Objectives/Endpoints:

Objectives	Endpoints
Primary *This is a prespecified objective that will be adjusted for multiplicity.*	
To test the hypothesis that baricitinib 2-mg QD is superior to placebo in the treatment of patients with moderate to severe AD	• Proportion of patients achieving EASI75 at Week 16
Key Secondary *These are prespecified objectives that will be adjusted for multiplicity.*	
To compare the efficacy of baricitinib 1-mg QD or 2-mg QD to placebo in AD during the 16-week, double-blind, placebo-controlled treatment period as measured by improvement in signs and symptoms of AD	• Proportion of patients achieving IGA of 0 or 1 with a ≥2-point improvement at Week 16 • Proportion of patients achieving EASI75 at Week 16 (1-mg) • Proportion of patients achieving EASI90 at Week 16 • Mean percent change from baseline in EASI score at Week 16 • Proportion of patients achieving SCORAD75 at Week 16
To compare the efficacy of baricitinib 1-mg QD or 2-mg QD to placebo in AD during the 16-week, double-blind, placebo-controlled treatment period as assessed by patient-reported outcome measures	• Proportions of patients achieving a 4-point improvement in Itch NRS at 1 week, 2 weeks, 4 weeks, and 16 weeks • Mean change from baseline in the score of Item 2 of the ADSS at 1 week and 16 weeks • Mean change from baseline in Skin Pain NRS at Week 16
Other Secondary Objectives *These are prespecified objectives that will not be adjusted for multiplicity.*	
To compare the efficacy of baricitinib 1-mg QD or 2-mg QD to placebo in AD during the 16-week, double-blind, placebo-controlled period as measured by improvement in signs and symptoms of AD	• Proportion of patients achieving IGA of 0 or 1 with a ≥2-point improvement at Week 4 • Proportion of patients achieving EASI50 at Week 16 • Proportion of patients achieving IGA of 0 at Week 16 • Mean change from baseline in SCORAD at Week 16 • Proportion of patients achieving SCORAD90 at Week 16 • Mean change from baseline in body surface area affected at Week 16 • Proportion of patients developing skin infections requiring antibiotic treatment by Week 16

	Period 1: Screening	Period 2: Double-Blinded Treatment Period														Period 3: PTFU
Visit Number	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15/ET[a]**	**801[b]**
Weeks from Randomization		**0**	**1**	**2**	**4**	**8**	**12**	**16**	**28**	**40**	**52**	**64**	**76**	**88**	**104**	**108**
Visit tolerance interval (days)	**-8 to -35**	**0**	**±2**	**±2**	**±2**	**±4**	**±4**	**±4**	**±7**	**±7**	**±7**	**±7**	**±7**	**±7**	**±4**	**28±4 after last dose**
IP returned and compliance assessed			X[h]	X[h]	X	X	X	X	X	X	X	X	X	X	X	
Dispense TCS[i]			X	X	X	X	X	X	X	X	X	X	X	X	X	
Weigh (tube with cap) and record returned TCS[i] (as needed)				X	X	X	X	X	X	X	X	X	X	X	X	X
Scales																
IGA	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
EASI	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
SCORAD	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Health Outcomes Measures and Other Questionnaires[j]																
Itch NRS	X	X	X	X	X	X	X	X[g]								
Skin Pain NRS	X	X	X	X	X	X	X	X[g]								
ADSS	X	X	X	X	X	X	X	X[g]								
PGI-S-AD	X	X	X	X	X	X	X	X[g]								
POEM	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
DLQI	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X
HADS	X	X		X	X	X		X	X	X	X	X	X	X	X	X
EQ-5D-5L		X	X	X	X	X	X	X		X		X		X	X	X
WPAI-AD		X	X	X	X	X	X	X		X		X		X	X	X
C-SSRS and Self-Harm Supplement[k]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Self-Harm Follow-Up Form[l]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Laboratory Assessments																
Lipids (fasting)[m]		X					X		X		X		X		X[n]	X
Clinical chemistry[o]	X	X			X	X	X	X	X	X	X	X	X	X	X	X
Hematology	X	X			X	X	X	X	X	X	X	X	X	X	X	X
Serum pregnancy[p]	X															
FSH[q]	X															
TSH	X															
HIV	X															
HCV antibody[r]	X															
HBV testing	X															
HBV DNA[s]	X							X	X	X	X	X	X	X	X	X

Objectives	Endpoints
To compare the efficacy of baricitinib 1-mg QD or 2-mg QD to placebo in AD during the 16-week, double-blind, placebo-controlled treatment period as assessed by patient-reported outcome/QoL measures	• Mean percent change from baseline in Itch NRS at 1 week and 16 weeks • Mean change from baseline in Itch NRS at 4 weeks and 16 weeks • Mean change from baseline in the total score of the POEM at Week 16 • Mean change in PGI-S-AD scores at Week 16 • Mean change from baseline in the HADS at Week 16 • Mean change in DLQI scores at Week 16 • Mean change in WPAI-AD scores at Week 16 • Mean change in EQ-5D-5L scores at Week 16
Other Secondary Objectives for responders beyond Week 16	
To describe the long-term efficacy of baricitinib 1-mg QD or 2-mg QD in AD as measured by improvement in signs and symptoms of AD	• Proportion of patients with a response of IGA 0 or 1 at Week 16 who maintain an IGA 0 or 1 at Weeks 28, 52, and 104 • Proportion of patients with a response of IGA 0 or 1 at Week 16 who achieve EASI75 assessed at Weeks 28, 52, and 104 • Proportion of patients with a response of IGA 0 or 1 at Week 16 who achieve SCORAD75 at Weeks 28, 52, and 104 • Mean percent change from baseline in EASI score at Weeks 28, 52, and 104 • Mean percent change from baseline in SCORAD score at Weeks 28, 52, and 104 • Mean percent change from baseline in SCORAD pruritus at Weeks 28, 52, and 104 • Mean percent change from baseline in POEM at Weeks 28, 52, and 104
Exploratory Objectives/Endpoints	
• Frequency of patient-reported “no itch” (Itch NRS score = 0) days from daily diaries from Week 12 to Week 16 • Frequency of patient-reported “no pain” (Skin Pain NRS score = 0) days from daily diaries from Week 12 to Week 16 • Mean change from baseline in the score of Item 1 of the ADSS at 1 week and 16 weeks • Mean change from baseline in the score of Item 3 of the ADSS at 1 week and 16 weeks • Exploratory objectives evaluating the response to baricitinib treatment regimens on other patient reported outcomes will be specified in the SAP. These endpoints may include dichotomous endpoints or change from baseline for the following measures: POEM, DLQI, Itch NRS, ADSS, Skin Pain NRS, PGI-S-AD, HADS. • To evaluate changes from baseline in IgE levels during the study • To evaluate changes from baseline in eosinophil levels during the study	

Patients who meet all of the inclusion and none of the exclusion criteria (Section 6.2) will continue to Visit 2.

5.1.2. Period 2: Double-Blind Placebo-Controlled Treatment, Weeks 0 to 104

At Visit 2 (Week 0, baseline), study eligibility for each patient will be reviewed, on the basis of all inclusion (Section 6.1) and exclusion criteria (Section 6.2) and laboratory test results. Patients who meet all criteria will proceed to randomization and begin the double-blind, placebo-controlled treatment period.

At Visit 2, after laboratory samples are collected and all assessments are completed, patients will take the first dose of investigational product.

Patients will be randomized at a 1:1:1 ratio into 1 of the 3 treatment groups (placebo QD, baricitinib 1–mg QD, or baricitinib 2–mg QD). Investigational product will be administered QD for 104 weeks (treatment period, Visits 2 through 15; see Section 7). All patients will be required to use emollients daily. Daily diaries will continue to be utilized through Week 16. Download of this data will be required at study visits. The use of topical corticosteroids (TCS), topical calcineurin inhibitors (TCNIs), topical phosphodiesterase-4 (PDE-4) inhibitor (crisaborole), and systemic therapies for the treatment of AD are not allowed during the first 16 weeks of the study, except as part of rescue therapy for patients not responding to treatment. After Week 16 (primary endpoint), patients are allowed to use low-potency TCS (for example, hydrocortisone 2.5% ointment) in combination with investigational product if they experience worsening in AD signs or symptoms. Details of rescue therapy and criteria are included in Section 7.7.3. Assessments of disease severity will be performed by the investigator at all study visits including unscheduled and ETVs.

The primary efficacy endpoint will be at Week 16 (Visit 8). All patients who permanently discontinue investigational product before the primary endpoint, including patients rescued with other systemic medications, should remain in the study to complete the schedule of study visits per protocol up to Week 16 (primary endpoint), when they will complete an ETV. If the patient refuses to continue up to Week 16 and wishes to withdraw consent, an ETV should be completed as soon as logistically possible.

At Week 16, all patients who meet IGA 0 or 1 and who have not required rescue therapy before Week 16 will be allowed to continue on this study. All other patients will be discontinued from this study and may be eligible to enroll in a separate open-label study (Study I4V-MC-JAIX [JAIX; BREEZE-AD6]). Patients need to complete at least 16 weeks in Study JAIW to be eligible to enroll in the open-label extension (Study JAIX). Patients who choose to enroll into the open-label study (Study JAIX) will not be required to complete the Study JAIW post-treatment follow up visit.

Patients experiencing worsening in disease severity resulting in an IGA score of ≥3 after Week 16 of Study JAIW will also have to be discontinued from this study and may be eligible to enroll in the open-label Study JAIX.

During the Screening Period (Period 1), a washout of systemic and topical treatments for AD is incorporated before randomization to minimize confounding effects because of background treatment. The double-blind, placebo-controlled treatment period (Period 2) is designed to minimize bias in the evaluation of the efficacy and safety of 2 baricitinib doses relative to placebo through the primary endpoint at Week 16.

In consideration of the disease severity, all patients in Study JAIW are eligible for rescue to TCS. Investigators are allowed to rescue patients who are experiencing unacceptable or worsening symptoms of AD. Once rescue medication is used, the patient will be determined to be a nonresponder (see Section 7.7.3).

Both EASI score and IGAs are commonly used in clinical trials, both for qualifying patients for enrollment and for evaluating treatment efficacy (Langley et al. 2015; Futamura et al. 2016; Bożek and Reich, 2017). There is no single "gold standard" disease severity scale for AD; however, IGA scales provide clinically meaningful measures to patients and investigators that are easily described and that correspond to disease severity categories (for example, moderate to severe), and a 75% improvement from Baseline (EASI75) is a commonly used measure of treatment effect in AD clinical trials.

The IGA scale that will be used in this trial, the validated Investigator's Global Assessment of Atopic Dermatis (vIGA-AD, referred to throughout the protocol as IGA; Appendix 11), has been developed internally and assesses AD severity using a 5-point scale.

The 16-week efficacy endpoint was chosen because it is probable that a robust clinical effect will be observed with baricitinib within this time frame on the basis of the Phase 2 study results in AD and previous studies in another inflammatory skin condition. Patients who do not achieve IGA 0 or 1 at Week 16 will be discontinued from the study. Similarly, patients who achieve an IGA of 0 or 1 at Week 16, and experience worsening in their disease resulting in an IGA score of ≥3 at any time after Week 16 will also be discontinued from the study. Patients who are discontinued from Study JAIW may be eligible to enroll in open-label Study JAIX.

The Post-Treatment Follow-Up Period (Period 3) is for safety monitoring after the patient has been off investigational product for approximately 28 days.

5.5. Justification for Dose

The doses proposed for Study JAIW are baricitinib 1-mg QD and 2-mg QD. These doses were chosen primarily on the basis of the recently completed Phase 2 AD study, Study JAHG, and are additionally supported by pharmacokinetic, safety, and efficacy data for baricitinib in Phase 2 and Phase 3 RA studies and a Phase 2 psoriasis study.

In the Phase 2 Study JAHG, both the 2-mg and 4-mg doses showed benefit on the primary and major secondary endpoints (EASI, IGA, SCORing Atopic Dermatitis [SCORAD], Patient-Oriented Eczema Measure [POEM], and Dermatology Life Quality Index [DLQI]) as compared to placebo, and both doses had an acceptable safety profile at Week 16. As baricitinib 2-mg QD and 4-mg QD doses showed similar efficacy for multiple endpoints, a lower dose,

9.1.2.4. Hospital Anxiety Depression Scale

The Hospital Anxiety Depression Scale (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 (for example, 0 to 3) for each question and is intended for ages 12 to 65 years (Zigmond and Snaith 1983; White et al. 1999). Scores for each domain (anxiety and depression) can range from 0 to 21, with higher scores indicating greater anxiety or depression (Zigmond and Snaith 1983; Snaith 2003).

9.1.3. Health Outcomes and Quality-of-Life Measures

The patient self-reported questionnaires will be administered via either an electronic patient diary or via an electronic tablet. Questionnaires will have been translated into the native language of the country and/or region and linguistically validated.

9.1.3.1. Patient-Oriented Eczema Measure

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).

9.1.3.2. Itch Numeric Rating Scale

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).

9.1.3.3. Atopic Dermatitis Sleep Scale

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 the previous night. Patient's 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 the previous night, item 2, by selecting the number of times they woke up each night, ranging from 0 to 29 times. The ADSS is designed to be completed each day with respondents thinking about sleep "last night." Each item is scored individually.

9.1.3.4. Skin Pain Numeric Rating Scale

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 previous 24 hours.

9.1.3.5. Patient Global Impression of Severity

The Patient Global Impression of Severity–Atopic Dermatitis (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."

9.1.3.6. Dermatology Life Quality Index

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 "not at all," "a little," "a lot," and "very much," with corresponding scores of 0, 1, 2, and 3, respectively, and unanswered ("not relevant") responses scored as 0 as applicable. Scores range from 0-30 with higher scores indicating greater impairment of quality of life. A DLQI total score of 0 to 1 is considered as having no effect on a patient's health-related quality of life (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).

9.1.3.7. European Quality of Life–5 Dimensions–5 Levels

The European Quality of Life–5 Dimensions–5 Levels (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 to 100 mm Visual Analog Scale (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 checking in the box associated with the most appropriate statement in each of the 5 dimensions. It should be noted that the numerals 1 to 5 have no arithmetic properties and should not be used as an ordinal score. The VAS records the respondent's self-rated health on a vertical VAS where the endpoints are labeled "best imaginable health state" and "worst imaginable health state." This information can be used as a quantitative measure of health outcome. The EQ-5D-5L health states, defined by the EQ-5D-5L descriptive system, may be converted into a single summary index by applying a formula that essentially attaches values (also called weights) to each of the levels in each dimension (Herdman et al. 2011; EuroQol Group 2015 [WWW]).

9.1.3.8. Work Productivity and Activity Impairment Questionnaire–Atopic Dermatitis

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.

9.1.4. Appropriateness of Assessments

All assessments utilized in this study are standard, widely used, and generally recognized as reliable, accurate, and relevant with the exception of ADSS and Skin Pain NRS, which are currently being developed and validated according to regulatory guidances.

9.2. Adverse Events

Investigators are responsible for monitoring the safety of patients who have entered this study and for alerting Lilly or its designee to any event that seems unusual, even if this event may be considered an unanticipated benefit to the patient.

The investigator is responsible for the appropriate medical care of patients during the study.

Investigators must document their review of each laboratory safety report.

The investigator remains responsible for following, through an appropriate health care option, AEs that are serious or otherwise medically important, considered related to the investigational product or the study, or that caused the patient to discontinue the investigational product before completing the study. The patient should be followed until the event resolves, stabilizes with appropriate diagnostic evaluation, or is reasonably explained. The frequency of follow-up evaluations of the AE is left to the discretion of the investigator.

Lack of drug effect is not an AE in clinical studies, because the purpose of the clinical study is to establish treatment effect.

After the ICF is signed, study site personnel will record via eCRF the occurrence and nature of each patient's preexisting conditions, including clinically significant signs and symptoms of the disease under treatment in the study. In addition, site personnel will record any change in the condition(s) and any new conditions as AEs. Investigators should record their assessment of the potential relatedness of each AE to investigational product, via eCRF.

The investigator will interpret and document whether or not an AE has a reasonable possibility of being related to study treatment, study device, or a study procedure, taking into account the disease, concomitant treatment, or pathologies.

A "reasonable possibility" means that there is a cause-and-effect relationship between the investigational product, study device, and/or study procedure and the AE.

The investigator answers yes/no when making this assessment.

Planned surgeries and nonsurgical interventions should not be reported as AEs unless the underlying medical condition has worsened during the course of the study.

If a patient's investigational product is discontinued as a result of an AE, study site personnel must report this to Lilly or its designee via eCRF, clarifying if possible the circumstances leading to any dosage modifications, or discontinuations of treatment.

9.2.1. Serious Adverse Events

An SAE is any AE from this study that results in 1 of the following outcomes:

- death
- initial or prolonged inpatient hospitalization
- a life-threatening experience (that is, immediate risk of dying)
- persistent or significant disability/incapacity
- congenital anomaly/birth defect
- important medical events that may not be immediately life threatening or result in death or hospitalization but may jeopardize the patient or may require intervention to prevent 1 of the other outcomes listed in the definition above. Examples of such medical events include allergic bronchospasm requiring intensive treatment in an emergency room or at home, blood dyscrasias or convulsions that do not result in inpatient hospitalization, or the development of drug dependency or drug abuse.

All AEs occurring after signing the ICF are recorded in the eCRF and assessed for serious criteria. The SAE reporting to the sponsor begins after the patient has signed the ICF and has received investigational product. However, if an SAE occurs after signing the ICF, but prior to receiving investigational product, the SAE should be reported to the sponsor as per SAE reporting requirements and timelines if it is considered reasonably possibly related to study procedure.

Study site personnel must alert Lilly or its designee of any SAE within 24 hours of investigator awareness of the event via a sponsor-approved method. If alerts are issued via telephone, they are to be immediately followed with official notification on study-specific SAE forms. This 24-hour notification requirement refers to the initial SAE information and all follow-up SAE information. Patients with a serious hepatic AE should have additional data collected using the hepatic safety eCRF.

Pregnancy (during maternal or paternal exposure to investigational product) does not meet the definition of an AE. However, to fulfill regulatory requirements, any pregnancy should be reported following the SAE process to collect data on the outcome for both mother and fetus.

Investigators are not obligated to actively seek AEs or SAEs in subjects once they have discontinued and/or completed the study (the patient summary CRF has been completed). However, if the investigator learns of any SAE, including a death, at any time after a subject has been discharged from the study, and he or she considers the event reasonably possibly related to the study treatment or study participation, the investigator must promptly notify Lilly.

9.2.1.1. Suspected Unexpected Serious Adverse Reactions

Suspected unexpected serious adverse reactions (SUSARs) are serious events that are not listed in the IB and that the investigator identifies as related to investigational product or procedure. United States 21 CFR 312.32 and European Union Clinical Trial Directive 2001/20/EC and the associated detailed guidances or national regulatory requirements in participating countries require the reporting of SUSARs. Lilly has procedures that will be followed for the recording

9.4.3. Physical Exam

For each patient, a complete physical examination (excluding pelvic and rectal examinations) will be performed at Visit 1 (Screening). A symptom-directed physical examination will be performed at other visits as specified in the Schedule of Activities (Section 2). A complete physical examination may be repeated at the investigator's discretion at any time a patient presents with physical complaints.

9.4.4. Laboratory Tests

For each patient, laboratory tests detailed in Appendix 2 should be conducted according to the Schedule of Activities (Section 2). With the exception of laboratory test results that may unblind the study, Lilly or its designee will provide the investigator with the results of laboratory tests analyzed by a central vendor, if a central vendor is used for the clinical trial.

9.4.5. Columbia Suicide Severity Rating Scale

The C-SSRS captures the occurrence, severity, and frequency of suicidal ideation and/or behavior during the assessment period. The scale includes suggested questions to solicit the type of information needed to determine if suicidal ideation and/or behavior occurred. The C-SSRS is administered by an appropriately trained health care professional with at least 1 year of patient care/clinical experience. The tool was developed by the National Institute of Mental Health trial group for the purpose of being a counterpart to the Columbia Classification Algorithm of Suicide Assessment categorization of suicidal events. For this study, the scale has been adapted (with permission from the scale authors) to include only the portion of the scale that captures the occurrence of the 11 preferred ideation and behavior categories.

The nonleading AE collection should occur before the collection of the C-SSRS. If a suicide-related event is discovered *during the C-SSRS* but was not captured during the nonleading AE collection, sites should not change the AE form. If an event is serious or leads to discontinuation, this is an exception, and the SAE and/or AE leading to discontinuation should be included on the AE form, and the process for reporting SAEs should be followed.

9.4.6. Self-Harm and Follow-Up Supplement Forms

Suicide-related events (behavior and/or ideations) will be assessed and evaluated at every visit with the administration of the C-SSRS and the Self-Harm Supplement Form. The Self-Harm Supplement Form is a single question to enter the number of suicidal behavior events, possible suicide behaviors, or nonsuicidal self-injurious behaviors. If the number of behavioral events is greater than zero, it will lead to the completion of the self-harm follow-up form. The self-harm follow-up form is a series of questions that provides a more detailed description of the behavior cases.

9.4.7. Chest X-Ray and Tuberculosis Testing

A posterior–anterior view chest x-ray will be obtained locally at screening (Visit 1), unless results from a chest x-ray obtained within 6 months before the study are available. The chest x-ray will be reviewed by the investigator or his or her designee to exclude patients with active

the DMC, including its operating characteristics, will be documented in a DMC charter and DMC analysis plan.

10.3.6.2. Unblinded Study Team for Early Regulatory Submission

Besides DMC members, a limited number of preidentified individuals may gain access to the limited unblinded data, as specified in the unblinding plan, before the interim or final database lock to initiate or for preparation of regulatory documents. The study may be terminated prematurely on the basis of futility following the Week 16 interim analysis. Although this is an interim analysis with respect to the entire study, it is the only and final analysis for the primary and all key secondary endpoints. Therefore, all the α is allocated for this interim analysis only and no other α is allocated for other endpoints assessed beyond Week 16. 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 will be specified in a separate unblinding plan document.

10.3.6.3. Adjudication Committee

A blinded Clinical Event Committee will adjudicate potential MACE (cardiovascular death, MI, stroke), other cardiovascular events (such as hospitalization for unstable angina, hospitalization for heart failure, serious arrhythmia, resuscitated sudden death, cardiogenic shock, coronary revascularization such as coronary artery bypass graft or percutaneous coronary intervention), VTEs, and noncardiovascular deaths. Details of membership, operations, recommendations from the Committee, and the communication plan will be documented in the Charter.

Investigational product	A pharmaceutical form of an active ingredient or placebo being tested or used as a reference in a clinical trial, including products already on the market when used or assembled (formulated or packaged) in a way different from the authorized form, or marketed products used for an unauthorized indication, or marketed products used to gain further information about the authorized form.
IWRS	interactive web-response system
JAK	Janus kinase
LSM	least squares mean
MACE	major adverse cardiovascular events
MI	myocardial infarction
MMRM	mixed-effects model of repeated measures
NRI	nonresponder imputation
NRS	Numeric Rating Scale
PDE-4 inhibitor	phosphodiesterase type 4 inhibitor
pMI	placebo multiple imputation
POEM	Patient-Oriented Eczema Measure
PPD	purified protein derivative
QD	once daily
RA	rheumatoid arthritis
SAE	serious adverse event
SAP	statistical analysis plan
SCORAD	SCORing Atopic Dermatitis
SUSAR	suspected unexpected serious adverse reaction
TB	Tuberculosis
TBL	total bilirubin level
TCNI	topical calcineurin inhibitor
TCS	topical corticosteroids
TEAE	Treatment-emergent adverse event: An untoward medical occurrence that emerges during a defined treatment period, having been absent pretreatment, or worsens relative to the pretreatment state, which does not necessarily have to have a causal relationship with this treatment.
TSH	thyroid-stimulating hormone
TSLP	thymic stromal lymphopoietin
TYK2	tyrosine kinase 2
ULN	upper limit of normal
VAS	Visual Analog Scale
vIGA-AD	validated Investigator's Global Assessment for Atopic Dermatitis
VTE	venous thromboembolic event (deep vein thrombosis or pulmonary embolism)
WPAI-AD	The Work Productivity and Activity Impairment–Atopic Dermatitis

Objectives	Endpoints
To compare the efficacy of baricitinib 1-mg QD or 2-mg QD to placebo in AD during the 16-week, double-blind, placebo-controlled treatment period as assessed by patient-reported outcome/QoL measures	• Mean percent change from baseline in Itch NRS at 1 week and 16 weeks • Mean change from baseline in Itch NRS at 4 weeks and 16 weeks • Mean change from baseline in the total score of the POEM at Week 16 • Mean change in PGI-S-AD scores at Week 16 • Mean change from baseline in the HADS at Week 16 • Mean change in DLQI scores at Week 16 • Mean change in WPAI-AD scores at Week 16 • Mean change in EQ-5D-5L scores at Week 16
Other Secondary Objectives for responders beyond Week 16	
To describe the long-term efficacy of baricitinib 1-mg QD or 2-mg QD in AD as measured by improvement in signs and symptoms of AD	• Proportion of patients with a response of IGA 0 or 1 at Week 16 who maintain an IGA 0 or 1 at Weeks 28, 52, and 104 • Proportion of patients with a response of IGA 0 or 1 at Week 16 who achieve EASI75 assessed at Weeks 28, 52, and 104 • Proportion of patients with a response of IGA 0 or 1 at Week 16 who achieve SCORAD75 at Weeks 28, 52, and 104 • Mean percent change from baseline in EASI score at Weeks 28, 52, and 104 • Mean percent change from baseline in SCORAD score at Weeks 28, 52, and 104 • Mean percent change from baseline in SCORAD pruritus at Weeks 28, 52, and 104 • Mean percent change from baseline in POEM at Weeks 28, 52, and 104

Abbreviations: AD = atopic dermatitis; ADSS = Atopic Dermatitis Sleep Scale; DLQI = Dermatology Life Quality Index; EASI = Eczema Area and Severity Index; EQ-5D-5L = the European Quality of Life–5 Dimensions–5 Levels; HADS = Hospital Anxiety Depression Scale; IGA = Investigator's Global Assessment; NRS = numeric rating scale; QD = once daily; PGI-S-AD = Patient Global Impression of Severity–Atopic Dermatitis; POEM = Patient-Oriented Eczema Measure; QoL = quality of life; SCORAD = SCORing Atopic Dermatitis; WPAI-AD = Work Productivity and Activity Impairment – Atopic Dermatitis.

Summary of Study Design:

Study I4V-MC-JAIW (JAIW) is a Phase 3, multicenter, randomized, double-blind, placebo-controlled, parallel-group, outpatient study evaluating the efficacy and safety of baricitinib 2-mg once daily (QD) and 1-mg QD as compared to placebo in adult patients with moderate to severe AD. The study population will include patients aged ≥18 years who have moderate to severe AD and a history of inadequate response or intolerance to existing topical therapies.

	Period 1: Screening	Period 2: Double-Blinded Treatment Period														Period 3: PTFU
Visit Number	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15/ET[a]**	**801[b]**
Weeks from Randomization		**0**	**1**	**2**	**4**	**8**	**12**	**16**	**28**	**40**	**52**	**64**	**76**	**88**	**104**	**108**
Visit tolerance interval (days)	**-8 to -35**	**0**	**±2**	**±2**	**±2**	**±4**	**±4**	**±4**	**±7**	**±7**	**±7**	**±7**	**±7**	**±7**	**±4**	**28±4 after last dose**
Urinalysis	X	X			X	X	X	X	X	X	X	X	X	X	X	X
Urine pregnancy[p]		X		X	X	X	X	X	X	X	X	X	X	X	X	X
Pharmacogenetics: blood		X														
Serum immunoglobulin (IgE)		X			X			X			X				X	
Exploratory storage samples (serum and plasma)		X			X			X			X				X	
RNA and biomarkers: blood		X			X			X			X				X	

Abbreviations: AD = atopic dermatitis; ADSS = Atopic Dermatitis Sleep Scale; BP = blood pressure; C-SSRS = Columbia Suicide Severity Rating Scale 11 categories suicidal ideation/suicidal behavior; DLQI = Dermatology Life Quality Index; EASI = Eczema Area and Severity Index; ECG = electrocardiogram; EQ-5D-5L = the European Quality of Life–5 Dimensions 5 Levels; ET = early termination; ePRO = electronic patient-reported outcomes (device); FSH = follicle-stimulating hormone; HADS = Hospital Anxiety Depression Scale; HBcAb = Hepatitis B core antibody; HBsAb = Hepatitis B surface antibody; HBV = hepatitis B virus; HCV = hepatitis C virus; HIV = human immunodeficiency virus; IGA = Investigator's Global Assessment; IgE = immunoglobulin E; IP = investigational product; IWRS = interactive web-response system; NRS = numeric rating scale; PGI-S-AD = Patient Global Impression of Severity–Atopic Dermatitis; POEM = Patient-Oriented Eczema Measure; PPD = purified protein derivative; PTFU = posttreatment follow-up; SCORAD = SCORing Atopic Dermatitis; TB = tuberculosis; TCS = topical corticosteroids; TSH = thyroid-stimulating hormone; WPAI-AD = Work Productivity and Activity Impairment-Atopic Dermatitis.

[a] An ET visit should be conducted if a patient discontinues from the study before Week 104. ET visit activities do not need to be duplicated if occurring at the time of a scheduled visit.

[b] Visit 801 is the PTFU visit, which occurs after the patient has been off baricitinib/IP for approximately 4 weeks. Patients who have permanently discontinued IP but remain in the study for more than 28 days without IP will only complete Visit 15/ET, and Visit 801 (follow-up visit) is not required. Patients choosing to enroll into the open-label study, Study I4V-MC-JAIX (JAIX), will not be required to complete Visit 801.

[c] The symptom-directed physical examination may be repeated at the investigator's discretion any time a patient presents with physical complaints.

[d] A posterior–anterior chest x-ray will be performed at screening unless one has been performed within the past 6 months and the x-ray and reports are available.

[e] TB test(s) include PPD, QuantiFERON®-TB Gold, and T SPOT®. See Exclusion Criterion [38] for description of TB testing. In countries where the QuantiFERON-TB Gold test or T-SPOT is available, either test may be used instead of the PPD TB test. The QuantiFERON-TB Gold test may be performed locally or centrally; the T-SPOT must be performed locally. (Note: Exception: Patients with a history of active or latent TB who have documented evidence of appropriate treatment, have no history of re-exposure since their treatment was completed, and have a screening chest x-ray with no evidence of active TB may be enrolled if other entry criteria are met. Such patients would not be required to undergo the protocol-specific TB testing but must have a chest x-ray at

Abbreviations: AD = atopic dermatitis; ADSS = Atopic Dermatitis Sleep Scale; DLQI = Dermatology Life Quality Index; EASI = Eczema Area and Severity Index; EQ-5D-5L = the European Quality of Life–5 Dimensions–5 Levels; HADS = Hospital Anxiety Depression Scale; IgE = immunoglobulin E; IGA = Investigator's Global Assessment; NRS = numeric rating scale; QD = once daily; QoL = quality of life; 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.

Patients who complete week 104 (Visit 15) will have the option to transition to open-label Study JAIX if eligibility criteria are met, regardless of responder status, or continue to the post-treatment follow-up.

5.1.3. Period 3: Post-Treatment Follow-Up

Patients who complete the study through Visit 15 (Week 104) will have a post-treatment follow-up visit (Visit 801) approximately 28 days after the last dose of investigational product, unless the patient is transitioning to open-label Study JAIX.

Patients who have received at least 1 dose of investigational product and discontinue early from the study must have an ETV and return for the post-treatment safety follow-up visit (Visit 801) approximately 28 days after the last dose of investigational product.

Patients who have discontinued investigational product but remain in the study for more than 28 days without investigational product will have an ETV if they choose to discontinue early; however, a separate follow-up visit (Visit 801) is not required.

Patients should not initiate new systemic AD treatment during this period. However, if patients or investigators must initiate treatment, investigators should make every attempt to conduct efficacy and safety assessments immediately before administering any new treatment. An unscheduled visit can be used for this purpose if necessary.

For patients who are discontinued from this study because of lack of efficacy (either at Week 16 or after) and who opt to enroll into the open-label extension (Study JAIX), a separate follow-up visit (Visit 801) is not required.

Figure JAIW.1 illustrates the study design. The 3 dosing regimens are described in Section 7.1. The blinding procedure is described in Section 7.3.

baricitinib 1-mg QD, was included in this study to assess if a lower dose can also provide reasonable efficacy in AD.

and expedited reporting of SUSARs that are consistent with global regulations and the associated detailed guidances.

9.2.2. Adverse Events of Special Interest

AEs of special interest will include the following:

- infections (including TB, herpes zoster, or opportunistic infections)
- malignancies
- hepatic events (see Section 9.4.9)
- major adverse cardiovascular events (MACE) (see Section 9.4.11)
- thrombotic events (such as deep vein thrombosis and pulmonary embolism) (see Section 9.4.10).

Sites will provide details on these AEs as instructed on the eCRF and may be asked for additional description by Lilly.

9.2.3. Complaint Handling

Lilly collects product complaints on investigational products and drug delivery systems used in clinical studies to ensure the safety of study participants, monitor quality, and to facilitate process and product improvements.

Patients will be instructed to contact the investigator as soon as possible if he or she has a complaint or problem with the investigational product so that the situation can be assessed.

9.3. Treatment of Overdose

Refer to the IB.

9.4. Safety

Any clinically significant findings from ECG testing, physical examination, vital signs measurements, or laboratory measurements that result in a diagnosis and that occur after the patient receives the first dose of study treatment should be reported to Lilly or its designee as an AE via eCRF.

9.4.1. Electrocardiograms

A single 12-lead standard ECG will be obtained locally at Visit 1 and read by a qualified physician (the investigator or qualified designee) at the site to determine whether the patient meets entry criteria.

ECGs may be obtained at additional times, when deemed clinically necessary.

9.4.2. Vital Signs

For each patient, vital signs should be measured according to the Schedule of Activities (Section 2).

TB infection. In addition, patients will be tested at screening (Visit 1) for evidence of active or latent TB as described in the exclusion criteria, Section 6.2.

Investigators should follow local guidelines for monitoring patients for TB if a patient is at high risk for acquiring TB or reactivation of latent TB.

9.4.8. Hepatitis B Virus DNA Monitoring

Patients who are HBcAb positive and HBV DNA negative (undetectable) at Visit 1 will require measurement of HBV DNA per Schedule of Activities, regardless of their hepatitis B surface antibody status.

The following actions should be taken in response to HBV DNA test results:

- If a single result is obtained with a value "below limit of quantitation," the test should be repeated within approximately 2 weeks. If the repeat test result is "target not detected," monitoring will resume as specified in the Schedule of Activities (Section 2).
- If the patient has 2 or more test results with a value "below limit of quantitation" or a test result above the limit of quantitation, the patient will be permanently discontinued from investigational product (see Section 8.1.2) and should be referred to a hepatology specialist.

9.4.9. Hepatic Safety Monitoring and Data Collection

If a study patient experiences elevated ALT ≥3 × ULN, ALP ≥2 × ULN, or elevated TBL ≥2 × ULN, liver testing (Appendix 4) should be repeated within 3 to 5 days including ALT, AST, ALP, TBL, direct bilirubin, gamma-glutamyl transferase, and creatine kinase to confirm the abnormality and to determine if it is increasing or decreasing. If the abnormality persists or worsens, clinical and laboratory monitoring should be initiated by the investigator in consultation with the study medical monitor. Monitoring of ALT, AST, TBL, and ALP should continue until levels normalize or return to approximate baseline levels.

Criteria for discontinuation of investigational products (either temporary interruption or permanent discontinuation) due to abnormal ALT, AST, TBL, or ALP are detailed in Section 8.1.

Additional safety data should be collected via the hepatic eCRF if 1 or more of the following conditions occur:

- elevation of serum ALT to ≥5 × ULN on 2 or more consecutive blood tests
- elevated serum TBL to ≥2 × ULN (except for cases of known Gilbert's syndrome)
- elevation of serum ALP to ≥2 × ULN on 2 or more consecutive blood tests
- patient discontinued from treatment due to a hepatic event or abnormality of liver tests
- hepatic event considered to be an SAE

11. References

Alosh M, Bretz F, Huque M. Advanced multiplicity adjustment methods in clinical trials. *Stat Med.* 2014;33(4):693-713.

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

Bieber T. Atopic Dermatitis. *Ann Dermatol*. 2010;22(2):125-137.

Bieber T, Straeter B. Off-label prescriptions for atopic dermatitis in Europe. *Allergy*. 2015;70(1):6-11.

Bożek A, Reich, A. Assessment of Intra- and Inter-Rater Reliability of Three Methods for Measuring Atopic Dermatitis Severity: EASI, Objective SCORAD, and IGA. *Dermatology*. 2017;233(1):16-22.

Bretz F, Posch M, Glimm E, Klinglmueller F, Maurer W, Rohmeyer K. Graphical approaches for multiple comparison procedures using weighted Bonferroni, Simes, or parametric tests. *Biom J*. 2011;53(6):894-913.

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.

Clark JD, Flanagan ME, Telliez JB. Discovery and development of Janus kinase (JAK) inhibitors for inflammatory diseases. *J Med Chem*. 2014;57(12):5023-5038.

Divekar R, Kita H. Recent advances in epithelium-derived cytokines (IL-33, IL-25, and thymic stromal lymphopoietin) and allergic inflammation. *Curr Opin Allergy Clin Immunol*. 2015;15(1):98-103.

Dougados M, van der Heijde D, Chen YC, Greenwald M, Drescher E, Liu J, Beattie S, Witt S, de la Torre I, Gaich C, Rooney T, Schlichting D, de Bono S, Emery P. Baricitinib in patients with inadequate response or intolerance to conventional synthetic DMARDs: results from the RA-BUILD study. *Ann Rheum Dis*. 2017;76(1):88-95.

Eichenfield LF, Tom WL, Chamlin SL, Feldman SR, Hanifin JM, Simpson EL, Berger TG, Bergman JN, Cohen DE, Cooper KD, Cordoro KM, Davis DM, Krol A, Margolis DJ, Paller AS, Schwarzenberger K, Silverman RA, Williams HC, Elmets CA, Block J, Harrod CG, Smith Begolka W, Sidbury R. Guidelines of care for the management of atopic dermatitis: section 1. Diagnosis and assessment of atopic dermatitis. *J Am Acad Dermatol*. 2014;70(2):338-351.

EuroQol Group. EQ-5D-5L User Guide. Version 2.1. Available at: https://euroqol.org/wp-content/uploads/2016/09/EQ-5D-5L_UserGuide_2015.pdf. Published April 2015. Accessed December 2017.

Fleischmann R, Schiff M, van der Heijde D, Ramos-Remus C, Spindler A, Stanislav M, Zerbini CA, Gurbuz S, Dickson C, de Bono S, Schlichting D, Beattie S, Kuo WL, Rooney T, Macias W, Takeuchi T. Baricitinib, methotrexate, or combination in patients with rheumatoid arthritis

Appendix 2. Clinical Laboratory Tests

Hematology[a,b]	**Clinical Chemistry**[a,b]
Hemoglobin	**Serum Concentrations of:**
Hematocrit	Sodium
Erythrocyte count (RBC)	Potassium
Absolute Reticulocyte Count	Total bilirubin
Mean cell volume	Direct bilirubin
Mean cell hemoglobin	Alkaline phosphatase
Mean cell hemoglobin concentration	Alanine aminotransferase (ALT)
Leukocytes (WBC)	Aspartate aminotransferase (AST)
Platelets	Blood urea nitrogen (BUN)
Absolute counts of:	Creatinine
Neutrophils, segmented	Cystatin C
Neutrophils, juvenile (bands)	Uric acid
Lymphocytes	Calcium
Monocytes	Glucose
Eosinophils	Albumin
Basophils	Total protein
	Estimated glomerular filtration rate (eGFR)[e]
Urinalysis[a,b,d]	Creatine phosphokinase (CPK)
Color	
Specific gravity	**Other Tests**[a]
pH	Hepatitis B Surface antigen (HBsAg)[f]
Protein	Anti-Hepatitis B Core antibody (HBcAb)[f]
Glucose	HBV DNA[k]
Ketones	Anti-Hepatitis B Surface antibody (HBsAb)[f]
Bilirubin	Human immunodeficiency virus (HIV)[f]
Urobilinogen	Hepatitis C antibody[f,g]
Blood	Thyroid-stimulating hormone (TSH)
Leukocyte esterase	Exploratory storage samples (serum, plasma and mRNA)
Nitrite	Pregnancy Test[h]
	Follicle-stimulating hormone[f,i]
Lipids[a,c]	Serum immunoglobulin (IgE)
Total cholesterol	QuantiFERON®-TB Gold or T-SPOT® TB[j]
Low-density lipoprotein	PPD (local testing)
High-density lipoprotein	
Triglycerides	

Abbreviations: FSH = follicle-stimulating hormone; HBV = hepatitis B virus; PPD = purified protein derivative; RBC = red blood cell; TB = tuberculosis; WBC = white blood cell.

a Assayed by sponsor-designated laboratory.

b Unscheduled or repeat blood chemistry, hematology, and urinalysis panels may be performed at the discretion of the investigator, as needed.

c Fasting lipid profile. Patients should not eat or drink anything except water for 12 hours prior to test. If a patient attends these visits in a nonfasting state, this will not be considered a protocol violation.

d Microscopic examination of sediment performed only if abnormalities are noted on the routine urinalysis.