: % change from baseline ≤ -75	Missing if baseline or observed value is missing.

Measure	Description	Variable	Derivation / Comment	Imputation Approach if Missing Components
		EASI90	% Improvement in baseline ≥90%: % change from baseline ≤ -90	Missing if baseline or observed value is missing.
		Rebound	If EASI percentage change from originating study baseline at any visit after Week 52 is observed to be ≥125, then disease rebound = 'Y'	Missing if baseline or observed value is missing.
Body Surface Area (BSA) Affected by AD	Body surface area 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%-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 (0-100%) in EASI as follows: $BSA = BSA_{\text{head and neck}} / 100/0.10 + BSA_{\text{trunk}} / 100 / 0.0333 + BSA_{\text{upper limbs}} / 100 / 0.05 + BSA_{\text{lower limbs}} / 100 / 0.025$	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.
SCORing Atopic Dermatitis (SCORAD)	The SCORing Atopic Dermatitis (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 0-3 (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 analogue scales (VAS) of 0-10 where 0 is no itch or sleep loss and 10 is worst-imaginable itch or sleep loss. These 3 aspects (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 ▪ Percent change from baseline in SCORAD score	Change from baseline: observed SCORAD score – baseline SCORAD score % change from baseline: $100 \times \frac{Observed\ score - Baseline}{Baseline}$	Missing if baseline or observed value is missing.
		SCORAD75	% Improvement in baseline ≥75%: % change from baseline ≤ -75	Missing if baseline or observed value is missing.
		SCORAD90	% Improvement in baseline ≥90%: % change from baseline ≤ -90	Missing if baseline or observed value is missing.

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

Table JAHN.6.5. Description of Primary, Secondary, and Exploratory Efficacy Analyses

Measure	Variable	Population[a] (Section 6.2.1)	Analysis Method[b] (Section 6.2.3)	Time Point	Analysis Type
Validated Investigator's Global Assessment for AD (IGA)[c]	Proportion of patients with a response of IGA 0 or 1 Proportion of patients with a response of IGA 0, 1, or 2	mITT population from originating combination therapy study -responders, partial responders	Descriptive using NRI and mLOCF	Weeks 0-52	Primary for IGA (0,1) at Weeks16, 36, and 52 Secondary for IGA (0,1,2) Primary for mLOCF Sensitivity for NRI
		mITT population from originating combination therapy study -non responders	Logistic regression using NRI Descriptive using mLOCF	Weeks 0-52	
		mITT population from open-label study	Descriptive using NRI and mLOCF	Weeks 0-52	Exploratory
		mITT population per pooled Bari 2 mg-2 mg; Bari 4 mg-4 mg	Descriptive using NRI and mLOCF	Weeks 0-52	Exploratory
		Long-term Bari 4 mg population	Descriptive using Observed, NRI and mLOCF	Weeks 52-200	Exploratory
		Randomized Downtitration Withdrawal Substudy Population Randomized Downtitration Withdrawal Substudy Population with Week 52 IGA of 0 or 1	Logistic regression using NRI and mLOCF	Week 52-200	Secondary for Week 68 Exploratory for Other weeks Primary for mLOCF Sensitivity for NRI
		Period 2 non-substudy population Open-label Period 2 non-substudy population	Descriptive using observed, mLOCF and NRI	Week 52 to 200	Secondary, exploratory Primary for mLOCF Sensitivity for NRI
	Time to retreatment (time to IGA ≥3)	Re-Treatment Substudy Population	Descriptive	Downtitration and withdrawal period	Secondary

Measure	Variable	Population[a] (Section 6.2.1)	Analysis Method[b] (Section 6.2.3)	Time Point	Analysis Type
	Proportion of patients with IGA ≤2 within 16 weeks after retreatment	Re-Treatment Substudy Population	Descriptive	Week 52 to 200	Secondary
Eczema Area and Severity Index (EASI)[c]	Proportion of patients achieving EASI75 (achieving EASI percent change from baseline < -75) Proportion of patients achieving EASI50 Proportion of patients achieving EASI90	mITT population from originating combination therapy study -responders, partial responders	Descriptive using NRI and mLOCF	Weeks 0-52	Secondary for EASI75 Exploratory for EASI50 and EASI90 Secondary for mLOCF Sensitivity for NRI
		mITT population from originating combination therapy study -non responders	Logistic regression using NRI Descriptive using mLOCF	Weeks 0-52	
		mITT population from open-label study	Descriptive using NRI and mLOCF	Weeks 0-52	Exploratory
		mITT population per pooled Bari 2 mg-2 mg; Bari 4 mg-4 mg	Descriptive using NRI and mLOCF	Weeks 0-52	Exploratory
		Long term Baricitinib 4 mg population	Descriptive using Observed, NRI and mLOCF	Weeks 52-200	Exploratory
		Randomized Downtitration Withdrawal Substudy Population Randomized Downtitration Withdrawal Substudy Population with Week 52 IGA of 0 or 1	Logistic regression using NRI and mLOCF	Week 52-200	Secondary for Week 68 Exploratory for Other weeks Secondary for mLOCF Sensitivity for NRI
		Period 2 non-substudy population Open-label Period 2 non-substudy population	Descriptive using observed, NRI and mLOCF	Week 52-200	Secondary, exploratory Secondary for mLOCF Sensitivity for NRI
	Change from baseline in EASI score Percent change from baseline in EASI score	mITT population from originating combination therapy study -responders, partial responders	Descriptive using observed and mLOCF	Weeks 0-52	Exploratory Sensitivity for NRI

Measure	Variable	Population[a] (Section 6.2.1)	Analysis Method[b] (Section 6.2.3)	Time Point	Analysis Type
		mITT population from originating combination therapy study -non responders	MMRM using observed Descriptive using mLOCF	Weeks 0-52	Exploratory Sensitivity for NRI
Eczema Area and Severity Index (EASI)[c]	Change from baseline in EASI score Percent change from baseline in EASI score	mITT population from open-label study	Descriptive using observed and mLOCF	Weeks 0-52	Exploratory
		Period 2 non-substudy population Open-label Period 2 non-substudy population	Descriptive using observed and mLOCF	Weeks 52-200	Exploratory
		Long term Baricitinib 4 mg population	Descriptive using observed and mLOCF	Weeks 52-200	Exploratory
		Randomized Downtitration Withdrawal Substudy Population Randomized Downtitration Withdrawal Substudy Population with Week 52 IGA of 0 or 1	ANCOVA using mLOCF	Week 52-200	Exploratory
Body Surface Area (BSA) Affected by AD	BSA Score Change from baseline in BSA Score	mITT population from originating combination therapy study - responders, partial responders	Descriptive using observed and mLOCF	Weeks 0-52	Exploratory
		mITT population from originating combination therapy study - nonresponders	MMRM using observed Descriptive using mLOCF	Weeks 0-52	Exploratory
		mITT population from open-label study	Descriptive using observed and MLOCF	Weeks 0-52	Exploratory
SCORing Atopic Dermatitis (SCORAD)	SCORAD score Change from baseline in SCORAD score Percent change from baseline in SCORAD score	mITT population from originating combination therapy study - responders, partial responders	Descriptive using observed and MLOCF	Weeks 0-52	Exploratory
		mITT population from originating combination therapy study - nonresponders	MMRM using observed Descriptive using mLOCF	Weeks 0-52	Exploratory
SCORing Atopic Dermatitis	Change from baseline in SCORAD score	mITT population from open-label study	Descriptive using observed and mLOCF	Weeks 0-52	Exploratory

Measure	Variable	Population[a] (Section 6.2.1)	Analysis Method[b] (Section 6.2.3)	Time Point	Analysis Type
(SCORAD)[c]	Percent change from baseline in SCORAD score	Period 2 non-substudy population Open-label Period 2 non-substudy population	Descriptive using observed and mLOCF	Weeks 52-200	Exploratory
	SCORAD score, Pruritus & Sleep Loss Percent change from baseline in SCORAD score, Pruritus & Sleep Loss	Long term Baricitinib 4 mg population	Descriptive using observed and mLOCF	Weeks 52-200	Exploratory
	Proportion of patients achieving SCORAD75 Proportion of patients achieving SCORAD90	mITT population from originating combination therapy study - responders, partial responders	Descriptive using NRI and mLOCF	Weeks 0-52	Exploratory
		mITT population from originating combination therapy study - nonresponders	Logistic regression using NRI Descriptive using mLOCF	Weeks 0-52	Exploratory
		mITT population from open-label study	Descriptive using NRI and mLOCF	Weeks 0-52	Exploratory
		Period 2 non-substudy population Open-label Period 2 non-substudy population	Descriptive using observed, NRI and mLOCF	Weeks 52-200	Exploratory
		Long term Baricitinib 4 mg population	Descriptive using NRI, observed and mLOCF	Weeks 52-200	Exploratory

Abbreviations: AD = atopic dermatitis; ANCOVA = analysis of covariance; Bari = baricitinib; DBL = database lock; EE = efficacy evaluable; Feb = February; mITT = modified intent-to-treat; mLOCF = modified last observation carried forward; MMRM = mixed-model repeated measures; NRI = nonresponder imputation.

[a] Populations from originating studies for reporting periods of Weeks 0-52 are on the mITT population; they are reported by responder status from originating study: responders, partial responders, responders + partial responders and nonresponders. Populations for visit-wise analyses for those who entered the randomized withdrawal and downtitration study are based on efficacy evaluable sets for each of Week 56, 60, 64, and 68 (defined in Section 6.2.1).

[b] All categorical endpoints use NRI, unless otherwise stated.

[c] Descriptive using observed data for the endpoints beyond Week 104, unless otherwise stated.

6.11.1. Primary Outcome and Methodology

The validated IGA for AD uses the clinical characteristics of erythema, papulation/induration, oozing/crusting, and lichenification to produce a single-item score ranging from 0 to 4. The primary analysis of the study is to estimate the effect of long-term therapy with baricitinib on responders and partial responders at the entry of Study JAHN by evaluating the proportion of patients with a response of IGA 0 or 1 assessed at Weeks 16, 36, and 52 using the mITT population, assuming that treatment response disappears after the patient is rescued or discontinued study or study treatment. This will serve as the primary estimand. In this estimand, missing data due to the application of the censoring rule and the occurrence of other noncensor intercurrent events will be imputed using the NRI method described in Section 6.4.1. A logistic regression analysis as described in Section 6.2.3 will be used for the comparisons. The odds ratio, the corresponding 95% CIs and p-value, as well as the treatment differences and the corresponding 95% CIs, will be reported.

For analyses on and after the February 2021 DBL, mLOCF will be used as Primary approach for all categorical endpoints (including PROs) and NRI will be used as supportive sensitivity.

6.11.2. Secondary and Exploratory Efficacy Analyses

There will be no adjustment for multiple comparisons for any analyses. The secondary and exploratory efficacy analyses are detailed in Table JAHN.6.4. Health outcomes analyses are described in Section 6.12.

For patients entering the downtitration and withdrawal substudy who need to be retreated, a summary of treatment responses containing proportion of patients with IGA (0, 1, 2), IGA (0, 1), and EASI75 response will be provided. A summary of the time to retreatment will also be provided.

Note that the exploratory efficacy analyses on a continuous outcome that is measured over time use a slightly different estimand which is to estimate the effect of long-term therapy with baricitinib on a specific endpoint evaluated at specified time points using the mITT population assuming all patients remained in their treatment throughout the defined period of the study.

6.11.3. Sensitivity Analyses

Sensitivity analyses are included to demonstrate robustness of analyses methods using different censoring rules, populations, and analyses assumptions. Sensitivity analyses for select outcomes have been previously described and include the following:

- analysis of select continuous outcomes with ANCOVA (Section 6.2.3), with missing data imputed using mLOCF (Section 6.4.3)
- analysis of select continuous outcomes in the randomized withdrawal and downtitration substudy using Week 52 as the baseline
- analysis of selected outcomes (IGA, EASI, ITCH-NRS), with missing data imputed using mLOCF

For analyses on and after the February 2021 DBL, NRI will be used as supportive sensitivity.

6.12. 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 6.2.

Health outcomes and quality-of-life measures will generally be analyzed according to the formats discussed in Section 6.2.

Table JAHN.6.6 includes the descriptions and derivations of the health outcomes and quality-of-life measures.

Table JAHN.6.7 provides the detailed analyses including analysis type, method and imputation, population, time point, and comparisons for health outcomes and quality-of-life measures.

Table JAHN.6.6. Description and Derivation of Health Outcomes and Quality-of-Life Measures

Measure	Description	Variable	Derivation / Comment	Imputation Approach if Missing Components
Itch Numeric Rating Scale (NRS)	The Itch Numeric Rating Scale (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 6.2.2 for details on how to calculate the weekly score which will be used in the continuous analysis.	▪ Itch NRS score	Single item. Range 0-10. Refer to Section 6.2.2 on how to derive the weekly score.	Refer to Section 6.2.2 on how to derive the weekly 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 Numeric Rating Scale (NRS)	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 6.2.2 for details on how to calculate the weekly score which will be used in the continuous analysis.	▪ Skin Pain NRS score	Single item; range 0-10. Refer to Section 6.2.2 on how to derive the weekly score.	Refer to Section 6.2.2 on how to derive the weekly 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.

Measure	Description	Variable	Derivation / Comment	Imputation Approach if Missing Components
Atopic Dermatitis Sleep Scale (ADSS)	The Atopic Dermatitis Sleep Scale (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 0-29 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 6.2.2 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 0-4; Item 2, range 0-29; Item 3, range 0-4. Refer to Section 6.2.2 on how to derive the weekly score.	Refer to Section 6.2.2 on how to derive the weekly 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.
Patient Global Impression of Severity–Atopic Dermatitis (PGI-S-AD)	The PGI-S-AD is a single-item question asking the patient how they would rate their overall AD symptoms over the past 24 hours. The 5 categories of responses range from "no symptoms" to "severe." Refer to Section 6.2.2 for details on how to calculate the weekly score which will be used in the continuous analysis.	PGI-S-AD score	Single item. Range 1-5. Refer to Section 6.2.2 on how to derive the weekly score.	Refer to Section 6.2.2 on how to derive the weekly score.
		Change from baseline in PGI-S-AD	Change from baseline: observed PGI-S-AD score – baseline PGI-S-AD score	Missing if baseline or observed value is missing.

Measure	Description	Variable	Derivation / Comment	Imputation Approach if 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 1-7, range 0-28.	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.
Hospital Anxiety Depression Scale (HADS)	The HADS is a 14-item self-assessment scale that determines the levels of anxiety and depression that a patient is experiencing over the past week. The HADS utilizes a 4-point Likert scale (eg, 0-3) for each question and is intended for ages 12-65 years (Zigmond and Snaith 1983; White et al. 1999). Scores for each domain (anxiety and depression) can range from 0-21, with higher scores indicating greater anxiety or depression (Zigmond and Snaith 1983; Snaith 2003).	HADS score for anxiety and depression domains	Anxiety domain score is sum of the 7 anxiety questions, range 0-21; Depression domain score is sum of the 7 depression questions, range 0-21.	N/A – partial assessments cannot be saved.
		• Change from baseline in HADS domain	Change from baseline: observed HADS domain score – baseline HADS domain score	Missing if baseline or observed value is missing.

Measure	Description	Variable	Derivation / Comment	Imputation Approach if Missing Components
Dermatology Life Quality Index (DLQI)	The Dermatology Life Quality Index (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 0-1 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, and 0 respectively.	N/A – partial assessments cannot be saved.
		• Personal relationships domain	Sum of Questions 8 and 9, range 0-6.	N/A – partial assessments cannot be saved.
		• Treatment domain	Question 10, range 0-3.	N/A – partial assessments cannot be saved.
		• DLQI total score	DLQI total score: sum of all 6 DLQI domain scores, range 0-30.	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.

Measure	Description	Variable	Derivation / Comment	Imputation Approach if Missing Components
Work Productivity and Activity Impairment: Atopic Dermatitis (WPAI-AD)	The Work Productivity and Activity Impairment Questionnaire–Atopic Dermatitis (WPAI-AD) records impairment due to AD during the past 7 days. The WPAI-AD consists of 6 items grouped into 4 domains: absenteeism (work time missed), presenteeism (impairment at work/reduced on-the-job effectiveness), work productivity loss (overall work impairment/absenteeism plus presenteeism), and activity impairment. Scores are calculated as impairment percentages (Reilly et al. 1993), with higher scores indicating greater impairment and less productivity.	▪ Employment status	Question (Q)1	Single item, missing if missing.
		• Change in employment status	Employed at baseline and remained employed: Q1 = 1 at postbaseline visit and at baseline visit. Not employed at baseline and remain unemployed: Q1 = 0 at postbaseline visit and at baseline visit.	Missing if baseline or observed value is missing.
		• Percentage of absenteeism	Percent work time missed due to problem: (Q2 / [Q2 + Q4]) * 100	If Q2 or Q4 is missing, then missing.
		• Change from baseline in absenteeism	Change from baseline: observed absenteeism – baseline absenteeism	Missing if baseline or observed value is missing.
		• Percentage of presenteeism	Percent impairment (reduced productivity while at work) while working due to problem: (Q5 / 10) * 100	If Q5 is missing, then missing.
		• Change from baseline in presenteeism	Change from baseline: observed presenteeism – baseline absenteeism	Missing if baseline or observed value is missing.
		• Overall work impairment	Percent overall work impairment (combines absenteeism and presenteeism) due to problem: (Q2 / [Q2+ Q4] + [(1 – Q2 / (Q2 + Q4)) * (Q5 / 10)]) * 100	If Q2, Q4, or Q5 is missing, then missing.
		• Change from baseline in work impairment	Change from baseline: observed work impairment – baseline work impairment	Missing if baseline or observed value is missing.
		• Percentage of impairment in activities	Percent activity impairment (performed outside of work) due to problem: (Q6 / 10) * 100	If Q6 is missing, then missing.
		• Change from baseline in impairment in activities	Change from baseline: observed impairment in activities – baseline impairment in activities	Missing if baseline or observed value is missing.

Measure	Description	Variable	Derivation / Comment	Imputation Approach if Missing Components
European Quality of Life–5 Dimensions–5 Levels (EQ-5D-5L)	The EQ-5D-5L is a standardized measure of health status that provides 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 or her current health state using a 0-100 mm VAS. The descriptive system comprises the following 5 dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each dimension has 5 levels: no problems, slight problems, moderate problems, severe problems, and extreme problems. The respondent is asked to indicate his or 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 1-5 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 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	▪ EQ-5D mobility ▪ EQ-5D self-care ▪ EQ-5D usual activities ▪ EQ-5D pain/discomfort ▪ EQ-5D anxiety/depression	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 1-5 have no arithmetic properties and should not be used as a primary score.	Each dimension is a single item, missing if missing.
		• EQ-5D VAS	Single item. Range 0-100. 0 represents "worst health you can imagine" 100 represents "best health you can imagine"	Single item, missing if missing.
		• Change from baseline in EQ-5D VAS	Change from baseline: observed EQ-5D VAS score – baseline EQ-5D VAS score	Missing if baseline or observed value is missing.
		• EQ-5D-5L UK population-based index score (health state index)	Derive EQ-5D-5L UK Population-based index score according to the link by using the UK algorithm to produce a patient-level index score between -0.59 and 1.0 (continuous variable).	N/A – partial assessments cannot be saved on the eCOA tablet.
		• Change from baseline in EQ-5D-5L UK population-based index score	Change from baseline: observed EQ-5D-5L UK score – baseline EQ-5D-5L UK score	Missing if baseline or observed value is missing.
		• EQ-5D-5L US population-based index score (health state index)	Derive EQ-5D-5L US Population-based index score according to the link by using the US algorithm to produce a patient-level index score between -0.11 and 1.0 (continuous variable).	N/A – partial assessments cannot be saved on the eCOA tablet.

Measure	Description	Variable	Derivation / Comment	Imputation Approach if Missing Components
	weights) to each of the levels in each dimension (Herdman et al. 2011; EuroQol Group 2015 [WWW]).	• Change from baseline in EQ-5D-5L US population-based index score	Change from baseline: observed EQ-5D-5L US score – baseline EQ-5D-5L US score	Missing if baseline or observed value is missing.

Abbreviations: AD = atopic dermatitis; eCOA = electronic clinical outcomes assessment; N/A = not applicable, QoL = quality of life; UK = United Kingdom; US = United States.

Table JAHN.6.7. Description of Health Outcomes and Quality-of-Life Measures Analyses

Measure	Variable	Population[a] (Section 6.2.1)	Analysis Method[b] (Section 6.2.3)	Time Point	Analysis Type
Itch Numeric Rating Scale (NRS)	Itch NRS score Change from baseline in Itch NRS score Percent change from baseline in Itch NRS score	mITT population entering the downtitration and withdrawal study (Period 2) mITT population entering the downtitration and withdrawal study (Period 2) with Week 52 IGA of 0 or 1	ANCOVA using mLOCF	Weeks 52-68	Exploratory analysis
	Proportion of patients with a 4-point improvement from baseline of originating study in Itch NRS in subgroup of patients who had baseline Itch NRS ≥4	mITT population entering the downtitration and withdrawal study (Period 2) mITT population entering the downtitration and withdrawal study (Period 2) with Week 52 IGA of 0 or 1	Logistic regression using NRI and mLOCF	Weeks 52-68	Exploratory analysis
Skin Pain NRS	Skin Pain NRS score Change from baseline in Skin Pain NRS score	mITT population entering the downtitration and withdrawal study (Period 2) mITT population entering the downtitration and withdrawal study (Period 2) with Week 52 IGA of 0 or 1	ANCOVA using mLOCF	Weeks 52-68	Exploratory analysis
Skin Pain NRS	Proportion of patients with a 4-point improvement from baseline of originating study in Skin Pain in subgroup of patients who had baseline Skin Pain NRS ≥4	mITT population entering the downtitration and withdrawal study (Period 2) mITT population entering the downtitration and withdrawal study (Period 2) with Week 52 IGA of 0 or 1	Logistic regression using NRI and mLOCF	Weeks 52-68	Exploratory analysis

Measure	Variable	Population[a] (Section 6.2.1)	Analysis Method[b] (Section 6.2.3)	Time Point	Analysis Type
Atopic Dermatitis Sleep Scale (ADSS)	ADSS item scores Change from baseline in ADSS item scores	mITT population entering the downtitration and withdrawal study (Period 2) mITT population entering the downtitration and withdrawal study (Period 2) with Week 52 IGA of 0 or 1	ANCOVA using mLOCF	Weeks 52-68	Exploratory analysis
	Proportion of patients with a 2-point improvement from baseline of originating study in item 2 score of ADSS, in a subgroup of patients with baseline item 2 score of ≥2 Proportion of patients with a 1.5-point improvement from baseline of originating study in Item 2 score of ADSS, in a subgroup of patients with baseline item 2 score of ≥1.5	mITT population entering the downtitration and withdrawal study (Period 2) mITT population entering the downtitration and withdrawal study (Period 2) with Week 52 IGA of 0 or 1	Logistic regression using NRI and mLOCF	Weeks 52-68	Exploratory analysis
Patient Global Impression of Severity–Atopic Dermatitis (PGI-S-AD)	PGI-S-AD score Change from baseline in PGI-S-AD score	mITT population entering the downtitration and withdrawal study (Period 2) mITT population entering the downtitration and withdrawal study (Period 2) with Week 52 IGA of 0 or 1	ANCOVA using mLOCF	Weeks 52-68	Exploratory analysis
Patient-Oriented Eczema Measure (POEM)[c]	POEM score Change from baseline in POEM score	mITT population from originating combination therapy study -responders, partial responders	Descriptive using observed and mLOCF	Weeks 0-52	Exploratory analysis

Measure	Variable	Population[a] (Section 6.2.1)	Analysis Method[b] (Section 6.2.3)	Time Point	Analysis Type
		mITT population from originating combination therapy study -non responders	MMRM using observed Descriptive using mLOCF	Weeks 0-52	Exploratory analysis
		mITT population from open-label study	Descriptive using observed and mLOCF	Weeks 0-52	Exploratory analysis
		Randomized Downtitration Withdrawal Substudy Population Randomized Downtitration Withdrawal Substudy Population with Week 52 IGA of 0 or 1	ANCOVA using mLOCF	Weeks 52-200	Exploratory analysis
		Period 2 non-substudy population Open-label Period 2 non-substudy population	Descriptive using observed and mLOCF	Week 52 to 200	Exploratory
	Proportion of patients with a 4-point improvement in POEM score from baseline of originating study in a subgroup of patients with baseline POEM ≥4	mITT population from originating combination therapy study - responders, partial responders	Descriptive using observed and mLOCF	Weeks 0-52	Exploratory analysis
		mITT population from originating combination therapy study - nonresponders	Logistic regression using NRI Descriptive using mLOCF	Weeks 0-52	Exploratory analysis
		mITT population from open-label study	Descriptive using observed and mLOCF	Weeks 0-52	Exploratory analysis

Measure	Variable	Population[a] (Section 6.2.1)	Analysis Method[b] (Section 6.2.3)	Time Point	Analysis Type
Patient-Oriented Eczema Measure (POEM)[c]	Proportion of patients with a 4-point improvement in POEM score from baseline of originating study in a subgroup of patients with baseline POEM ≥4	All patients Randomized Downtitration Withdrawal Substudy Population Randomized Downtitration Withdrawal Substudy Population with Week 52 IGA of 0 or 1	Logistic regression using NRI and mLOCF	Weeks 52-200	Exploratory analysis
		Period 2 non-substudy population Open-label Period 2 non-substudy population	Descriptive using Observed, mLOCF and NRI	Week 52 to 200	Exploratory Primary for mLOCF Sensitivity for NRI
Hospital Anxiety Depression Scale (HADS)	Observed and change from baseline in HADS domain scores Change from baseline in HADS total score	mITT population from originating combination therapy study - responders, partial responders	Descriptive using observed and mLOCF	Weeks 0-52	Exploratory analysis
		mITT population from originating combination therapy study - nonresponders	MMRM using observed Descriptive using mLOCF	Weeks 0-52	Exploratory analysis
		mITT population from open-label study	Descriptive using observed and mLOCF	Weeks 0-52	Exploratory analysis
		Period 2 non-substudy population Open-label Period 2 non-substudy population	Descriptive using observed and mLOCF	Week 52 to 104	Exploratory
Dermatology Life Quality Index (DLQI)[c]	DLQI total score Observed and change from baseline in domain scores	mITT population from originating combination therapy study - responders, partial responders	Descriptive using observed and mLOCF	Weeks 0-52	Exploratory analysis

Measure	Variable	Population[a] (Section 6.2.1)	Analysis Method[b] (Section 6.2.3)	Time Point	Analysis Type
		mITT population from originating combination therapy study - nonresponders	MMRM using observed Descriptive using mLOCF	Weeks 0-52	Exploratory analysis
		mITT population from open-label study	Descriptive using observed and mLOCF	Weeks 0-52	Exploratory analysis
		Randomized Downtitration Withdrawal Substudy Population Randomized Downtitration Withdrawal Substudy Population with Week 52 IGA of 0 or 1	ANCOVA using mLOCF	Weeks 52-200	Exploratory analysis
		Period 2 non-substudy population Open-label Period 2 non-substudy population	Descriptive using observed and mLOCF	Week 52 to 200	Exploratory
		Long term Baricitinib 4 mg population	Descriptive using observed and mLOCF	Weeks 52-200	Exploratory
	Proportion of patients with DLQI (0,1) Proportion of patients with a 4-point improvement from baseline in originating study in a subgroup of patients with baseline DLQI score ≥4	mITT population from originating combination therapy study - responders, partial responders	Descriptive using observed and mLOCF	Weeks 0-52	Exploratory analysis

Measure	Variable	Population[a] (Section 6.2.1)	Analysis Method[b] (Section 6.2.3)	Time Point	Analysis Type
		mITT population from originating combination therapy study - nonresponders	Logistic regression using NRI Descriptive using mLOCF	Weeks 0-52	Exploratory analysis
		mITT population from open-label study	Descriptive using observed and mLOCF	Weeks 0-52	Exploratory analysis
Dermatology Life Quality Index (DLQI)[c]	Proportion of patients with DLQI (0,1) Proportion of patients with a 4-point improvement from baseline in originating study in a subgroup of patients with baseline DLQI score ≥4	Randomized Downtitration Withdrawal Substudy Population Randomized Downtitration Withdrawal Substudy Population with Week 52 IGA of 0 or 1	Logistic regression using NRI and mLOCF	Weeks 52-200	Exploratory analysis
		Period 2 non-substudy population Open-label Period 2 non-substudy population	Descriptive using mLOCF and NRI	Week 52 to 200	Exploratory Primary for mLOCF Sensitivity for NRI
		Long term Baricitinib 4 mg population	Descriptive using observed, NRI and mLOCF	Weeks 52-200	Exploratory
Work Productivity and Activity Impairment: Atopic Dermatitis (WPAI-AD)	Observed and change from baseline in employment status: ▪ absenteeism ▪ presenteeism ▪ overall work impairment impairment in activities	mITT population from originating combination therapy study - responders, partial responders	Descriptive using observed and mLOCF	Weeks 0-52	Exploratory analysis
		mITT population from originating combination therapy study - nonresponders	MMRM using observed Descriptive using mLOCF	Weeks 0-52	Exploratory analysis

Measure	Variable	Population[a] (Section 6.2.1)	Analysis Method[b] (Section 6.2.3)	Time Point	Analysis Type
		mITT population from open-label study	Descriptive using observed and mLOCF	Weeks 0-52	Exploratory analysis
European Quality of Life–5 Dimensions–5 Levels (EQ-5D-5L)	Observed and change from baseline in: ▪ EQ-5D mobility ▪ EQ-5D self-care ▪ EQ-5D usual activities ▪ EQ-5D pain/discomfort EQ-5D anxiety/depression	mITT population from originating combination therapy study - responders, partial responders	Descriptive using observed and mLOCF	Weeks 0-52	Exploratory analysis
		mITT population from originating combination therapy study - nonresponders	MMRM using observed Descriptive using mLOCF	Weeks 0-52	Exploratory analysis
		mITT population from open-label study	Descriptive using observed and mLOCF	Weeks 0-52	Exploratory analysis

Abbreviations: ANCOVA = analysis of covariance; EE = efficacy evaluable; IGA = Investigator's Global Assessment; mITT = modified intent-to-treat; mLOCF = modified last observation carried forward; MMRM = mixed-model repeated measures; NRI = nonresponder imputation; VAS = visual analog scale.

[a] Populations from originating studies for reporting periods of Week 0-52 are on the mITT population; they are reported by responder status from originating study: responders, partial responders, responders + partial responders, and nonresponders.

[b] All categorical endpoints use NRI, unless otherwise stated.

[c] Descriptive using observed data for the endpoints beyond Week 104, unless otherwise stated.

6.13. Bioanalytical and Pharmacokinetic/Pharmacodynamic Methods

Pharmacokinetic (PK), pharmacodynamic (PD), and biomarker analyses will not be conducted.

6.14. Safety Analyses

Detailed analyses and discussion of Study JAHN safety data are more thoroughly assessed in the context of combining the safety data from the originating Studies, JAHL, JAHM, and JAIY with the safety data from Study JAHN. The planned analyses are described in the compound level safety standards.

For the purpose of Study JAHN alone, the following are planned to be analyzed:

- duration of exposure by Treatment Periods 1 and 2
- overview of AEs by Treatment Periods 1 and 2
- treatment-emergent adverse events (TEAEs) by PT nested within System Organ Class (SOC) by Treatment Periods 1 and 2
- serious adverse events (SAEs) by PT nested within SOC by Treatment Periods 1 and 2, and listing of SAEs in Treatment Periods 1 and 2 combined (one listing)
- AEs leading to permanent study drug discontinuation by PT nested within SOC for Treatment Period 1, and listing of AEs leading to permanent study drug discontinuation and discontinuation from the study in Treatment Periods 1 and 2 combined (one listing)
- listing of deaths in Treatment Periods 1 and 2 combined (1 listing), if any
- clinical laboratory evaluation and vital signs in terms of box plots, descriptive and treatment-emergent summaries in Treatment Period 1
- clinical laboratory and vital signs in terms of box plots and descriptive summaries in Treatment Periods 1 and 2 combined

The general methods used to summarize safety data, including the definition of baseline value, are described in Section 6.2.

Safety analyses will be conducted according to the safety population defined in Section 6.2.1 and the treatment groups defined in Table JAHN.6.1 in Section 6.2.1.

Unless otherwise specified, by-visit summaries will include planned on-treatment visits. For tables that summarize events (such as AEs), on-treatment and follow-up data will be included up to 30 days after the last dose of treatment or up to dose or treatment change, where applicable.

For the February 2021 DBL and Final DBL summary and analyses, safety analyses will be done for the disposition summary and the below AE tables summarized in Period 1 and Period 2, respectively. Other detailed safety assessments will be implemented in the integrated safety analysis plan with other studies (I4V-MC-JAIN, I4V-MC-JAIW, and I4V-MC-JAIX).

- overview of AEs
- TEAEs
- AEs leading to permanent discontinuation of study drug
- SAEs

6.14.1. Extent of Exposure

Duration of exposure (in days) for Treatment Periods 1 and 2 will be calculated as follows:

- duration of exposure to IP in Treatment Period 1: minimum (date of last dose of study drug, Week 52 date) – date of first dose of study drug in Study JAHN + 1.
- duration of exposure to IP in Treatment Period 2: minimum (date of last dose of study drug [>Week 52 date], >Week 52 date) – date of first dose of study drug in Study JAHN in Treatment Period 2 + 1

Last dose of study drug is calculated as last date on study drug or censored at dose or treatment change according to the treatment groups defined in Section 6.2.1. For patients discontinuing treatment and the treatment stop date is missing, the duration of exposure will use date of last visit in exposure calculation. For patients whose treatment discontinuation is unknown because they were lost to follow-up, the date of the visit (scheduled or unscheduled) just prior to the last visit (if the treatment stop date is missing) is used for this calculation.

Descriptive statistics will be provided for patient-days of exposure. Total patient-years of exposure (PYE) will be reported for each treatment group by Treatment Periods 1 and 2 for duration of exposure.

Exposure ranges for Treatment Period 1 will be summarized as follows:

- >0 weeks, ≥4 weeks, ≥8 weeks, ≥12 weeks, ≥16 weeks, ≥24 weeks, ≥32 weeks and ≥52 weeks
- >0 to <4 weeks, ≥4 weeks to <8 weeks, ≥8 weeks to <12 weeks, ≥12 weeks to <16 weeks, ≥16 weeks to <24 weeks, ≥24 weeks to <32 weeks, ≥32 weeks to <52 weeks, and ≥52 weeks

Exposure ranges for Treatment Period 2 will be summarized as follows:

- >52 weeks, ≥56 weeks, ≥60 weeks, ≥64 weeks, ≥68 weeks, ≥76 weeks, ≥84 weeks, ≥92 weeks, ≥104 weeks, ≥120 weeks, ≥136 weeks, ≥152 weeks, ≥168 weeks, ≥184 weeks, and ≥200 weeks
- ≥52 to <56 weeks, ≥56 weeks to <60 weeks, ≥60 weeks to <64 weeks, ≥64 weeks to <68 weeks, ≥68 weeks to <76 weeks, ≥76 weeks to <84 weeks, ≥84 weeks to <92 weeks, ≥92 weeks to <104 weeks, ≥104 weeks to <120 weeks, ≥120 weeks to <136 weeks, ≥136 weeks to <152 weeks, ≥152 weeks to <168 weeks, ≥168 weeks to <184 weeks, ≥184 weeks to <200 weeks, and ≥200 weeks

Overall exposure by Treatment Periods 1 and 2 will be summarized in total PYE, which is calculated according to the following formula:

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

6.14.2. 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 DBL. Severity of AEs is recorded as mild, moderate, or severe. TEAEs will be analyzed by Treatment Periods 1 and 2.

A TEAE is defined as an event that first occurred or worsened in severity after the first dose of study treatment in Study JAHN 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 upon 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 JAHN 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 treatment emergence (ie, 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 rating 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 relative frequency (ie, 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.

In an overview table, the number and percentage of patients in the safety population who experienced death, an SAE, any TEAE, discontinuation from the study due to an AE, permanent discontinuation from study drug due to an AE, or a severe TEAE will be summarized by treatment group.

The number and percentage of patients with TEAEs will be summarized by treatment group for both Treatment Periods 1 and 2 by MedDRA PT nested within SOC with decreasing frequency in SOC, and events ordered within each SOC by decreasing frequency in the baricitinib 4-mg group.

6.14.2.1. Common Adverse Events

Common TEAEs will not be analyzed for Study JAHN data alone.

6.14.2.2. Serious Adverse Event Analyses

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

- death
- initial or prolonged inpatient hospitalization
- a life-threatening experience (ie, 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. See examples in the ICH E2A guideline Section 3B.

The number and percentage of patients who experienced any SAE will be summarized by treatment group for each Treatment Periods 1 and 2 using MedDRA PT nested within SOC. Events will be ordered by decreasing frequency in the baricitinib 4-mg group within decreasing frequency in 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.

6.14.2.3. 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 group for Treatment Period 1 using MedDRA PT nested within SOC. Events will be ordered by decreasing frequency in the baricitinib 4-mg group within decreasing frequency in SOC.

A listing of all AEs leading to permanent discontinuation from the study drug or from the study will be provided for both Treatment Periods 1 and 2.

6.14.2.4. Criteria for Notable Patients

Patient narratives will be provided for all patients who experience certain “notable” events prior to data cutoff for the submission. See compound level safety standards for list of criteria.

6.14.3. Clinical Laboratory Evaluation

For the categorical laboratory analyses (shift and treatment emergent), 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 laboratory analyses (eg, change from baseline by time point) is defined as the treatment period excluding off-drug follow-up time.

All laboratory tests will be presented using the International Système (SI) and US conventional (CN) units. The performing central laboratory reference ranges will be used to define the low and high limits.

There is one special circumstance for laboratory values to be derived based on regularly scheduled, protocol-specified analytes. The low-density lipoprotein/high-density lipoprotein (LDL/HDL) ratio will be derived as the ratio of LDL cholesterol to HDL cholesterol. There are no central lab reference ranges for the LDL/HDL ratio.

The following will be conducted for the laboratory analytes collected quantitatively:

- Box plots: Values at each visit (starting at first visit in Study JAHN) and change from last baseline to each visit and to last postbaseline measure will be displayed in box plots for patients who have both a baseline and at least 1 postbaseline visit. The last nonmissing observation in the treatment period will be used as the last observation. Box plots will be provided for Treatment Period 1 and for Treatment Periods 1 and 2 combined. Individual measurements outside of reference limits will also be displayed using distinct symbols overlaying the box plot. Original-scale data will be used for the display but for some analytes (eg, immunoglobulins) a logarithmic scale may be used to aid in viewing the measures of central tendency and dispersion. Unplanned measurements will be excluded. Descriptive summary statistics will be included below the box plot along with p-values resulting from between-treatment comparison in change from last baseline to last observation. An ANCOVA model with explanatory term for treatment and the baseline value as a covariate will be used. These box plots will be used to evaluate trends over time and to assess a potential impact of outliers on central tendency summaries.
- Treatment-emergent high/low analyses: The number and percentage of patients with treatment-emergent high and low laboratory results at any time during Treatment Period 1 will be summarized by treatment group. Planned and unplanned measurements will be included. A treatment-emergent **high** result is defined as a change from a value less than or equal to the high limit at all baseline visits to a value greater than the high limit at any time during the treatment period. A treatment-emergent **low** result is defined as a change from a value greater than or equal to the low limit at all baseline visits to a value less than the low limit at any time during the treatment period. The Fisher's exact test will be used for the treatment comparison.

For laboratory analyte measurements collected qualitatively, a listing of abnormal findings will be provided for both treatment periods combined. The listing will include but will not be limited to patient ID, treatment group, laboratory collection date, analyte name, and analyte finding.

6.14.4. Vital Signs and Other Physical Findings

For the treatment-emergent categorical analyses (shift and treatment emergent), 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 analyses (eg, change from baseline by time point) is defined as the treatment period excluding off-drug follow-up time.

Vital signs and physical characteristics include systolic blood pressure, diastolic blood pressure, pulse, weight, and BMI. Original-scale data will be analyzed. When these parameters are analyzed as continuous numerical variables, unplanned measurements will be excluded. When these parameters are analyzed as categorical outcomes and/or treatment-emergent abnormalities, planned and unplanned measurements will be included.

The planned analyses described for the laboratory analytes in Section 6.14.3 will be used to analyze the vital signs and physical characteristics.

Table JAHN.6.8 defines the low and high baseline values, as well as the criteria used to define treatment emergence based on postbaseline values. The blood pressure and pulse rate criteria are consistent with the document *Selected Reference Limits for Pulse/Heart Rate, Arterial Blood Pressure (Including Orthostasis), and Electrocardiogram Numerical Parameters for Use in Analyses of Phase 2-4 Clinical Trials Version 1.3* approved on April 29, 2015 as recommended by the Lilly Cardiovascular Safety Advisory Committee.

Table JAHN.6.8. Categorical Criteria for Abnormal Treatment-Emergent Blood Pressure and Pulse Measurement, and Categorical Criteria for Weight Changes for Adults

Parameter (Units of Measure)	Low	High
Systolic Blood Pressure (mm Hg)	≤90 (low limit) and decrease from lowest value during baseline ≥20 if >90 at each baseline visit	≥140 (high limit) and increase from highest value during baseline ≥20 if <140 at each baseline visit
Diastolic Blood Pressure (mm Hg)	≤50 (low limit) and decrease from lowest value during baseline ≥10 if >50 at each baseline visit	≥90 (high limit) and increase from highest value during baseline ≥10 if <90 at each baseline visit
Pulse (BPM)	<50 (low limit) and decrease from lowest value during baseline ≥15 if ≥50 at each baseline visit	>100 (high limit) and increase from highest value during baseline ≥15 if ≤100 at each baseline visit
Weight (kg)	(Loss) decrease ≥7% from lowest value during baseline	(Gain) increase ≥7% from highest value during baseline

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, JAHL, JAHM, and JAIY with the safety data from Study JAHN. The planned analyses are defined in the compound level safety standards.

6.15. Subgroup Analyses

Subgroup analyses with descriptive statistics will be provided for the responder and partial responder mITT population at Week 16. Subgroup analyses with comparisons for the nonresponder mITT population will be performed at Week 16. Analysis with censoring will be for the following endpoints:

- proportion of patients achieving IGA 0 or 1
- proportion of patients achieving EASI75 Response Rate
- proportion of patients achieving Itch NRS 4-point improvement

The following subgroups categorized into disease-related characteristics and demographic characteristics will be evaluated using baseline of originating study where applicable:

- patient demographic and characteristics subgroups:

  - gender (male, female)
  - age group (<65, ≥65 years old)
  - age group (<65, ≥65 to <75, ≥75 to <85, ≥85 years old)
  - baseline weight (<60 kg, ≥60 to <100 kg, ≥100 kg)
  - baseline BMI (<25 kg/m^2, ≥25 to <30 kg/m^2, ≥30 kg/m^2)
  - race (American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, White, Multiple)
  - baseline renal function status: impaired (eGFR <60 mL/min/1.73 m^2) or not impaired (eGFR ≥60 mL/min/1.73 m^2)
- geographic region subgroups:
  - region: (as defined in Table JAHN.5.1)
  - specific regions (Europe, other)
  - specific country (Japan, other)
- previous and concomitant therapy subgroups:
  - prior use of TCNI (yes, no)
  - prior systemic therapy use (yes, no)
  - cyclosporine contraindication [ineligible] (yes, no)
- baseline disease-related characteristics subgroup
  - baseline disease severity (IGA score)

Descriptive statistics will be provided for each treatment and stratum of a subgroup as outlined, regardless of sample size. If patient and even numbers allow, subgroup analyses for categorical outcomes will be performed using logistic regression using Firth's correction to accommodate (potential) sparse response rates. The model will include the categorical outcome as the dependent variable and baseline value from originating study (for EASI and itch), baseline severity from originating study, treatment, subgroup, and treatment-by-subgroup interaction as explanatory variables. Missing data will be imputed using NRI (Section 6.4.1). The treatment-by-subgroup interaction comparing treatment groups will be tested at the 0.1 significance level. The p-value from the logistic regression model will be reported for the interaction test and the subgroup test, unless the model did not converge. Response counts and percentages will be summarized by treatment for each subgroup category. The difference in percentages and 95% CI of the difference in percentages using the Newcombe-Wilson without continuity correction will be reported. The corresponding p-value from the Fisher's exact test will also be produced.

In case any level of a subgroup comprises <10% of the overall sample size, only descriptive summary statistics will be provided for nonresponder treatment arms, and no treatment group comparisons will be performed within these subgroup levels.

Additional subgroup analyses on efficacy may be performed as deemed appropriate and necessary.

6.16. Protocol Deviations

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

Potential examples of deviations include patients who receive excluded concomitant therapy, significant noncompliance with study medication (<80% of assigned doses taken, failure to take study medication and taking incorrect study medication), patients incorrectly enrolled in the study, and patients whose data are questionable due to significant site quality or compliance issues. Refer to a separate document for the important protocol deviations (IPDs).

The number and percentage of patients having IPD(s) will be summarized within category and subcategory of deviation by treatment group. The summary will be presented for the mITT population. Individual patient listings of IPDs will be provided.

6.17. Interim Analyses and Data Monitoring

A data monitoring committee (DMC) will oversee the conduct of this trial. The DMC will consist of members external to Lilly. This DMC will follow the rules defined in the DMC charter, focusing on potential, and identified risks for this molecule and for this class of compounds. Data Monitoring Committee membership will include, at a minimum, specialists with expertise in dermatology, statistics, and other appropriate specialties.

The DMC will be authorized to review unblinded results of analyses by treatment group prior to DBL, including study discontinuation data, AEs including SAEs, clinical laboratory data, vital sign data, etc. The DMC may recommend continuation of the study, as designed; temporary suspension of enrollment; or the discontinuation of a particular dose regimen or the entire study. While the DMC may request to review efficacy data to investigate the benefit/risk relationship in the context of safety observations for ongoing patients in the study, no information regarding efficacy will be communicated. Moreover, the study will not be stopped for positive efficacy results nor will it be stopped for futility. Hence, no alpha is spent. Details of the DMC, including its operating characteristics, are documented in the Baricitinib Atopic Dermatitis DMC charter and further details are given in the Interim Analysis Plan in Section 6.17.1.

Besides DMC members, a limited number of preidentified individuals may gain access to the limited unblinded data, as specified in the unblinding plan, prior to the interim or final DBL, to initiate the final population PK/PD model development processes or for preparation of regulatory documents. Information that may unblind the study during the analyses will not be reported to study sites or blinded study team until the study has been unblinded.

Unblinding details are given in Section 7.

In addition to the DMC analyses, several interim analyses will be conducted to support regulatory interactions and submission activities (at Week 16 and for the EMA, US and/or Japan submissions, and potentially a 52-week interim CSR Addendum). Refer to the separate Blinding and Unblinding Plan for details.

6.17.1. Interim Analysis Plan

Analyses for the DMC will include listings and/or summaries of the following information:

- patient disposition, demographics, and baseline characteristics
- exposure

- AEs, to include the following:
  - TEAEs
  - SAEs, including deaths
  - selected special safety topics
- clinical laboratory results
- vital signs
- Columbia-Suicide Severity Rating Scale

Summaries will include TEAEs, SAEs, special-topics AEs, and treatment-emergent high and low laboratory and vital signs in terms of counts and percentages where applicable. For continuous analyses, box plots of laboratory analytes will be provided by time point and summaries will include descriptive statistics.

The DMC may request efficacy data if they feel there is value and to confirm a reasonable benefit/risk profile for ongoing patients in the studies. If efficacy data is requested, it will be mean change from baseline of EASI score. Further details are given in the DMC charter.

6.18. Planned Exploratory Analyses

The planned exploratory analyses are described in Sections 6.11 and 6.12. Additional exploratory analyses may be conducted and will be documented in a supplemental SAP, as deemed appropriate. Health Technology Assessment toolkit analyses, which may need to be produced, will also be documented in the supplemental SAP.

6.19. Annual Report Analyses

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

6.20. 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' AE 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 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).
- AE reporting is consistent with other document disclosures (eg, the CSR, manuscripts, and so forth).

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

7. Unblinding Plan

Refer to a separate Blinding and Unblinding Plan document for details.

8. References

Basra MK, Salek MS, Camilleri L, Sturkey R, et al. Determining the minimal clinically important difference and responsiveness of the Dermatology Life Quality Index (DLQI): further data. *Dermatology*. 2015;230(1):27-33. https://doi.org/10.1159/000365390

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. *Arch Dermatol.* 2004;140(12):1513-1519. https://doi.org/10.1001/archderm.140.12.1513

EuroQol Group. EQ-5D-5L User Guide. Version 2.1. Available at: http://www.euroqol.org/fileadmin/user_upload/Documenten/PDF/Folders_Flyers/EQ-5D-5L_UserGuide_2015.pdf. Accessed March 2018.

Hanifin JM, Thurston M, Omoto M, Cherill R, et al. The eczema area and severity index (EASI): assessment of reliability in atopic dermatitis. EASI Evaluator Group. *Exp Dermatol*. 2001;10(1):11-18. https://doi.org/10.1034/j.1600-0625.2001.100102.x

Herdman M, Gudex C, Lloyd A, Janssen MF, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res*. 2011;20(10):1727-1736. https://doi.org/10.1007/s11136-011-9903-x

Hongbo Y, Thomas CL, Harrison MA, Salek MS, et al. 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. https://doi.org/10.1111/j.0022-202x.2005.23621.x

[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 21 July 2017.

[ICH] International Conference on Harmonisation. Addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials. E9(R1) Available at: https://database.ich.org/sites/default/files/E9-R1_Step4_Guideline_2019_1203.pdf. Accessed 20 January 2020.

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, et al. Psychometric properties of the Itch Numeric Rating Scale in patients with moderate-to-severe plaque psoriasis. *Br J Dermatol.* 2016;175(1):157-162. https://doi.org/10.1111/bjd.14464

Kunz B, Oranje A, Labrèze L, Stalder JF, et al. Clinical Validation and Guidelines for the SCORAD Index: Consensus Report of the European Task Force Atopic Dermatitis. *Dermatology*. 1997;195(1):10-19. https://doi.org/10.1159/000245677

Naegeli AN, Flood E, Tucker J, Devlen J, et al. 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. https://doi.org/10.1111/ijd.12645

Protocol I4V-MC-JAHN(a). A Phase 3 Multicenter, Double-Blind Study to Evaluate the Long-Term Safety and Efficacy of Baricitinib in Adult Patients with Atopic Dermatitis. Report on file, Eli Lilly and Company.

Protocol I4V-MC-JAHN(c). A Phase 3 Multicenter, Double-Blind Study to Evaluate the Long-Term Safety and Efficacy of Baricitinib in Adult Patients with Atopic Dermatitis. Report on file, Eli Lilly and Company.

Protocol I4V-MC-JAHN(7.1). A Phase 3 Multicenter, Double-Blind Study to Evaluate the Long-Term Safety and Efficacy of Baricitinib in Adult Patients with Atopic Dermatitis. Report on file, Eli Lilly and Company.

Reilly MC, Zbrozek AS, Dukes EM. The validity and reproducibility of a work productivity and activity impairment instrument. *Pharmacoeconomics*. 1993;4(5):353-365. https://doi.org/10.2165/00019053-199304050-00006

Schram ME, Spuls PI, Leeflang MMG, Lindeboom R, et al. EASI, (objective) SCORAD and POEM for atopic eczema: responsiveness and minimal clinically important difference. *Allergy.* 2012;67(1):99-106. https://doi.org/10.1111/j.1398-9995.2011.02719.x

Snaith RP. The Hospital Anxiety and Depression Scale. *Health Quality Life Outcomes*. 2003;1:29. https://doi.org/10.1186/1477-7525-1-29

Stalder J, Taieb A, Atherton D, Bieber P, et al. Severity scoring of atopic dermatitis: the SCORAD index. Consensus Report of the European Task Force on Atopic Dermatitis. *Dermatology*. 1993;186(1):23-31. https://doi.org/10.1159/000247298

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.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRsearch.cfm?CFRPart=312. Accessed 21 July 2017.

White D, Leach C, Sims R, Atkinson M, et al. Validation of the Hospital Anxiety and Depression Scale for use with adolescents. *Br J Psychiatry*. 1999;175(5):452-454. https://doi.org/10.1192/bjp.175.5.452

Zigmond AS, Snaith RP. The Hospital Anxiety and Depression Scale. *Acta Psychiatr Scand*. 1983;67(6):361-370. https://doi.org/10.1111/j.1600-0447.1983.tb09716.x

Yong Lin, Ilya Lipkovich, Fanni Natanegara, Mary E Nilsson, Patrick M Peterson, Yongming, Statistical Analyses for Clinical Studies Impacted by the Novel Coronavirus (COVID-19) – A White Paper, April 22, 2020, Statistical Analyses for Clinical Studies Impacted by the Novel Coronavirus (COVID-19)

9. Appendices

Appendix 1. Statistical Analysis Plan Addendum

The statistical analysis plan described in this appendix is based on Treatment Period 1 of Protocol I4V-MC-JAHN(7.1) (JAHN[7.1]).

APP.1.1. Study Objectives

APP.1.1.1 Primary Objective

The **primary** objective of this study is to estimate the effect of starting baricitinib 2-mg on clinical measures and patient-reported outcomes as assessed by the proportion of patients with a response of Investigator's Global Assessment (IGA) 0, 1, or 2 at Week 16.

The associated estimand for this objective is to measure the effect of starting baricitinib 2-mg as assessed by the proportion of patients with a response of IGA 0, 1, or 2 at Week 16 assuming that treatment response disappears after patients are rescued or discontinue from study or treatment. See also Section APP.1.2.12.1 on how this estimand handles outcome after the occurrence of any intercurrent event through nonresponder imputation (NRI).

APP.1.1.2. Secondary Objectives

The secondary objectives of this study are:

Objectives	Endpoints
Weeks 0-52	
To estimate the effect of starting baricitinib 2-mg on clinical measures and patient-reported outcomes.	• Proportion of patients with a response of IGA 0, 1, or 2 assessed at Weeks 4, 24, and 52 • Proportion of patients with a response of IGA 0 or 1 assessed at Weeks 4, 16, 24, and 52 • Proportion of patients achieving response of EASI75 from baseline assessed at Weeks 4, 16, 24, and 52 • Proportion of patients with a 4-point improvement from baseline in Itch NRS at Week 16

Abbreviations: AD = atopic dermatitis; EASI = Eczema Area and Severity and Severity Index score; IGA = Investigator's Global Assessment for AD; NRS = Numeric Rating Scale.

APP.1.1.3. Exploratory Objectives

The exploratory objectives of this study are:

Objectives	Endpoints
Weeks 0-52	
To estimate the effect of starting baricitinib 2-mg on clinical measures and patient-reported outcomes.	• Proportion of patients with a 4-point improvement from baseline in Itch NRS at Weeks 1 and 4 • Proportion of patients achieving EASI90 at Weeks 4, 16, 24 and 52 • Percent change from baseline in EASI score at Weeks 4, 16, 24, and 52 • Proportion of patients achieving EASI50 at Weeks 4, 16, 24, and 52 • Proportion of patients achieving SCORAD75 at Weeks 4, 16, 24, and 52 • Mean change from baseline in SCORAD at Weeks 4, 16, 24, and 52 • Proportion of patients achieving SCORAD90 at Weeks 4, 16, 24, and 52 • Mean change from baseline in Skin Pain NRS at Weeks 1, 4, and 16 • Mean change from baseline in the score of the items of the ADSS at Weeks 1, 4, and 16 • Mean change from baseline in BSA affected at Weeks 4, 16, 24, and 52 • Percent change from baseline in Itch NRS at Weeks 1, 4 and 16 • Mean change from baseline in the total score of the POEM at Weeks 4, 16, 24, 52 • Mean change from baseline in the PGI-S-AD scores at Weeks 1, 4, and 16 • Mean change from baseline in the HADS at Weeks 4, 16, 24, and 52 • Mean change from baseline in the DLQI scores at Weeks 4, 16, 24, and 52 • Mean change from baseline in the WPAI-AD components (absenteeism, presenteeism, work productivity loss, and activity impairment) at Weeks 4, 16, 24, and 52 • Mean change from baseline in the EQ-5D-5L scores at Weeks 4, 16, 24, and 52

Abbreviations: ADSS = Atopic Dermatitis Sleep Scale; BSA = body surface area; DLQI = Dermatology Life Quality Index; EASI = Eczema Area and Severity Index; EQ-5D-5L = European Quality of Life–5 Dimensions–5 Levels; HADS = Hospital Anxiety Depression Scale; NRS = Numeric Rating Scale; PGI-S-AD = Patient Global Impression of Severity–Atopic Dermatitis; POEM = Patient-Oriented Eczema Measure; SCORAD = SCORing Atopic Dermatitis; WPAI-AD = Work Productivity and Activity Impairment-Atopic Dermatitis.

APP.1.2. Study Design

APP.1.2.1 Summary of Study Design

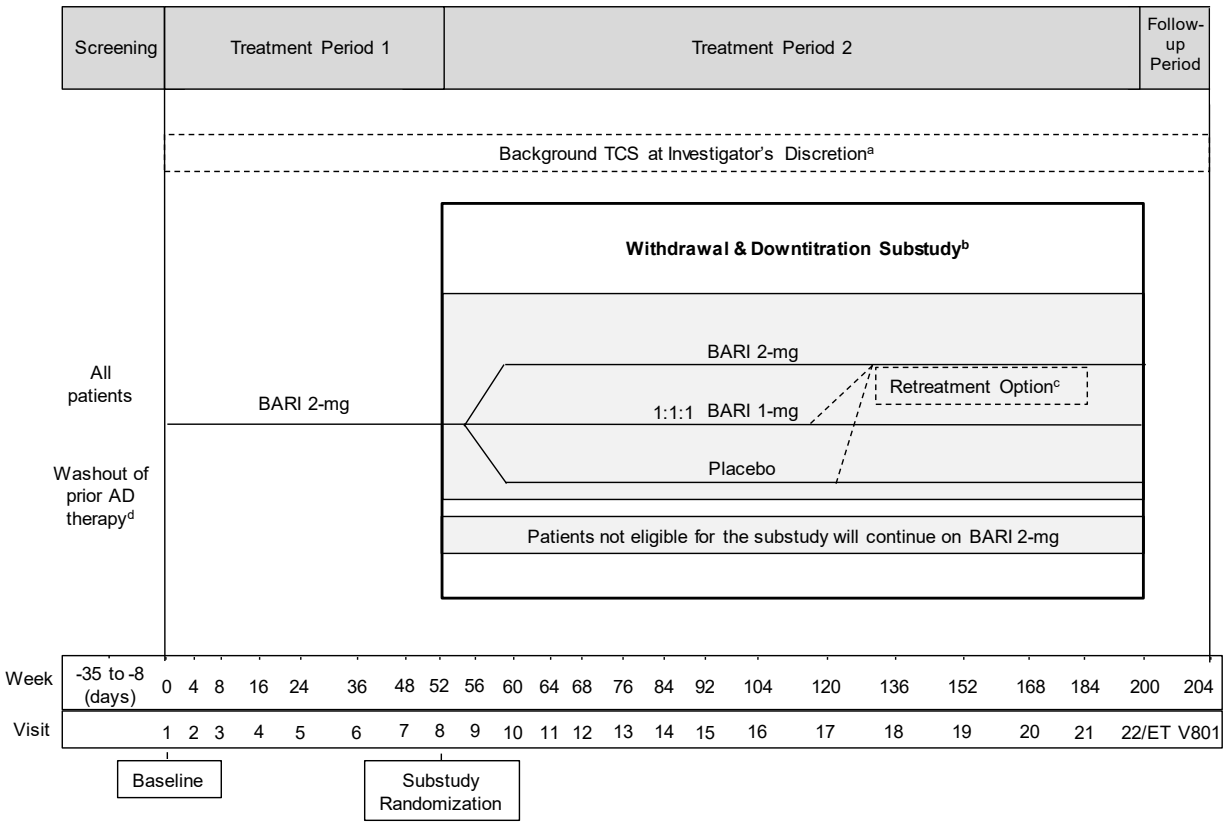
Study JAHN(7.1) is a Phase 3, multicenter, open-label study to evaluate the effects of starting baricitinib 2-mg in patients with atopic dermatitis (AD). The study will consist of 4 study periods and 1 substudy: randomized treatment withdrawal and downtitration. Unlike the main protocol, patients do not have to complete an originating study.

Figure APP.1.1 illustrates the study design. The full visit schedule is outlined in the Study Schedule of Activities in Protocol JAHN(7.1). After the screening period, all safety and efficacy assessments will be collected at identical time points as in the main Study JAHN protocol.

Patients can participate in Study JAHN(7.1) if all the criteria are met during the screening period (up to 5 weeks in duration). All patients enrolled in Study JAHN(7.1) will receive open-label baricitinib 2-mg for 52 weeks (Treatment Period 1). During Treatment Period 2 (Weeks 52 to 200), patients may participate in the randomized withdrawal and downtitration substudy if all eligibility criteria are met, otherwise patients will continue to be treated with open-label baricitinib 2-mg. Patients will use emollients daily throughout the study.

Study JAHN(7) will consist of 4 periods:

- Screening Period: Duration is between 8 and 35 days prior to Visit 1 (Week 0).
- Treatment Period 1: Treatment period from Week 0 to Week 52. Patients who meet all study eligibility criteria at Visit 1 (Week 0, baseline) will proceed to enrollment and begin the 52-week open-label treatment period.
- Treatment Period 2: Treatment period from Week 52 to 200 and will contain a randomized withdrawal and downtitration substudy. At Week 52, all patients will be evaluated for substudy eligibility and the criteria are given in the Protocol. Patients not entered in the substudy will continue on their current treatment.
- Posttreatment follow-up period: Period from Week 200 (Visit 22) or Early Termination Visit (ETV) to approximately 28 days after the last dose of investigational product.



Abbreviations: AD = atopic dermatitis; Bari = baricitinib; ET = early termination; IGA = Investigator's Global Assessment; TCS = topical corticosteroids; V = visit; W = week.

a Background TCS may be initiated or reinitiated at any time during the study, however, we recommend waiting at least 1 week after initial treatment assignment since everyone will be on active, open-label treatment. For those eligible for the substudy, TCS will also be provided as part of retreatment any time a patient's IGA score becomes ≥3. Please refer to the main protocol.

b Eligible patients will be randomized in the withdrawal and downtitration substudy as described in the main protocol. Patients who do not enroll in the substudy will remain on their treatment as described in the main protocol.

c Patients enrolled in the substudy will automatically be retreated if their IGA score becomes ≥3 as described in the main protocol.

d Applicable to patients taking topical treatments (excluding emollients) or systemic treatments for AD at the time of screening.

Figure APP.1.1. Illustration of study design for Clinical Protocol I4V-MC-JAHN.

APP.1.2.2. Method of Assignment to Treatment

At entry into Study JAHN(7.1), patients who meet all criteria for enrollment will be assigned treatment at Visit 1 (Week 0). The interactive web-response system will assign open-label bottles starting at Visit 1 (Week 0), and at each visit up to and including Visit 7 (Week 48) for all patients.

At Week 52, patients eligible for the withdrawal and downtitration substudy will be assigned to treatment as described in the main protocol.

Patients who do not qualify for the substudy will continue to be assigned open-label bottles from Visit 8 (Week 52) until Visit 15 (Week 92).

This study will be conducted internationally at multiple sites. Table APP.1.1 describes how regions were defined.

Table APP.1.1. Geographic Regions

Region	Country
Europe	Austria, Czech Republic, Denmark, Germany, Hungary, Italy, Spain, Switzerland
Rest of World	Argentina, Australia, Israel, Mexico, Russia

APP.1.2.3. A Priori Statistical Methods

APP.1.2.3.1. Determination of Sample Size

Approximately 250 patients may be enrolled in Study JAHN(7.1). Given that this addendum was planned, these patients are included in the anticipated number of participants enrolling into the randomized withdrawal and downtitration substudy as discussed in Section 6.1.

This study is intended to estimate the effect of starting baricitinib 2-mg and thus, the analyses will be descriptive.

Patients at Week 52 will be stratified by responder status (IGA 0 or 1 versus IGA 2) when entering the randomized withdrawal and downtitration substudy. It is estimated that there will be approximately 60 patients entering into the randomized withdrawal and downtitration substudy from JAHN(7.1). The substudy is meant to evaluate the change in clinical response after treatment withdrawal or downtitration and does not account for whether the sample size is sufficient to detect a difference between baricitinib and placebo. Maintenance of treatment benefit is defined as response of IGA 0, 1, or 2.

APP.1.2.3.2. General Considerations

Refer to Section 6.2.

APP.1.2.3.2.1. Analysis Populations

Intent-to-treat (ITT) population: The ITT population analysis set is defined as all enrolled patients in Study JAHN(7.1).

Unless otherwise specified, the efficacy and health outcome analyses will be conducted on the ITT population (Gillings and Koch 1991).

Safety population: The Safety population for this addendum is defined the same as the safety population in Section 6.2.1.

Safety analyses for this addendum will be performed on the safety population and data will be pooled with the data from Study JAHN.

For the database lock that supports the US submission, the population for this addendum, Week 24 efficacy and health outcome analyses will be conducted on the **Week 24 efficacy evaluable se**t. This subset of the ITT Population is anchored on the database cut-off date for the US DBL (13 December 2019). Specifically, a patient will be included in the Week *24* Efficacy Evaluable set if their Week *24* visit has occurred or their expected Week *24* visit date plus a 15-day buffer is on or prior to the database cutoff date. The expected Week *24* visit date will be calculated as follows: Date of first dose date + (*24* weeks * 7 days) + 15 days.

APP.1.2.3.2.2. Definition of Baseline and Postbaseline Measures

The baseline value for the efficacy and health outcome analyses is defined similarly to Section 6.2.2 except it uses the date of first study drug administration (expected at Week 0, Visit 1) in place of the first study drug administration in the originating study. The baseline value for the daily diary assessments (Itch Numeric Rating Scale (NRS), Atopic Dermatitis Sleep Scale, Skin Pain NRS, Patient Global Impression of Severity-AD) is also defined similarly (using the date of first study administration in Study JAHN).

For the purpose of Study JAHN addendum, the safety analyses will be conducted as described in Section 6.2.2.

Postbaseline measurements are described in Section 6.2.2.

APP.1.2.3.2.3. Analysis Methods

All discrete efficacy and health outcome variables will be summarized using frequencies and percentages. The 95% confidence intervals will also be provided.

APP.1.2.3.2.4. Derived Data

Refer to Section 6.2.4, using the Study JAHN(7.1) baseline where applicable.

APP.1.2.4. Adjustments for Covariates

Not applicable.

APP.1.2.5. Handling of Dropouts or Missing Data

Refer to Section 6.4.

The censoring rule will be applied to all efficacy and health outcome endpoints conducted for the ITT population.

Sections 6.4.1 and 6.4.2 summarize the imputation methods for the various efficacy and health outcome endpoints. Table APP.1.2 summarizes the various imputation techniques being used for the efficacy and health outcomes analyses.

Table APP.1.2. Imputation Techniques for Various Variables

Efficacy and Health Outcome Endpoints[a]	Imputation Method[b]
IGA(0,1), IGA(0,1,2), EASI75, 4-point Itch NRS improvement	NRI
All remaining categorical endpoints	NRI

Abbreviations: AD = atopic dermatitis; EASI = Eczema Area and Severity Index score; IGA = Investigator's Global Assessment for AD; NRI = nonresponder imputation, NRS = Numeric Rating Scale.

[a] Refer to Table JAHN.6.4 and Table JAHN.6.6.

[b] Analysis utilizing the censoring rule.

APP.1.2.6. Multicenter Studies

This study will be conducted by multiple investigators at multiple sites internationally. The countries will be categorized into geographic regions, as described in Section APP.1.2.2.

APP.1.2.7. Multiple Comparisons/Multiplicity

As this study is designed to assess the effect of starting baricitinib 2-mg in patients with AD, no adjustments for multiple comparisons will be utilized in the statistical analyses for this study.

APP.1.2.8. Patient Disposition

Patient disposition will be described as discussed in Section 6.7 for the Study JAHN(7.1) ITT population only. A listing of the ITT population with their treatment assignments will not be produced as all patients are on open-label baricitinib 2-mg.

APP.1.2.9. Patient Characteristics

Patient characteristics including demographics, baseline characteristics, and pre-existing conditions will be summarized descriptively.

APP.1.2.9.1. Demographics

Patient demographics will be summarized as described above. Refer to Section 6.8.1 for the list of demographics. A listing of patient demographics will also be provided for the ITT population.

APP.1.2.9.2. Baseline Disease Characteristics

Refer to Section 6.8.2 for the baseline AD clinical characteristics, baseline health outcome measures, and other baseline demographic and disease characteristics which will be presented

using the Study JAHN(7.1) baseline where applicable. In addition, the following will be summarized:

- 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)
- Prior use of Cyclosporine (yes, no)
- Reasons Cyclosporine was not used
- Prior use of topical calcineurin inhibitor (TCNI) (yes, no)
- Reasons TCNI was not used

APP.1.2.9.3. Historical Illness and Pre-existing Conditions

Historical illnesses are defined as those conditions recorded in the Pre-existing Conditions and Medical History electronic case report form (eCRF) or from the Prespecified Medical History: Comorbidities eCRF with an end date prior to the informed consent date. The number and percentage of patients with selected historical diagnoses will be summarized by treatment group using the ITT population. Historical diagnoses will be categorized using the Medical Dictionary for Regulatory Activities (MedDRA®, most current available version) algorithmic standardized MedDRA queries or similar pre-defined lists of preferred terms of interest.

Pre-existing conditions are defined in Section 6.8.3 and will be summarized using the ITT population.

APP.1.2.10. Treatment Compliance

Patient compliance with study medication will be assessed from Week 0 (Visit 2) to Week 104 (Visit 16) or ETV using the ITT population.

All patients are expected to take 1 tablet daily from a bottle for Treatment Period 1 as described in APP.1.2.2. Each bottle contains 36 tablets.

Refer to Section 6.9 for further details on compliance. Analyses will be presented for the ITT population.

APP.1.2.10.1. Background Therapy

Analyses will be presented as described in Section 6.9.1 using the ITT population.

APP.1.2.11. Previous and Concomitant Therapy

Summaries of previous and concomitant medications will be based on the ITT population.

At screening, previous and current AD treatments are recorded for each patient. A summary of previous medications used for AD, including zoster immunization and tuberculosis vaccine and medications that are discontinued after screening and before the first dose of study drug, will be prepared using frequency counts and percentages by preferred medication name, with preferred medication names sorted by frequency. Concomitant therapy will be recorded at each visit and will be classified similarly. An additional summary for previous medications used for AD will be created containing the reason of discontinuation.

Concomitant therapy for the treatment period is defined as therapy that starts before or during the treatment period and ends during the treatment period or is ongoing (has no end date or ends after the treatment period). Should there be insufficient data to make this comparison (for example, 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.

Summaries of previous medications will be provided for the following categories:

- Previous AD therapies
- Previous AD therapies including reason for discontinuation

A summary of concomitant therapy including concomitant therapies of special interest will be provided for the ITT population.

APP.1.2.12. Efficacy Analyses

The general methods used to summarize efficacy data, including the definition of baseline value for assessments are described in Section APP.1.2.3.2.

Efficacy analyses will generally be analyzed according to the following formats:

- Week 0 to Week 52

Table JAHN.6.4 includes the descriptions and derivations of the primary, secondary, and exploratory efficacy outcomes. Table APP.1.3 provides the detailed analyses including analysis type, method and imputation, population, and time point for efficacy analyses.

Table APP.1.3. Description of Primary, Secondary and Exploratory Efficacy Analyses

Measure	Variable	Analysis Method (Section 6.2.3)	Population (Section 6.2.1)	Time Point	Analysis Type
Validated Investigator's Global Assessment for AD (IGA)	▪ Proportion of patients with a response of IGA [0,1,2]	Descriptive statistics (Observed and NRI)	ITT	Weeks 0 to 16	Secondary analysis
			ITT Week 24 EE set	Week 24	Exploratory analysis
	▪ Proportion of patients with a response of IGA [0,1]	Descriptive statistics (Observed and NRI)	ITT	Weeks 0 to 16	Primary analysis
			ITT Week 24 EE set	Week 24	Exploratory analysis
Eczema Area and Severity Index (EASI)	▪ EASI score ▪ Change from baseline in EASI score ▪ Percent change from baseline in EASI score	Descriptive statistics (Observed)	ITT	Weeks 0 to 24	Exploratory analysis
	▪ Proportion of patients achieving EASI75	Descriptive statistics (Observed and NRI)	ITT	Weeks 0 to 16	Secondary analysis
			ITT Week 24 EE set	Week 24	Exploratory analysis
	▪ Proportion of patients achieving EASI90 ▪ Proportion of patients achieving EASI50	Descriptive statistics (Observed and NRI)	ITT	Weeks 0 to 16	Exploratory analysis
			ITT Week 24 EE set	Week 24	Exploratory analysis
Body Surface Area (BSA) Affected by AD	▪ BSA score ▪ Change from baseline in BSA score	Descriptive statistics (Observed)	ITT	Weeks 0 to 24	Exploratory analysis
SCORing Atopic Dermatitis (SCORAD)	▪ SCORAD score ▪ Change from baseline in SCORAD score ▪ Percent change from baseline in SCORAD score	Descriptive statistics (Observed)	ITT	Weeks 0 to 24	Exploratory analysis
	▪ Proportion of patients achieving SCORAD75 ▪ Proportion of patients achieving SCORAD90	Descriptive statistics (Observed and NRI)	ITT	Weeks 0 to 16	Exploratory analysis
			ITT Week 24 EE set	Week 24	Exploratory analysis

Abbreviations: AD = atopic dermatitis; EE = efficacy evaluable; ITT = intent-to-treat; MMRM = mixed-model repeated measures; NRI = nonresponder imputation.

APP.1.2.12.1. Primary Outcome and Methodology

The validated IGA for AD uses the clinical characteristics of erythema, papulation/induration, oozing/crusting and lichenification to produce a single-item score ranging from 0 to 4. The primary analysis of the study is to estimate the effect of starting baricitinib 2-mg on clinical measures and patient-reported outcomes as assessed by the proportion of patients with a response of IGA 0, 1, or 2 assessed at Week 16 using the ITT population assuming that treatment response disappears after the patient is rescued or discontinued study or study treatment. This will serve as the primary estimand. In this estimand, missing data due to the application of the censoring rule and the occurrence of other noncensor intercurrent events will be imputed using the NRI method described in Section 6.4.1.

APP.1.2.12.2. Secondary and Exploratory Efficacy Analyses

There will be no adjustment for multiple comparisons for any analyses. The secondary and exploratory efficacy analyses are detailed in Section 6.2.3 and Table JAHN.6.4. Health outcomes analyses are described in Section APP.1.2.13.

Note that the exploratory efficacy analyses on a continuous outcome that is measured over time uses a slightly different estimand which is to estimate the effect of starting baricitinib 2-mg on a specific endpoint evaluated at specified time points using the ITT population assuming all patients remained in their treatment throughout the defined period of the study.

APP.1.2.13. 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 APP.1.2.3.2.2.

Health outcomes and quality-of-life measures will generally be analyzed according to the formats discussed in Section APP.1.2.12.

Table JAHN.6.6 includes the descriptions and derivations of the health outcomes and quality-of-life measures. Table APP.1.4 provides the detailed analyses including analysis type, method and imputation, population, and time point for health outcomes and quality-of-life measures.

Table APP.1.4. Description of Health Outcomes and Quality-of-Life Measures Analyses

Measure	Variable	Analysis Method (Section 6.2.3)	Population (Section 6.2.1)	Time Point	Analysis Type
Itch Numeric Rating Scale (NRS)	▪ Itch NRS score ▪ Change from baseline in Itch NRS score ▪ Percent change from baseline Itch score	Descriptive statistics	ITT	Weeks 0 to 16	Exploratory analysis
	▪ Proportion of patients with a 4-point improvement from baseline of originating study in Itch NRS in subgroup of patients who had baseline Itch NRS ≥4	Descriptive statistics (Observed and NRI)	ITT	Weeks 0 to 16	Secondary analysis
Skin Pain Numeric Rating Scale (NRS)	▪ Skin Pain NRS score ▪ Change from baseline in Skin Pain NRS score	Descriptive statistics (Observed)	ITT	Weeks 0 to 16	Exploratory analysis
	▪ Proportion of patients with a 4-point improvement from baseline of originating study in Skin Pain NRS in subgroup of patients who had baseline Skin Pain NRS ≥4	Descriptive statistics (Observed and NRI)	ITT	Weeks 0 to 16	Exploratory analysis
Atopic Dermatitis Sleep Scale (ADSS)	▪ ADSS item scores ▪ Change from baseline in ADSS item scores	Descriptive statistics (Observed)	ITT	Weeks 0 to 16	Exploratory analysis
	▪ Proportion of patients with a 1.5 point improvement from baseline of originating study in ADSS2 score in subgroup of patients who had baseline ADSS2 score ≥ 1.5 ▪ Proportion of patients with a 2 point improvement from baseline of originating study in ADSS2 score in subgroup of patients who had baseline ADSS2 score ≥ 2	Descriptive statistics (Observed and NRI)	ITT	Weeks 0 to 16	Exploratory analysis

Measure	Variable	Analysis Method (Section 6.2.3)	Population (Section 6.2.1)	Time Point	Analysis Type
Patient Global Impression of Severity–Atopic Dermatitis (PGI-S-AD)	▪ PGI-S-AD score ▪ Change from baseline in PGI-S-AD score	Descriptive statistics (Observed)	ITT	Weeks 0 to 16	Exploratory analysis
Patient-Oriented Eczema Measure (POEM)	▪ POEM score ▪ Change from baseline in POEM score	Descriptive statistics (Observed)	ITT	Weeks 0 to 24	Exploratory analysis
	▪ Proportion of patients with a 4 point improvement from baseline of originating study in POEM score in subgroup of patients who had baseline POEM score ≥4	Descriptive statistics (Observed and NRI)	ITT	Weeks 0 to 16	Exploratory analysis
			ITT Week 24 EE set	Week 24	Exploratory analysis
Hospital Anxiety Depression Scale (HADS)	▪ HADS domain scores and total scores ▪ Change from baseline in HADS domain and total scores	Descriptive statistics (Observed)	ITT	Weeks 0 to 24	Exploratory analysis
Dermatology Life Quality Index (DLQI)	▪ DLQI total score ▪ Change from baseline in DLQI ▪ Observed and change from baseline in domain scores - Symptoms and feelings - Daily activities - Leisure - Work and school - Personal relationships - Treatment	Descriptive statistics (Observed)	ITT	Weeks 0 to 24	Exploratory analysis
	▪ Proportion of patients with a DLQI (0, 1) response ▪ Proportion of patients with a 4 point improvement from baseline of originating study in DLQI in subgroup of patients who had baseline DLQI ≥4	Descriptive statistics (Observed and NRI)	ITT	Weeks 0 to 16	Exploratory analysis
			ITT Week 24 EE set	Weeks 24	Exploratory analysis

Measure	Variable	Analysis Method (Section 6.2.3)	Population (Section 6.2.1)	Time Point	Analysis Type
Work Productivity and Activity Impairment: Atopic Dermatitis (WPAI-AD)	Observed and Change from baseline in employment status	Descriptive statistics (observed)	ITT	Weeks 0 to 24	Exploratory analysis
	Observed and Change from baseline in: ▪ absenteeism ▪ presenteeism ▪ overall work impairment ▪ impairment in activities	Descriptive statistics (Observed)	ITT	Weeks 0 to 24	Exploratory analysis
European Quality of Life–5 Dimensions–5 Levels (EQ-5D-5L)	Observed values in ▪ EQ-5D mobility ▪ EQ-5D self-care ▪ EQ-5D usual activities ▪ EQ-5D pain/ discomfort ▪ EQ-5D anxiety/ depression	Descriptive statistics (Observed)	ITT	Weeks 0 to 24	Exploratory analysis
	Observed and Change from baseline in ▪ EQ-5D VAS ▪ EQ-5D-5L UK Population-based index score ▪ EQ-5D-5L US Population-based index score	Descriptive statistics (Observed)	ITT	Weeks 0 to 24	Exploratory analysis

Abbreviations: AD = atopic dermatitis; EE = efficacy evaluable; ITT = intent-to-treat; NRI = nonresponder imputation.

APP.1.2.14. Bioanalytical and Pharmacokinetic / Pharmacodynamic Methods

Pharmacokinetic, pharmacodynamic, and biomarker analyses will not be conducted.

APP.1.2.15. Safety Analyses

The safety data from this addendum will be pooled with the bari 2-mg treatment group from Study JAHN and will follow the same planned analyses given in Section 6.14. Refer to Section 6.14 for the planned analyses.

APP.1.2.15.1. Extent of Exposure

Refer to Section 6.14.1 for the planned analyses.

APP.1.2.15.2. Adverse Events

Refer to Section 6.14.2 for the planned analyses.

APP.1.2.15.2.1. Common Adverse Events

Not applicable.

APP.1.2.15.2.2. Serious Adverse Event Analyses

Refer to Section 6.14.2.2 for the planned analyses.

APP.1.2.15.2.3. Other Significant Adverse Events

Refer to Section 6.14.2.3 for the planned analyses.

APP.1.2.15.3. Clinical Laboratory Evaluation

Refer to Section 6.14.3 for the planned analyses.

APP.1.2.15.4. Vital Signs and Other Physical Findings

Refer to Section 6.14.3 for the planned analyses.

APP.1.2.15.5. Special Safety Topics, including Adverse Events of Special Interest

Refer to Section 6.14.5 for the planned analyses.

APP.1.2.16. Subgroup Analyses

Subgroup analyses will not be performed for Study JAHN(7.1).

APP.1.2.17. Protocol Deviations

Protocol deviations will be analyzed as described in Section 6.16 using the ITT population.

APP.1.2.18. Interim Analyses and Data Monitoring

Refer to Section 6.17 for the planned analyses.

APP.1.2.19. Planned Exploratory Analyses

The planned exploratory analyses are described in Section 6.18.

APP.1.2.20. Annual Report Analyses

Refer to Section 6.19.

APP.1.2.21. Clinical Trial Registry Analyses

Refer to Section 6.20.

APP.1.3. Unblinding Plan

Refer to Section 7.

APP.1.4. References

Gillings D, Koch G. The Application of the Principle of Intention–to–Treat to the Analysis of Clinical Trials. *Drug Inf J*. 1991;25:411-424.

Protocol I4V-MC-JAHN(7.1). A Phase 3 Multicenter, Double-Blind Study to Evaluate the Long-Term Safety and Efficacy of Baricitinib in Adult Patients with Atopic Dermatitis. Report on file, Eli Lilly and Company.

Appendix 2. Statistical Analysis for 4-Month Safety Update

This appendix describes the efficacy analyses to be performed for the 4-month safety update (20 May 2020).

APP.2.1. *4-Month Safety Update Lock*

For the DBL that supports the 4-month safety update, the population originating from combination therapy Study JAIY has a definition for Week 36 efficacy evaluable set, similar to Week 24 efficacy evaluable set defined in Section 6.2.1. Week 36 efficacy and health outcome analyses will be conducted on the **Week 36 efficacy evaluable set**. The following efficacy summaries/analyses will be provided:

- Proportion of patients achieving IGA (0,1) and IGA (0,1,2) assessed at Weeks 0 to 24 for nonresponders (NR). Analysis to be repeated for responders and partial responders (RPR).
- Proportion of patients achieving IGA (0,1) and IGA (0,1,2) assessed at Week 36 efficacy evaluable set for NR. Analysis to be repeated for RPR.
- Proportion of patients achieving EASI50, EASI75, and EASI90 response assessed at Weeks 0 to 24 for NR. Analysis to be repeated for RPR.
- EASI change from baseline and percent change from baseline mixed-model repeated measures assessed at Weeks 0 to 24 for NR.
- Proportion of patients achieving EASI50, EASI75, and EASI90 response assessed at Week 36 efficacy evaluable set for NR. Analysis to be repeated for PRP.
- EASI change and percent change from baseline assessed at Week 36 efficacy evaluable set for NR. Analyses to be repeated for RPR.
- Summary of sponsor provided topical corticosteroids (TCS) background therapy weight (g) by potency for weeks 0 to 52.

For the 4-month safety update database lock, the open label addendum has a definition for **Week 36 and 48 efficacy evaluable sets**, similar to Week 24 efficacy evaluable set defined in App.1.2.3.2.1. Week 36 and 48 efficacy and health outcome analyses will be conducted on the **Week 36 and 48 efficacy evaluable sets**. The following efficacy summaries/analyses will be provided:

- Proportion of patients achieving IGA (0,1) and IGA (0,1,2) assessed at Weeks 0 to 24.
- Proportion of patients achieving IGA (0,1) and IGA (0,1,2) assessed at Week 36 and Week 48 efficacy evaluable sets, respectively.
- Proportion of patients achieving EASI50, EASI75, and EASI90 response assessed at Weeks 0 to 24.
- Proportion of patients achieving EASI50, EASI75, and EASI90 response rate assessed at Week 36 and Week 48 efficacy evaluable sets, respectively.
- EASI change and percent change from baseline assessed at Weeks 0 to 48 based on observed data.

- Summary of sponsor provided TCS background therapy weight (g) by potency for weeks 0 to 52.

For the substudy the following analyses/summaries will be provided:

- Proportion of patients achieving IGA (0,1) and IGA (0,1,2) assessed at Weeks 52 to 68.
- Proportion of patients achieving EASI50, EASI75, and EASI90 assessed at Weeks 52 to 68.
- EASI change from baseline and percent change from baseline (mLOCF) assessed at Weeks 52 to 68; Week 0 used as baseline.
- EASI change from baseline and percent change from baseline (mLOCF) assessed at Weeks 52 to 68; Week 52 used as baseline.
- Proportions of patients achieving a ≥4-point improvement assessed at Weeks 52 to 68.
- Itch NRS change from baseline and percent change from baseline (mLOCF) assessed at Weeks 52 to 68; Week 0 used as baseline.
- Itch NRS change from baseline and percent change from baseline (mLOCF) assessed at Weeks 52 to 68; Week 52 used as baseline.

Signature Page for VV-CLIN-127789 v1.0

Approval	PPD 25-Aug-2023 13:26:21 GMT+0000

Signature Page for VV-CLIN-127789 v1.0