Objectives	Endpoints
To compare the efficacy of baricitinib 2-mg QD + TCS or baricitinib 4-mg QD + TCS to placebo + TCS 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 2 days, 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 16 weeks.
Other Secondary Objectives *These are prespecified objectives that will not be adjusted for multiplicity.*	
To compare the efficacy of baricitinib 2-mg QD + TCS or baricitinib 4-mg QD + TCS to placebo + TCS 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 16 weeks • Proportion of patients achieving IGA of 0 at 16 weeks • Mean change from baseline in SCORAD at 16 weeks • Proportion of patients achieving SCORAD90 at 16 weeks • Mean change from baseline in body surface area affected at 16 weeks • Proportion of patients developing skin infections requiring antibiotic treatment by Week 16 • Mean gram quantity of background TCS used over 16 weeks (tube weights)
To compare the efficacy of baricitinib 2-mg QD + TCS or baricitinib 4-mg QD + TCS to placebo + TCS in AD during the 16-week, double-blind, placebo-controlled treatment period as assessed by patient-reported outcome/QoL measures.	• Percent change from baseline in Itch NRS at 2 days, 1 week, 4 weeks, and 16 weeks • Mean change from baseline in Itch NRS at 2 days, 1 week, 4 weeks, and 16 weeks • Mean change from baseline in the total score of the POEM at 16 weeks • Mean change in PGI-S-AD scores at 16 weeks • Mean change from baseline in the HADS at 16 weeks • Mean change in DLQI scores at 16 weeks • Mean change in WPAI scores at 16 weeks • Mean change in EQ-5D-5L scores at 16 weeks. • Mean number of days without use of background TCS over 16 weeks

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; TCS = topical corticosteroids; WPAI = Work Productivity and Activity Impairment.

Summary of Study Design:

Study I4V-MC-JAIY (JAIY) is a Phase 3, multicenter, randomized, double-blind, placebo-controlled, parallel-group, outpatient study evaluating the efficacy and safety of baricitinib 4-mg once daily (QD) plus topical corticosteroids (TCS) and 2-mg QD plus TCS, as compared to placebo plus TCS in adult patients with moderate to severe AD. The study population will include patients aged 18 years or older who have moderate to severe AD and a history of inadequate response to existing topical therapies.

	Screening	Double-Blind, Placebo-Controlled							Post-treatment Follow-up
	Period 1	Period 2							Period 3
Visit Number	1	2	3	4	5	6	7	8	801[a]
Weeks from Randomization		0	1	2	4	8	12	16 or ET	
Days from Randomization		0	7	14	28	56	84	112	
Visit Tolerance Interval (days)	-8 to -35		±2	±2	±2	±4	±4	±4	28 ± 4 after last dose
Randomization		X							
IWRS	X	X	X	X	X	X	X	X	X
IP dispensed		X	X	X	X	X	X		
IP returned and compliance assessed			X	X	X	X	X	X	
Dispense background TCS[h]		X	X	X	X	X	X	X	
Weigh (tube with cap) and record returned background TCS[h]			X	X	X	X	X	X	X
Scales									
IGA	X	X	X	X	X	X	X	X	X
EASI	X	X	X	X	X	X	X	X	X
SCORAD	X	X	X	X	X	X	X	X	X
Health Outcome Measures and Other Questionnaires[i]									
POEM	X	X	X	X	X	X	X	X	X
DLQI	X	X	X	X	X	X		X	X
HADS	X	X		X	X	X		X	
EQ-5D-5L		X	X	X	X	X	X	X	X
WPAI-AD		X	X	X	X	X	X	X	X
Itch NRS	X	X	X	X	X	X	X	X	X
Skin Pain NRS	X	X	X	X	X	X	X	X	X
ADSS	X	X	X	X	X	X	X	X	X
PGI-S-AD	X	X	X	X	X	X	X	X	X
PIQ – General[j]	X	X	X	X	X	X	X	X	X
PIQ – Activity and Clothing[j]	X	X	X	X	X	X	X	X	X
PIQ – Mood and Sleep[j]	X	X	X	X	X	X	X	X	X
PIQ – Scratching Behavior[j]	X	X	X	X	X	X	X	X	X

Objectives	Endpoints
To compare the efficacy of baricitinib 2-mg QD + TCS or baricitinib 4-mg QD + TCS to placebo + TCS in AD during the 16-week, double-blind, placebo-controlled treatment period as assessed by patient-reported outcome/QoL measures.	• Percent change from baseline in Itch NRS at 2 days, 1 week, 4 weeks, and 16 weeks • Mean change from baseline in Itch NRS at 2 days, 1 week, 4 weeks, and 16 weeks • Mean change from baseline in the total score of the POEM at 16 weeks • Mean change in PGI-S-AD scores at 16 weeks • Mean change from baseline in the HADS at 16 weeks • Mean change in DLQI scores at 16 weeks • Mean change in WPAI scores at 16 weeks • Mean change in EQ-5D-5L scores at 16 weeks • Mean number of days without use of background TCS over 16 weeks

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 PIQ – General score
- Mean change from baseline in PIQ – Activity and Clothing score
- Mean change from baseline in PIQ – Mood and Sleep score
- Mean change from baseline in PIQ – Scratching Behavior score
- Mean change from baseline in PROMIS – Sleep-Related Impairment score
- Mean change from baseline in Neuro-QoL – Cognitive Function score
- Mean change from baseline in Patient Benefit Index score – global score plus the following subscales:
    - Reducing social impairments
    - Reducing psychological impairments
    - Reducing impairments due to therapy
    - Reducing physical impairments
    - Having confidence in healing.
- 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
- To evaluate changes from baseline in IgE levels during the study
- To evaluate changes from baseline in eosinophil levels during the study

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; Neuro-QoL = Quality of Life in Neurological Disorders; NRS = numeric rating scale; PIQ = Pain Impact Questionnaire; PROMIS = Patient-Reported Outcomes Measurement Information System; 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; TCS = topical corticosteroids; WPAI = Work Productivity and Activity Impairment.

guidelines for vaccination of those >18 years of age with nonlive vaccines intended to prevent infectious disease prior to entering patients into the study.

Patients who meet all of the inclusion and none of the exclusion criteria (Section 6) will continue to Visit 2.

5.1.2. Period 2: Double–Blind, Placebo-Controlled Treatment

At Visit 2 (Week 0, baseline), study eligibility for each patient will be reviewed, based on all inclusion and exclusion criteria (Section 6), and laboratory test results. Patients who meet all criteria will proceed to randomization and begin the 16-week 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 at the clinic.

Patient will be randomized at a 1:1:1 ratio into 1 of the 3 treatment groups (placebo QD, baricitinib 2-mg QD, or baricitinib 4-mg QD). Investigational product will be administered daily for 16 weeks (treatment period Visits 2 through 8; Section 7). All patients will be required to use emollients daily. Daily diaries will continue to be utilized throughout the treatment period. Download of this data will be required at study visits. TCS will be dispensed at V2 and used on affected areas as described in section 7.7.2. Topical calcineurin inhibitors (TCNIs) is also allowed, but TCNI use should be limited to problem areas (e.g. face and skin folds). The use of higher potency TCS and systemic therapies for the treatment of AD are not allowed, except as part of rescue therapy for patients not responding to treatment. Details of background topical therapy, as well as rescue therapy and rescue criteria are included in Section 7.7. Assessments of disease severity will be performed by the investigator at all study visits including unscheduled and early termination visits (ETVs).

The primary efficacy endpoint and final visit in the treatment period will be at Week 16 (Visit 8). Patients who complete through the Week 16 study visit may be eligible for inclusion in the long-term extension study JAHN (up to 2 additional years of treatment).

If a patient discontinues investigational product for any reason, the patient should remain in the study through Week 16 (Visit 8). If the patient refuses and wishes to withdraw consent, an ETV should be performed as soon as logistically possible.

5.1.3. Period 3: Post-Treatment Follow-up

Patients who complete the study through Visit 8 (Week 16) and do not enter the long-term extension study will have a post-treatment follow-up visit (Visit 801) approximately 28 days after the last dose of investigational product.

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.

In the Phase 2 Study JAHG, both the 2-mg and 4-mg doses showed benefit on the primary and major secondary endpoints (Eczema Area and Severity Index [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 (Guttman-Yassky et al. 2018). However, the 4-mg dose appeared to demonstrate a more rapid benefit (at 4 weeks) on the more stringent endpoints (improvement of at least 75% in EASI score [EASI75], improvement of at least 90% in EASI score [EASI90], and IGA 0 or 1) compared to 2-mg dose particularly in the subgroup of patients with baseline EASI scores ≥16. The 4-mg dose resulted in statistically significant improvement in these endpoints at Week 4 and this level of response was maintained through Week 16. A similar trend between the baricitinib 4-mg and 2-mg doses was observed in patients with RA. Although in Study JAHG, the 4-mg dose seemed to perform better than the 2-mg dose on more stringent endpoints, on other endpoints, including an improvement of at least 50% in EASI score [EASI50], and EASI change from baseline, 2-mg and 4-mg doses showed similar efficacy compared to placebo. Therefore, both doses will be tested in Study JAIY.

5.5.1. Dose Adjustment for Renal Impairment

Baricitinib exposure increases with decreased renal function. Based on PK simulations of baricitinib exposures for the mild and moderate categories of renal function (stratified as estimated glomerular filtration rate [eGFR] 60 to <90 mL/min/1.73 m^2 and eGFR 30 to <60 mL/min/1.73 m^2, respectively), dose adjustment is not required for patients with eGFR ≥60 mL/min/1.73 m^2. Patients with eGFR <60 mL/min/1.73 m^2 who are randomized to the 4-mg dose will receive a dose of 2-mg QD, which will ensure that exposures do not exceed those of the 4-mg QD dose in patients with eGFR ≥60 mL/min/1.73 m^2. For patients randomized to the 2-mg dose, there will be no dose adjustment based on renal function. The dose adjustment for renal impairment will be managed by interactive web-response system (IWRS) to ensure maintenance of the treatment blind. This study will not enroll patients with screening eGFR <40 mL/min/1.73 m^2. See Section 8.1.1 for eGFR thresholds that trigger interruption of investigational product.

The procedure of dose adjustment based on renal function (eGFR) during the study is detailed in Section 7.2.2.

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 and in countries where the questionnaires have been translated into the native language of the 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 to 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 last 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 last night, item 2, by selecting the number of times they woke up each night, ranging from 0 to 29 times. The ADSS is designed to be completed each day with respondents thinking about sleep "last night." Each item is scored individually.

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 past 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 lot," and "very much," with corresponding scores of 1, 2, and 3, respectively, and unanswered ("not relevant") responses scored as 0. Scores range from 0 to 30 with higher scores indicating greater impairment of quality of life. A DLQI total score of 0 to 1 is considered as having no effect on a patient's health-related QoL (Hongbo et al. 2005), and a 4-point change from baseline is considered as the minimal clinically important difference threshold (Khilji et al. 2002; Basra et al. 2015).

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 ticking (or placing a cross) in the box associated with the most appropriate statement in each of the 5 dimensions. It should be noted that the numerals 1 to 5 have no arithmetic properties and should not be used as an ordinal score. The VAS records the respondent's self-rated health on a vertical VAS where the endpoints are 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.

9.1.3.9. PROMIS Itch Questionnaire General Short Form 8a v1.0

Patient-Reported Outcomes Measurement Information System (PROMIS) is a set of person-centered measures that evaluates and monitors physical, mental, and social health in adults and children. It can be used with the general population and with individuals living with chronic conditions. The Itch General Short Form 8a (PIQ – General) within the PROMIS itch item bank consists of 8 items assessing the impact of itch on various aspects of life. Response

options range from 1=Never, 2=Rarely; 3=Sometimes; 4=Often; to 5=Almost always; total raw scores are converted to T-Scores with higher scores representing greater impact because of itch.

9.1.3.10. PROMIS Itch Questionnaire Activity and Clothing Short Form 8a v1.0

The Activity and Clothing Short Form 8a (PIQ – Activity and Clothing) within the PROMIS itch item bank consists of 8 items assessing activity and clothing related quality of life impairment from itch in adults "in the past 7 days". Response options range from 1=Never, 2=Rarely; 3=Sometimes; 4=Often; to 5=Almost always; total raw scores are converted to T-Scores with higher scores representing greater impact on activity and clothing because of itch.

9.1.3.11. PROMIS Itch Questionnaire Mood and Sleep Short Form 8a v1.0

The Mood and Sleep Short Form 8a (PIQ – Mood and Sleep) within the PROMIS Itch item bank consists of 8 items assessing mood and sleep related quality of life impairment from itch and impact of itch "in the past 7 days". Response options range from 1=Never, 2=Rarely; 3=Sometimes; 4=Often; to 5=Almost always; total raw scores are converted to T-Scores with higher scores representing greater impact on mood and sleep because of itch.

9.1.3.12. PROMIS Itch Questionnaire Scratching Behavior Short Form 5a v1.0

The Scratching Behavior Short Form 5a (PIQ – Scratching Behavior) within the PROMIS Itch item bank consists of 5 items assessing quality of life impairment from scratching behavior and the physical manifestations of itch in adults "in the past 7 days". Response options for the frequency of scratching behaviors range from 1=Never, 2=Rarely, 3=Sometimes; 4=Often, to 5=Almost always. The response options for the worry related to scratching items range from 1=Not at all; 2=A little bit; 3=Somewhat; 4=Quite a bit; to 5=Very much. Total raw scores are converted to T-Scores with higher scores representing more scratching behavior.

9.1.3.13. PROMIS Sleep Related Impairment Short Form 8a v1.0

The Sleep Related Impairment Short Form 8a (PROMIS – Sleep Impairment [PROMIS 2015]) within the PROMIS bank consists of 8 items measuring self-reported perceptions of alertness, sleepiness, and tiredness during usual waking hours, and the perceived functional impairments during wakefulness associated with sleep problems or impaired alertness "in the past 7 days". Response options range from 1=Not at all; 2=A little bit; 3=Somewhat; 4=Quite a bit; to 5=Very much. Total raw scores are converted to T-Scores with higher scores representing greater sleep impairment.

9.1.3.14. Neuro-QoL Cognitive Function Short Form v2.0

Neuro-QoL is a set of self-reported measures that assess the health-related quality of life (HRQOL) of adults and children with neurological disorders and is comprised of item banks and scales that evaluate symptoms, concerns, and issues that are relevant across disorders along with instruments that assess areas most relevant for specific patient populations (NINDS 2015). The Cognitive Function Short Form v2.0 (Neuro-QoL – Cognitive Function) domain within Neuro-QoL bank consists of 8 items measuring Executive Function (perceived difficulties in applications of mental health function related to planning, organizing, calculating, remembering and learning) "in the past 7 days" and General Concerns (perceived difficulties in everyday cognitive abilities such as memory, attention, and decision making) using the lead-in phrase

"how much difficulty do you currently have…". The response options for the Executive Function items range from 1=Very Often (several times a day); 2=Often (once a day); 3=Sometimes (2-3 times); 4=Rarely (once); to 5=Never). The response options for the General Concerns items range from 1=Cannot do; 2=A lot; 3=Somewhat; 4=A little, to 5=None. The total raw scores and are converted to T-Scores with higher scores indicating better (desirable) self-reported health.

9.1.3.15. Patient Benefit Index

The Patient Benefit Index (PBI) measures patient-defined treatment objectives and benefits, particularly in the course of a treatment (Augustin et al. 2009; Blome et al. 2011). It consists of 2 questionnaires. Before therapy, patients complete the standardized "Patient Needs Questionnaire" indicating individual importance of treatment objectives. This reflects their personal preferences with respect to therapeutic benefit. During the study, patients rate the extent to which the treatment objectives have been achieved in the "Patient Benefit Questionnaire". Response options range from 0=not at all, 1=somewhat; 2=moderately; 3=quite; 4=very; to 5=does/did not apply to me. A global score is calculated for each patient by weighing the achievement values of the treatment objectives by their importance to the individual patient. Moreover, the PBI will be supplemented by 6 rating scales assessing the following areas: physical well-being, emotional well-being, performance capacity on the job and in everyday living, social contacts, leisure activities, and quality of life. Patients with PBI $\geq$1 are considered having at least minimum patient-relevant treatment benefit.

9.1.4. Appropriateness of Assessments

All assessments utilized in this study are standard, widely used, and generally recognized as reliable, accurate, and relevant except 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 or stabilizes with appropriate diagnostic evaluation; for events that are not anticipated to resolve or stabilize, the patient should be followed until the treating physician (in consultation with the sponsor) determines that appropriate followup has been completed. 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. In addition, site personnel will record any change in the condition(s) and any new conditions as AEs.

Investigators should record the following via eCRF for each AE: time of onset, time of termination, severity, and their assessment of the potential relatedness of each AE to investigational product.

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 (i.e., 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 (e.g. investigator space portal, telephone, or fax). 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. Investigators can contact the sponsor via telephone at any time using the qualified medical personnel or Lilly affiliate medical contact details which are provided in the site study file.

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

9.2.2. Adverse Events of Special Interest

Adverse events of special interest will include the following:

- infections (including TB, herpes zoster, or opportunistic infections)
- malignancies (except for successfully treated basal or squamous cell carcinoma)
- hepatic events (Section 9.4.9)
- major adverse cardiovascular events (MACE) (Section 9.4.10)
- thrombotic events (such as deep vein thrombosis [DVT] and pulmonary embolism)

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.

Electrocardiograms 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).

9.4.3. Physical Examination

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

The DMC will be authorized to review unblinded results of analyses by treatment group prior to database lock, 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, will be documented in a DMC charter and DMC analysis plan.

Besides DMC members, a limited number of pre-identified individuals may gain access to the limited unblinded data, as specified in the unblinding plan, prior to the final database lock 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 will be specified in a separate unblinding plan document.

10.3.7.2. Adjudication Committee

A blinded Clinical Event Committee will adjudicate potential MACE (cardiovascular death, MI, and 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), venous thrombotic events, and noncardiovascular deaths. Details of membership, operations, recommendations from the Committee, and the communication plan will be documented in the Charter.

5.4. Scientific Rationale for Study Design

This study will enroll moderate to severe AD patients with a history of inadequate response to existing topical therapies for whom a systemic treatment such as baricitinib may therefore be appropriate.

Topical corticosteroids are the first-line anti-inflammatory treatment, even for patients treated with systemic treatments. For this reason, this study will assess the efficacy of baricitinib in combination with background mild-to-moderate potency TCS, used as determined appropriate by the investigator.

During the screening period (Period 1), a washout of systemic and topical treatments for AD was incorporated prior to randomization to minimize confounding effects of patients receiving a wide range of different background treatments prior to study entry as well as potential safety issues related to the use of other therapies, including rebound effects after discontinuation of systemic therapies. 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 16 weeks of treatment.

In consideration of the disease severity, all patients in Study JAIY are eligible for rescue to higher potency 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 (Section 7.7.4).

Investigator's Global Assessments 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). 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 (e.g., moderate to severe). The scale that will be used in this study, the validated Investigator's Global Assessment of Atopic Dermatitis (vIGA-AD, referred to throughout the protocol as IGA), has been developed internally and assesses AD severity using a 5-point scale.

The 16-week efficacy endpoint was chosen because it is likely that a robust clinical effect will be observed with baricitinib within this timeframe based on the Phase 2 study results in AD and from previous studies in another inflammatory skin condition.

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 AD Phase 3 studies are baricitinib 2-mg, and 4-mg QD. These doses were chosen primarily based on the recently completed Phase 2 AD study, JAHG, and are additionally supported by pharmacokinetics (PK), safety, and efficacy data for baricitinib in Phase 2 and Phase 3 RA studies and a Phase 2 psoriasis study.

HADS	Hospital Anxiety Depression Scale
HBcAb	hepatitis B core antibody
HBsAg	hepatitis B surface antigen
HBV	hepatitis B virus
HCV	hepatitis C virus
HIV	human immunodeficiency virus
IB	Investigator's Brochure
ICF	informed consent form
ICH	International Council for Harmonisation
IGA	Investigator's Global Assessment
IL	interleukin
interim analysis	An interim analysis is an analysis of clinical study data, separated into treatment groups, that is conducted before the final reporting database is created/locked.
INR	international normalized ratio
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
LOCF	last observed carried forward
LSM	least squares mean
MACE	major adverse cardiovascular events
MI	myocardial infarction
MMRM	mixed-effects model of repeated measures
Neuro-QoL	Quality of Life in Neurological Disorders
NRI	nonresponder imputation
NRS	Numeric Rating Scale
PBI	Patient Benefit Index
PD	Pharmacodynamic(s)
PDE-4 inhibitor	phosphodiesterase type 4 inhibitor
PIQ	Pain Impact Questionnaire
PK	pharmacokinetic(s)
POEM	Patient-Oriented Eczema Measure
PPD	purified protein derivative
PRO/ePRO	patient-reported outcomes/electronic patient-reported outcomes
PROMIS	Patient-Reported Outcomes Measurement Information System
QD	once daily
QoL	quality of life
RA	rheumatoid arthritis
SAE	serious adverse event
SAP	statistical analysis plan
SCORAD	SCORing Atopic Dermatitis
STAT	signal transducer and activator of transcription
SUSAR	suspected unexpected serious adverse reaction
TB	tuberculosis
TBL	total bilirubin level
TCNI	topical calcineurin inhibitor

The study duration will be up to 25 weeks over 3 study periods:

- Period 1: Screening Period lasting from 8 to 35 days prior to Week 0 (baseline, Visit 2).
- Period 2: Double-Blinded Treatment Period, lasting from Week 0 (baseline, Visit 2) through Week 16 (Visit 8) inclusive:
    - At completion of the double-blind treatment period, eligible patients will be provided the option to participate in the long-term extension study I4V-MC-JAHN (JAHN). Those not eligible or who chose not to participate will proceed to the post-treatment follow-up period.
- Period 3: Post-Treatment Follow-Up Period, spanning the period from the last treatment visit at Week 16 (Visit 8) or Early Termination Visit (ETV) to approximately 4 weeks following the last dose of investigational product.

Treatment Arms and Duration

Patients will be randomized at Week 0 to 1 of 3 treatment groups: placebo once daily (QD), baricitinib 2-mg QD, or baricitinib 4-mg QD. Use of TCS as a background therapy is allowed during the study. The study duration will be up to 25 weeks (Screening Period: up to 5 weeks prior to randomization; Double-Blinded Treatment Period: 16 weeks; Follow-up Period: approximately 4 weeks after the last dose of investigational product).

Number of Patients

This study will include approximately 300 patients with AD who will be randomized in a 1:1:1 ratio to receive placebo QD, baricitinib 2-mg QD or baricitinib 4-mg QD (100 patients in each treatment group).

Statistical Analysis

Unless otherwise specified, the efficacy and health outcome analyses will be conducted on the intent-to-treat population and safety analyses will be conducted on those patients who receive at least 1 dose of investigational product.

Treatment comparisons of discrete efficacy variables will be made using a logistic regression analysis with treatment, baseline disease severity, and region in the model. The proportions and 95% confidence interval (CI) will be reported. If a patient needs to use rescue medication, the data after rescue onward will be considered missing and missing data will be imputed using the nonresponder imputation (NRI) method. All patients who discontinue the study or study treatment at any time for any reason will be defined as nonresponders for the NRI analysis for categorical variables after discontinuation onward. Additional analyses will be done using all observed data whether rescue medication was used or not.

Treatment comparisons of continuous efficacy and health outcome variables will be made using mixed-effects model of repeated measures (MMRM) with treatment, region, baseline severity, visit, and treatment-by-visit interaction as fixed categorical effects and baseline score and baseline score-by-visit interaction as fixed continuous effects. An unstructured covariance matrix will be used to model the within-patient variance–covariance errors. Type III sums of squares for the least squares means (LSMs) will be used for the statistical comparison and contrasts will be set up within the model to compare treatment groups at specific time points of interest.

Fisher exact test will be used for all adverse events (AEs), baseline, discontinuation, and other categorical safety data. Continuous vital signs and laboratory values will be analyzed by an analysis of covariance (ANCOVA) with treatment and baseline values in the model.

	Screening	Double-Blind, Placebo-Controlled							Post-treatment Follow-up
	Period 1	Period 2							Period 3
Visit Number	1	2	3	4	5	6	7	8	801[a]
Weeks from Randomization		0	1	2	4	8	12	16 or ET	
Days from Randomization		0	7	14	28	56	84	112	
Visit Tolerance Interval (days)	-8 to -35		±2	±2	±2	±4	±4	±4	28 ± 4 after last dose
PROMIS – Sleep-Related Impairment[i]	X	X	X	X	X	X	X	X	X
Neuro-QoL – Cognitive Function[j]	X	X						X	
Patient Benefit Index[j]	X	X						X	
C-SSRS[k]/Self-Harm Supplement Form	X	X	X	X	X	X	X	X	X
Self-Harm Follow-up Form[l]	X	X	X	X	X	X	X	X	X
Laboratory Assessment									
Clinical Chemistry[m]	X	X			X	X	X	X	X
Hematology	X	X			X	X	X	X	X
Lipids (fasting)[n]		X					X	X[o]	X
Serum Pregnancy[p]	X								
FSH[q]	X								
TSH	X								
HIV	X								
HCV antibody[r]	X								
HBV testing	X								
HBV DNA[s]	X							X	
Urinalysis	X	X			X	X	X	X	X
Urine Pregnancy[p]		X		X	X	X	X	X	X
Pharmacogenetics: blood		X							
Serum immunoglobulins		X			X			X	
Exploratory storage samples (serum and plasma)		X			X			X	
RNA and biomarkers: blood		X			X			X	

5. Study Design

5.1. Overall Design

Study JAIY is a Phase 3, multicenter, randomized, double-blind, placebo-controlled, parallel-group, outpatient study evaluating the efficacy and safety of baricitinib 2-mg QD and 4-mg QD, in combination with TCS, as compared to placebo in combination with TCS, in adult patients with moderate to severe AD. The study is divided into 3 periods, a 5-week Screening period, a 16-week Double-Blinded Treatment period, and a 4-week Post-Treatment Follow-Up period. For those patients who complete the 16-week treatment period, there is an option to participate in the long-term extension Study JAHN.

Approximately 300 patients ≥18 years of age who have responded inadequately to topical therapy will be randomized in a 1:1:1 ratio to receive placebo QD, baricitinib 2-mg QD, or baricitinib 4-mg QD in combination with TCS (100 patients in each treatment group). Patients will be stratified at randomization according to disease severity (Investigator's Global Assessment [IGA] 3 vs. 4) and geographic region.

All procedures to be conducted during the study, including timing of all procedures, are indicated in the Schedule of Activities (Section 2). Section 9.4.4 describes collection of laboratory samples; Appendix 2, Appendix 4, and Appendix 5 list the specific laboratory tests that will be performed for this study. Laboratory samples listed in Appendix 4, Appendix 5, and Appendix 6 are collected when possible in the event of specific AEs. Study governance considerations are described in detail in Appendix 3. Section 10.3.7.1 outlines information regarding the data monitoring committee (DMC) and interim analyses.

5.1.1. Period 1: Screening

The duration of the Screening Period is between 8 and 35 days prior to Visit 2 (Week 0). At Visit 1, the patient will sign the informed consent form (ICF) prior to any study assessments, examinations, or procedures being performed (Appendix 3). All screening procedures will be performed according to the Schedule of Activities (Section 2). Patients who receive a purified protein derivative (PPD) skin test at Visit 1 will need to return within 48 to 72 hours later to read the skin test. Prior to randomization, treatments for AD will be washed out: 4 weeks for systemic treatments and 2 weeks for topical treatments (not including emollients). Patients will be required to use emollients daily during the 14 days preceding randomization and throughout the study. If patients have been using emollients daily at the time of screening, then those cumulative days can be utilized to meet inclusion criterion [8]. Additionally, collection of data through daily diaries will be required throughout the screening period. The baseline for the daily PRO assessments will be the average score of the 7 days prior to randomization; thus, the minimum screening window was set at 8 days.

All patients who have not previously received the herpes zoster vaccine by screening will be encouraged (per local guidelines) to do so prior to randomization. Refer to the exclusion criterion [28] in Section 6 for additional information regarding herpes zoster vaccinations. In addition, investigators should review the vaccination status of their patients and follow the local

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 chose to discontinue early; however, a separate follow-up visit (V801) is not required.

Patients should not initiate new systemic AD treatment during this period. However, if patients or investigators must initiate treatment, patients should complete an unscheduled visit prior to the first dose of the new therapy.

Figure JAIY.5.1 illustrates the study design. The 3 dosing regimens are described in Section 7.1. The blinding procedure is described in Section 7.3.

6. Study Population

Prospective approval of protocol deviations to recruitment and enrollment criteria, also known as protocol waivers or exemptions, are not permitted.

Study investigator(s) will review patient history and screening test results at Visit 1 and Visit 2 to determine if the patient meets all inclusion and none of the exclusion criteria to qualify for randomization in the study. All screening activities must be completed and reviewed before the patient is randomized.

6.1. Inclusion Criteria

Informed Consent

[1] are at least 18 years of age at the time of informed consent.

Note: Use local requirements to provide consent if the age of adulthood is defined as >18 years

[2] are able to read, understand, and give documented (electronic or paper signature) informed consent.

Type of Patient and Disease Characteristics

[3] have a diagnosis of AD at least 12 months prior to screening, as defined by the American Academy of Dermatology: Guidelines of care for the management of AD; Section 1. Diagnosis and assessment of atopic dermatitis (Appendix 7).

[4] have moderate to severe AD, including all of the following:

a. Eczema Area and Severity Index (EASI) score ≥16 at screening (Visit 1) and at randomization (Visit 2)

b. IGA score of ≥3 at screening (Visit 1) and at randomization (Visit 2)

c. ≥10% of BSA involvement at screening (Visit 1) and at randomization (Visit 2).

[5] have a documented history by a physician and/or investigator of inadequate response to existing topical medications within 6 months preceding screening as defined by at least 1 of the following:

a. inability to achieve good disease control defined as mild disease or better (e.g., IGA ≤2) after use of at least a moderate potency TCS for at least 4 weeks, or for the maximum duration recommended by the product prescribing information (e.g., 14 days for super-potent TCS), whichever is shorter. Topical corticosteroids may be used with or without TCNIs.

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 prior to 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 where 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 0, 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 prior to the study are available. The chest x-ray will be reviewed by the investigator or his or her designee to exclude patients with active 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 at Week 16 (Visit 8) or early termination visit, regardless of their hepatitis B surface antibody (HBsAb) 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 for those patients enrolling in the long-term extension study, JAHN.
- 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 (Section 8.1.2) and should be referred to a hepatology specialist.

11. References

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

Augustin M, Radtke MA, Zschocke I, Blome C, Behechtnejad J, Schäfer I, Reusch M, Mielke V, Rustenbach SJ. The patient benefit index: a novel approach in patient-defined outcomes measurement for skin diseases. *Arch Dermatol Res*. 2009;301(8):561-571.

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 Derm*. 2010;22(2):125-137.

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

Blome C, Augustin M, Behechtnejad J, Rustenbach SJ. Dimensions of patient needs in dermatology: subscales of the patient benefit index. *Arch Dermatol Res.* 2011;303(1):11-17.

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. *Arch Dermatol.* 2004;140:1513-1519.

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: http://www.euroqol.org/fileadmin/user_upload/Documenten/PDF/Folders_Flyers/EQ-5D-5L_UserGuide_2015.pdf. April 2015. Accessed March 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

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
VAS	Visual Analog Scale
ULN	upper limit of normal
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