Protocol (c) I4V-MC-JAIW

A Multicenter, Randomized, Double-Blind, Placebo Controlled, Phase 3 Study to Evaluate the Efficacy and Safety in Adult Patients with Moderate to Severe Atopic Dermatitis

NCT03435081

Approval Date: 10-Oct-2019

Protocol I4V-MC-JAIW(c)
A Multicenter, Randomized, Double-Blind, Placebo-Controlled, Phase 3 Study to Evaluate the Efficacy and Safety of Baricitinib in Adult Patients with Moderate to Severe Atopic Dermatitis

BREEZE-AD5

Confidential Information

The information contained in this document is confidential and is intended for the use of clinical investigators. It is the property of Eli Lilly and Company or its subsidiaries and should not be copied by or distributed to persons not involved in the clinical investigation of baricitinib (LY3009104), unless such persons are bound by a confidentiality agreement with Eli Lilly and Company or its subsidiaries.

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

Baricitinib (LY3009104)

Study I4V-MC-JAIW is a Phase 3, multicenter, randomized, double-blind, placebo-controlled, parallel-group, outpatient, 113-week study designed to evaluate the efficacy and safety of baricitinib 1-mg and 2-mg in patients with moderate to severe atopic dermatitis.

Eli Lilly and Company
Indianapolis, Indiana USA 46285

Protocol Electronically Signed and Approved by Lilly on 13 December 2017.
Amendment (a) Electronically Signed and Approved by Lilly on 6 July 2018.
Amendment (b) Electronically Signed and Approved by Lilly on 1 April 2019.
Amendment (c) Electronically Signed and Approved by Lilly on approval date provided below

Approval Date: 10-Oct-2019 GMT

Table of Contents

Section **Page**

List of Tables

Table **Page**

List of Figures

Figure **Page**

List of Appendices

1. Synopsis

Title of Study:
A Multicenter, Randomized, Double-Blind, Placebo-Controlled, Phase 3 Study to Evaluate the Efficacy and Safety of Baricitinib in Adult Patients with Moderate to Severe Atopic Dermatitis

Rationale:
Atopic dermatitis (AD) is a pruritic, chronic or chronically relapsing, highly symptomatic inflammatory skin disease characterized by excessive T cell activation leading to significant skin infiltration by T cells and dendritic cells (Bieber 2010). Presentation is varied but includes skin manifestations and pruritus, with associated sleep disturbances and subsequent skin infections. The course of disease includes relapses of varying duration and severity.

Baricitinib is an orally available, selective Janus kinase (JAK) inhibitor with potency and selectivity for JAK1 and JAK2 and less potency for JAK3 or tyrosine kinase 2 (TYK2) (Fridman et al. 2010). The pathogenesis of AD is thought to be modulated through thymic stromal lymphopoietin (TSLP), interleukin (IL)-13, IL-4, IL-5, IL-22, and IL-31, many of which activate receptors with downstream signaling through intracellular JAK1/JAK2/TYK2 (Nomura and Kabashima 2015). This activity profile suggests that baricitinib would inhibit many of the cytokines involved in AD pathogenesis.

Clinical studies have established that baricitinib is effective in autoimmune/autoinflammatory diseases involving the joints, kidneys, and skin. Baricitinib was effective at reducing swollen and tender joints in patients with rheumatoid arthritis (RA) (Genovese et al. 2016; Dougados et al. 2017; Fleischmann et al. 2017; Taylor et al. 2017); was effective at reducing disease severity in patients with moderate to severe plaque psoriasis (Papp et al. 2016); was effective at reducing the urinary albumin-to-creatinine ratio in patients with diabetic kidney disease (Tuttle et al. 2015); and in a recently completed Phase 2 study (Study I4V-MC-JAHG) was effective at reducing disease severity in patients with moderate to severe AD. The mechanism of action, combined with demonstration of clinical benefit in inflammatory diseases involving joints, kidneys, and skin, provides the rationale for evaluating baricitinib in moderate to severe AD.

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

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.

The study duration will be up to 113 weeks over 3 study periods:

- Period 1: Screening Period lasting from 8 to 35 days before Week 0 (baseline, Visit 2)
- Period 2: Double-Blinded Treatment Period, lasting from Week 0 (baseline, Visit 2) through Week 104 (Visit 15)
- Period 3: Post-Treatment Follow-Up Period (Visit 801), spanning the period from the last treatment visit at Week 104 (Visit 15) 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 QD, baricitinib 1-mg QD, or baricitinib 2-mg QD. The study duration will be up to 113 weeks (Screening Period: up to 5 weeks before randomization; Double-Blinded Treatment Period: 104 weeks; Post-Treatment Follow-up Period: approximately 4 weeks after the last dose of investigational product).

Number of Patients:

This study will include approximately 450 patients with AD who will be randomized 1:1:1 to placebo QD, baricitinib 1-mg QD, or baricitinib 2-mg QD (150 patients per arm).

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 (IGA) and baseline score in the model. Region may be added to the model if patient numbers allow. The proportions and 95% 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, baseline disease severity (IGA),, visit, and treatment-by-visit interaction as fixed categorical effects and baseline and baseline -by-visit interaction as fixed continuous effects. 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's exact test will be used for all adverse event (AE), 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.

2. Schedule of Activities

Table JAIW.1. Schedule of Activities

	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**
Procedure																
Inclusion and exclusion review	X	X														
Informed consent	X															
Clinical Assessments																
Demographics	X															
Medical history	X															
Substance Use (alcohol, tobacco)	X															
Previous and current AD treatments	X															
Weight	X	X			X	X	X	X	X	X	X	X	X	X	X	X
Height	X															
Body mass index	X	X														
Vital signs (BP and pulse)	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Physical exam	X															
Symptom-directed physical exam[c]		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
12-lead ECG (single)	X															
Chest x-ray[d] (posterior–anterior view)	X															
TB test[e]	X															
Read PPD if applicable (48 to 72 hours after PPD)	X[f]															
Pre-existing conditions	X															
Adverse events		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Concomitant medications	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
ePRO (patient diary) dispensed	X	X	X	X	X	X	X									
ePRO (patient diary) returned		X	X	X	X	X	X	X							X[g]	
Randomization		X														
IWRS	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
IP dispensed		X	X[h]	X[h]	X	X	X	X	X	X	X	X	X	X		

	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

	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

screening.)

[f] If PPD testing was chosen to test for TB, then the patient must return and have her or his PPD test read 48 to 72 hours after Visit 1 (post-PPD).

[g] Applies to ET Visit if conducted prior to Week 16 only.

[h] New bottles of study drug tablets will NOT be dispensed at Weeks 1 and 2 (Visits 3 and 4, respectively) unless the patient requires additional tablets for the next visit window. Patients will continue to take tablets from the bottles dispensed at Week 0 (Visit 2). Compliance will be checked, and the bottle with remaining tablets will be returned to the patient for dosing the following visit window. If a new bottle is required, the patient will not have their previous bottle returned.

[i] Only required if patient met rescue criteria before Week 16 or at investigator discretion after Week 16 for low-potency TCS.

[j] The following measures (Itch NRS, Skin Pain NRS, ADSS, PGI-S-AD, POEM, DLQI, HADS, EQ-5D-5L, WPAI-AD) should be completed before any clinical assessments being performed on days when study visits occur.

[k] Suicidal ideation and behavior subscales excerpt. Adapted for the assessment of 11 preferred ideation and behavior categories.

[l] The Self-Harm Follow-Up Form is required only if triggered by the Self-Harm Supplement Form.

[m] Fasting lipid profile. Patients should not eat or drink anything except water for 12 hours before sample collection. If a patient attends these visits in a nonfasting state, this will not be considered a protocol violation. Unscheduled lipid testing can be performed at the discretion of the investigator.

[n] For ET visits, collect fasting lipids when possible.

[o] Clinical chemistry will include the following value calculated from serum creatinine: estimated glomerular filtration rate (calculated using the Chronic Kidney Disease Epidemiology Collaboration Creatinine 2009 equation).

[p] For all women of childbearing potential, a serum pregnancy test (central laboratory) will be performed at Visit 1. Urine pregnancy tests (local laboratory) will be performed at Visit 2 and at all subsequent study visits after Visit 3. If required per local regulations and/or institutional guidelines, pregnancy testing can occur at other times during the study treatment period.

[q] For female patients aged ≥40 and <60 years who have had a cessation of menses for ≥12 months, an FSH test will be performed to confirm nonchildbearing potential (FSH ≥40 mIU/mL).

[r] For patients who are positive for HCV antibody, a follow-up test for HCV RNA will be performed automatically. Patients who are positive for HCV antibody and negative for HCV RNA may be enrolled.

[s] Patients who are positive for HBcAb and negative for HBV DNA may be enrolled. Any enrolled patient who is HBcAb positive, regardless of HBsAb status or level, must undergo HBV DNA testing per the schedule.

3. Introduction

3.1. Background

AD, also known as eczema or atopic eczema, is a common, chronic, relapsing, highly symptomatic inflammatory skin disease (Bieber 2010). Patients with AD may present with skin lesions that can be acute with oozing, crusted, eroded vesicles or papules on erythematous plaques. Patients may also present with lesions that have a subacute appearance, with thick and excoriated plaques, or chronic appearance, with lichenified, slightly pigmented, excoriated plaques (Bieber 2010). AD causes pruritus attacks throughout the day, which is the primary source of morbidity in this disorder (Simpson 2012). Pruritus often leads to an "itch–scratch" cycle, further compromising the epidermal barrier and resulting in dry skin, microbial colonization, and secondary infections (Krakowski et al. 2008), with 36% of patients reporting that they often or always scratch until their skin bleeds (Langenbruch et al. 2014). Pruritus from AD can worsen at night, resulting in sleep disturbances, with approximately 27% of adult patients with AD experiencing sleep disturbance as a result of itching (Langenbruch et al. 2014). In adult patients with moderate to severe AD, poor sleep quality and latency were significantly associated with poor quality of life (Yano et al. 2013).

In clinical practice, AD is classified as mild, moderate, or severe based on a variety of clinical features, including severity of skin lesions and pruritus and extent of disease (body surface area [BSA] involved).

Until recently, there were no US Food and Drug Administration-approved systemic treatments for patients with moderate to severe AD, with the exception of systemic corticosteroids. In March 2017, Dupixent® (dupilumab) injection, a IgG4 monoclonal antibody that inhibits IL-4 and IL-13, was approved by the US Food and Drug Administration for this patient population. In the European Union, only cyclosporine has been approved for the treatment of patients with severe AD (Bieber and Straeter 2015). A recently completed Phase 2 study (Study I4V-MC-JAHG [JAHG]) evaluated the safety and efficacy of baricitinib (a JAK inhibitor) in AD, and results showed significant improvement in disease severity compared to placebo, and no new safety concerns were identified.

In addition to AD, baricitinib has also been studied in Phase 3 in patients with RA and in Phase 2 in patients with diabetic nephropathy, moderate to severe psoriasis, and systemic lupus erythematosus.

Through 13 February 2019, baricitinib has been studied in more than 548 healthy volunteers and 6555 patients have received baricitinib in clinical studies. As of 13 February 2018, more than 2700 patients have been treated with baricitinib for more than a year and more than 1800 patients have been treated with baricitinib for more than 2 years at doses of 2-mg once daily or 4-mg once daily across the RA clinical program. Baricitinib has been administered as single doses ranging from 1- to 40-mg and as repeat oral doses ranging from 2- to 20-mg to healthy subjects. Baricitinib has also been administered to patients with RA at doses up to 15-mg daily for 4 weeks, 10-mg daily for 24 weeks, 8-mg daily for 76 weeks, and lower doses up to 4-mg daily for up to approximately 7 years.

3.2. Study Rationale

The underlying cause of AD is not completely understood. Loss of function mutations in the gene for *filaggrin* (filament aggregating protein), a key protein in terminal differentiation of the epidermis contributing to barrier function, has been identified as the strongest genetic risk factor for AD in European populations (Palmer et al. 2006). At a cellular level, AD is characterized by excessive T cell activation caused by genetic and environmental factors, leading to significant skin infiltration by T cells and dendritic cells. The cytokine TSLP is thought to act as a master switch that triggers the initiation and maintenance of AD (Moniaga et al. 2013; Ziegler et al. 2013). Overexpression of TSLP in keratinocytes, the most prevalent cell type in the skin, triggers robust itch-evoked scratching and the development of an AD-like skin phenotype in mice (Li et al. 2005). In addition to directly inducing itch by activating sensory neurons in the skin, TSLP also enhances maturation and differentiation of dendritic cells and naive CD4+ T cells and induces production of Th2-related cytokines involved in AD pathogenesis (Wilson et al. 2013; Divekar and Kita 2015). TSLP and other key cytokines involved in AD pathogenesis, such as IL-13, IL-5, IL-22, and IL-31, signal through receptors associated with intracellular JAK1/JAK2/ TYK2 signaling (Ziegler et al. 2013; Nomura and Kabashima 2015).

JAKs are a family of tyrosine kinases that mediate cytokine receptor signalling through phosphorylation and activation of signal transducers and activators of transcription proteins. There are 4 known JAK family members: JAK1, JAK2, JAK3, and TYK2 (Clark et al. 2014). The relative affinity of JAK inhibitors for different members of the JAK kinase family allows for differentiation of JAK inhibitors in relation to their enzymatic inhibitory profile. In vitro assays indicate that baricitinib is a selective inhibitor of JAKs with potency and selectivity for JAK1/2 and less potency for JAK3 or TYK2 (Fridman et al. 2010). The balanced JAK1/JAK2 inhibitory profile of baricitinib suggests that baricitinib will have the greatest modulatory effect in cytokines signalling through a JAK1/JAK2 heterodimer intracellularly (or a JAK1/JAK2/TYK2), such as IL-6, TSLP, IL-13, or IL-31 (Vaddi and Luchi 2012).

The recently completed Phase 2 study of baricitinib in AD, Study JAHG, met its primary objective of proportion of patients achieving a 50% improvement from baseline in Eczema Area and Severity Index (EASI) scores compared to placebo. Baricitinib also showed statistically significant improvements for other disease severity analyses as well as multiple different patient-reported outcome scales compared to placebo, further validating the hypothesis that JAK1/JAK2 signaling plays a key role in AD pathogenesis.

CCI

[redacted]

CCI

[redacted]

4. Objectives and Endpoints

Table JAIW.2 shows the objectives and endpoints of the study.

Table JAIW.2. Objectives and 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

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	

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.

5. Study Design

5.1. Overall Design

Study JAIW is a Phase 3, multicenter, randomized, double-blind, placebo-controlled, parallel-group, outpatient study evaluating the efficacy and safety of baricitinib 1-mg QD and 2-mg QD as compared to placebo in adult patients with moderate to severe AD. The study is divided into 3 periods: a 5-week Screening Period, a 104-week Double-Blinded Treatment Period, and a 4-week Post-Treatment Follow-Up Period.

Approximately 450 patients aged ≥18 years who have responded inadequately to or who are intolerant of topical therapy will be randomized at a 1:1:1 ratio to receive placebo QD, baricitinib 1-mg QD, or baricitinib 2-mg QD (approximately 150 patients per group). Patients will be stratified at randomization according to disease severity (Investigator's Global Assessment [IGA] 3 versus 4) and geographic region if the planned country allocation justifies.

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, Appendix 5, and Appendix 6 list the specific laboratory tests that will be performed for this study. Study governance considerations are described in detail in Appendix 3. Section 10.3.6.1 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 before Visit 2 (Week 0). At Visit 1, the patient will sign the informed consent form (ICF) before any study assessments, examinations, or procedures are performed (see 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 have the skin test read. Before 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 patient-reported outcome assessments will be the average score of the 7 days before randomization; thus, the minimum screening window was set at 8 days.

All patients who are eligible to receive the herpes zoster vaccine (per local guidelines) and who have not received it by screening will be encouraged to do so before randomization. Investigators should review the vaccination status of their patients and follow the local guidelines for vaccination of those aged ≥18 years with nonlive vaccines intended to prevent infectious disease before entering patients into the study. Refer to Exclusion Criterion [27] in Section 6.2 for additional information regarding herpes zoster vaccinations.

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.

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.

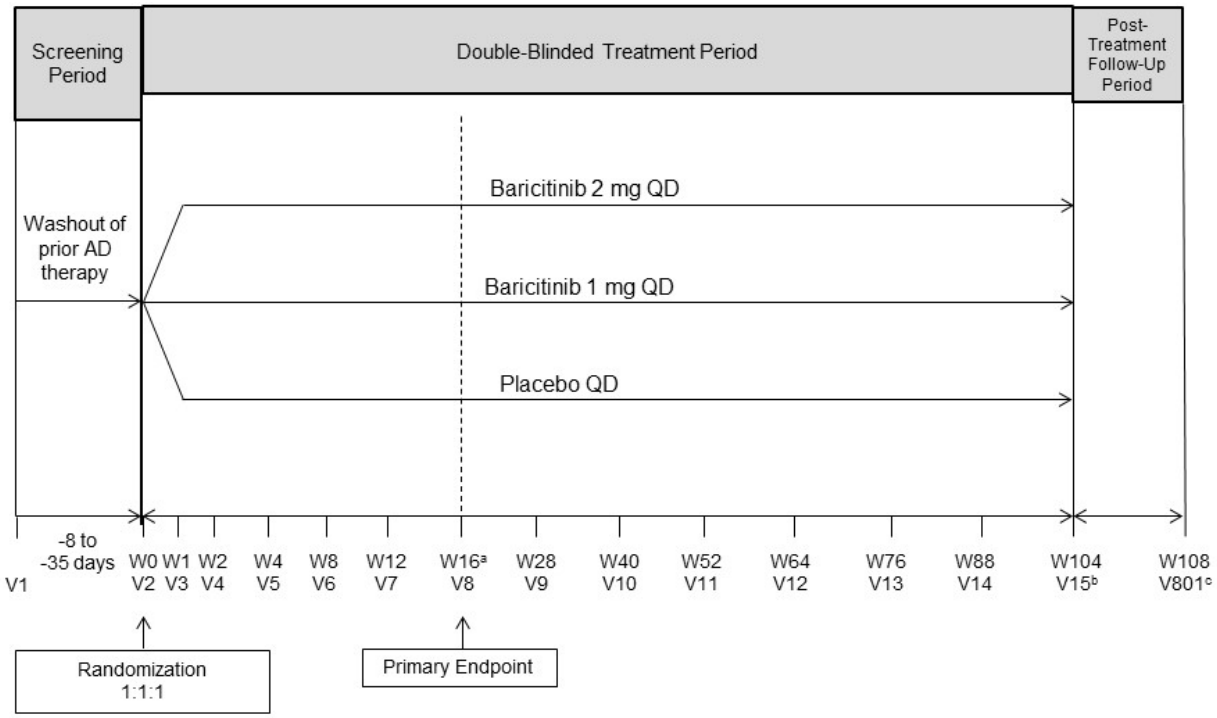
Patients who have received at least 1 dose of investigational product and discontinue early from the study must have an ETV and return for the post-treatment safety follow-up visit (Visit 801) approximately 28 days after the last dose of investigational product.

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.



Abbreviations: AD = atopic dermatitis; QD = once daily; V = visit; W = week.

[a] At Week 16, all patients who achieve an IGA 0 or 1 and who have not required rescue therapy before Week 16 will be allowed to continue in this study. All other patients will be discontinued from this study and may be eligible to enroll in the separate open-label Study JAIX.

[b] Patients who complete this study will be eligible for assessment to enroll in open-label Study JAIX.

[c] Occurs approximately 28 days after the last dose of investigational product.

Figure JAIW.1. Illustration of study design for Clinical Protocol I4V-MC-JAIW.

5.2. Number of Participants

Approximately 450 participants will be enrolled; approximately 750 patients will be screened to achieve this enrollment.

5.3. End of Study Definition

End of the study is the date of the last visit or last scheduled procedure shown in the Schedule of Activities (Section 2) for the last patient.

5.4. Scientific Rationale for Study Design

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

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,

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

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 criteria 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 aged ≥18 years 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 before screening, as defined by the American Academy of Dermatology: Criteria for the Diagnosis and Assessment of Atopic Dermatitis (see Appendix 7).

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

a. 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 of screening, or history of intolerance to topical therapy as defined by at least 1 of the following:

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

b. documented history of clinically significant adverse reactions with the use of TCS such as skin atrophy, allergic reactions, or systemic effects that in the opinion of the investigator outweigh the benefits of retreatment.

c. patients who failed systemic therapies intended to treat AD (such as cyclosporine, methotrexate, azathioprine, or mycophenolate mofetil) within 6 months of screening will also be considered as surrogates for having inadequate response to topical therapy.

[6] agree to discontinue use of the following excluded medications/treatments for at least 4 weeks before randomization (Visit 2):

a. oral systemic corticosteroids and leukotriene inhibitors

b. systemic immunomodulators, including but not limited to cyclosporine, methotrexate, mycophenolate mofetil, and azathioprine

c. sedating antihistamines, including but not limited to alimemazine, chlorphenamine, clemastine, cyproheptadine, diphenhydramine, hydroxyzine, ketotifen, and promethazine

Note: Patients may use newer, less sedating antihistamines (for example, fexofenadine, loratadine, cetirizine).

d. any other systemic therapy used to treat AD or symptoms of AD (approved or off-label use)

e. phototherapy, includes therapeutic phototherapy (psoralen plus ultraviolet-A, ultraviolet-B), excimer laser, as well as self-treatment with tanning beds.

[7] agree to discontinue use of the following excluded medications for at least 2 weeks before randomization (Visit 2):

a. TCS or topical immune modulators (for example, tacrolimus or pimecrolimus)

b. Topical PDE-4 inhibitor (crisaborole).

[8] have applied emollients daily for at least 14 days before randomization and agree to use emollient daily throughout the treatment period.

[9] Patients who are receiving chronic treatments to improve sleep should be on a stable dose for at least 2 weeks before screening as determined by the investigator. Sedating antihistamines (see above) are not permitted.

Patient Characteristics

[10] are male or nonpregnant, nonbreastfeeding female patients, and

a. Male patients will either remain abstinent (if this is their preferred and usual lifestyle) or agree to use 2 forms of birth control (1 must be highly effective, see below) while engaging in sexual intercourse with female partners of childbearing potential while enrolled in the study and for at least 4 weeks following the last dose of investigational product.

Men who are in exclusively same sex relationships (when it is their preferred and usual lifestyle) are not required to use contraception.

b. Female patients of childbearing potential who are abstinent (if this is complete abstinence, as their preferred and usual lifestyle) or in a same-sex relationship (as part of their preferred and usual lifestyle) must agree to either remain abstinent or stay in a same-sex relationship without sexual relationships with males. Periodic abstinence (for example, calendar, ovulation, symptothermal, post-ovulation methods), declaration of abstinence just for the duration of a trial, and withdrawal are not acceptable methods of contraception.

   Otherwise, female patients of childbearing potential must agree to use 2 forms of birth control when engaging in sexual intercourse with a male partner while enrolled in the study and for at least 4 weeks following the last dose of investigational product.

   The following birth control methods are considered acceptable (the patient should choose 2 to be used with their male partner, and 1 must be highly effective):

   - Highly effective birth control methods: oral, injectable, or implanted hormonal contraceptives (combined estrogen/progesterone or progesterone only, associated with inhibition of ovulation); intrauterine device or intrauterine system (for example, progestin-releasing coil); or, vasectomized male (with appropriate post-vasectomy documentation of the absence of sperm in the ejaculate).
   - Effective birth control methods: condom with a spermicidal foam, gel, film, cream, or suppository; occlusive cap (diaphragm or cervical/vault caps) with a spermicidal foam, gel, film, cream, or suppository; or, oral hormonal contraceptives.

     Note: When local guidelines concerning highly effective or effective methods of birth control differ from the above, the local guidelines must be followed.

c. Females of nonchildbearing potential are not required to use birth control. They are defined as:

   - women aged ≥60 years or women who are congenitally sterile, or
   - women aged ≥40 and <60 years who have had a cessation of menses for ≥12 months and a follicle-stimulating hormone test confirming nonchildbearing potential (≥40 mIU/mL or ≥40 IU/L), or women who are surgically sterile (that is, have had a hysterectomy or bilateral oophorectomy or tubal ligation).

6.2. Exclusion Criteria

Patients will be excluded from study enrollment if they meet any of the following criteria:

Medical Conditions Related to Atopic Dermatitis

[11] are currently experiencing or have a history of other concomitant skin conditions (for example, psoriasis or lupus erythematosus) that would interfere with evaluations of the effect of study medication on AD.

[12] patients who are currently experiencing or have a history of erythrodermic, refractory, or unstable skin disease that requires frequent hospitalizations and/or intravenous treatment for skin infections that, in the opinion of the investigator, may interfere with participation in the study.

[13] a history of eczema herpeticum within 12 months of screening.

[14] a history of 2 or more episodes of eczema herpeticum.

[15] patients who are currently experiencing a skin infection that requires treatment or is currently being treated with topical or systemic antibiotics.

Note: Patients may not be rescreened until at least 4 weeks after the date of their screen failure and at least 2 weeks after resolution of the infection.

[16] have any serious concomitant illness that is anticipated to require the use of systemic corticosteroids or otherwise interfere with study participation or require frequent monitoring (for example, unstable chronic asthma).

[17] have been treated with the following therapies:

a. monoclonal antibody (for example, ustekinumab, omalizumab, dupilumab) for less than 5 half-lives before randomization.

b. received prior treatment with any oral JAK inhibitor (for example, tofacitinib, ruxolitinib) less than 4 weeks before randomization.

c. received any parenteral corticosteroid administered by intramuscular or intravenous injection within 6 weeks of planned randomization (Visit 2) or are anticipated to require parenteral injection of corticosteroids during the study.

d. have had an intra-articular corticosteroid injection within 6 weeks of planned randomization (Visit 2).

Note: Intranasal or inhaled steroid use is allowed during the trial.

e. probenecid at the time of randomization (Visit 2) that cannot be discontinued for the duration of the study.

Medical Conditions in General

[18] are largely or wholly incapacitated permitting little or no self-care, such as being bedridden.

[19] have uncontrolled arterial hypertension characterized by a repeated systolic blood pressure >160 mm Hg or diastolic blood pressure >100 mm Hg in a seated position.

[20] have had any major surgery within 8 weeks of screening or will require major surgery during the study that, in the opinion of the investigator in consultation with Lilly or its designee, would pose an unacceptable risk to the patient.

[21] are immunocompromised and, in the opinion of the investigator, at an unacceptable risk for participating in the study.

[22] have experienced any of the following within 12 weeks of screening: myocardial infarction (MI), unstable ischemic heart disease, stroke, or New York Heart Association Stage III/IV heart failure.

[23] have a history of VTE, or are considered at high risk for VTE as deemed by the investigator, or have 2 or more of the following risk factors for VTE:

a. Aged >65 years.

b. BMI >35 kg/m^2.

c. Oral contraceptive use and current smoker.

[24] have a history or presence of cardiovascular, respiratory, hepatic, gastrointestinal, endocrine, hematological, neurological, or neuropsychiatric disorders or any other serious and/or unstable illness that, in the opinion of the investigator, could constitute an unacceptable risk when taking investigational product or interfere with the interpretation of data.

[25] have a history of lymphoproliferative disease; or have signs or symptoms suggestive of possible lymphoproliferative disease, including lymphadenopathy or splenomegaly; or have active primary or recurrent malignant disease; or have been in remission from clinically significant malignancy for <5 years.

a. Patients with cervical carcinoma in situ that has been appropriately treated with no evidence of recurrence or metastatic disease for at least 3 years may participate in the study.

b. Patients with basal cell or squamous epithelial skin cancers that have been appropriately treated with no evidence of recurrence for at least 3 years may participate in the study.

[26] have a current or recent clinically serious viral, bacterial, fungal, or parasitic infection, including but not limited to the following:

Note: A recent viral upper respiratory tract infection or uncomplicated urinary tract infection should not be considered clinically serious.

a. symptomatic herpes zoster infection within 12 weeks of screening.

b. a history of disseminated/complicated herpes zoster (for example, multidermatomal involvement, ophthalmic zoster, central nervous system involvement, or post-herpetic neuralgia).

c. symptomatic herpes simplex at the time of randomization.

d. active or chronic viral infection from hepatitis B virus (HBV), hepatitis C virus (HCV), or human immunodeficiency virus (HIV).

e. household contact with a person with active tuberculosis (TB) and have not received appropriate and documented prophylaxis for TB.

f. evidence of active TB or have previously had evidence of active TB and have not received appropriate and documented treatment.

g. clinically serious infection or received intravenous antibiotics for an infection, within 4 weeks of randomization.

h. any other active or recent infection within 4 weeks of randomization that, in the opinion of the investigator, would pose an unacceptable risk to the patient if participating in the study.

[27] have been exposed to a live vaccine within 12 weeks of planned randomization or are expected to need/receive a live vaccine during the course of the study (with the exception of herpes zoster vaccination).

Note: All patients who are eligible to receive the herpes zoster vaccine (per local guidelines) and who have not received it by screening will be encouraged to do so before randomization; vaccination must occur >4 weeks before randomization and first dose of investigational product. Patients will be excluded if they were exposed to herpes zoster vaccination within 4 weeks of planned randomization.

[28] have a history of chronic alcohol abuse, IV drug abuse, or other illicit drug abuse within 2 years of screening.

[29] presence of significant uncontrolled neuropsychiatric disorder, are clinically judged by the investigator to be at risk for suicide, or have a "yes" answer to any of the following within the 2 months before Visit 1:

a. Question 4 (Active Suicidal Ideation with Some Intent to Act, Without Specific Plan) on the "Suicidal Ideation" portion of the Columbia Suicide Severity Rating Scale (C-SSRS) or

b. Question 5 (Active Suicidal Ideation with Specific Plan and Intent) on the "Suicidal Ideation" portion of the C-SSRS or

c. Any of the suicide-related behaviors (actual attempt, interrupted attempt, aborted attempt, preparatory act or behavior) on the "Suicidal Behavior" portion of the C-SSRS.

Note: A patient does not necessarily have to be excluded if they have self-injurious behavior that would be classified as nonsuicidal self-injurious behavior. If this situation arises, the subject should be referred to a psychiatrist or appropriately trained professional as indicated.

[30] have donated more than a single unit of blood within 4 weeks of screening or intend to donate blood during the course of the study.

Other Exclusions

[31] are unable or unwilling to make themselves available for the duration of the study and/or are unwilling to follow study restrictions/procedures, including use of data collection devices.

[32] are currently enrolled in any other clinical trial involving an investigational product or any other type of medical research judged not to be scientifically or medically compatible with this study.

[33] have participated in a clinical study involving an investigational product within the last 30 days. If the previous investigational product has a long half-life (2 weeks or longer), at least 3 months or 5 half-lives (whichever is longer) must have passed.

[34] have previously been randomized in this study or any other study investigating baricitinib.

[35] are investigator site personnel directly affiliated with this study and/or their immediate families. Immediate family is defined as a spouse, parent, child, or sibling, whether biological or legally adopted.

[36] are Lilly or Incyte employees or their designee.

Diagnostic Assessments

[37] have screening electrocardiogram (ECG) abnormalities that, in the opinion of the investigator, are clinically significant and indicate an unacceptable risk for the patient's participation in the study.

[38] have evidence of active TB or latent TB

a. have evidence of active TB, defined in this study as the following:

- documented by a positive PPD test (≥5 mm induration between approximately 48 and 72 hours after application, regardless of vaccination history), medical history, clinical features, and abnormal chest x-ray at screening.
- The QuantiFERON-TB Gold test or T-SPOT TB test (as available and if compliant with local TB guidelines) may be used instead of the PPD test. Patients are excluded from the study if the test is not negative and there is clinical evidence of active TB.

Exception: Patients with a history of active 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 for PPD, QuantiFERON-TB Gold test, or T-SPOT TB test but must have a chest x-ray at screening.

b. have evidence of untreated/inadequately or inappropriately treated latent TB, defined in this study as the following:

- documented to have a positive PPD test (≥5 mm induration between approximately 48 and 72 hours after application, regardless of vaccination history), no clinical features consistent with active TB, and a chest x-ray with no evidence of active TB at screening; or
- PPD test is positive, and if the patient has no medical history or chest x-ray findings consistent with active TB, the patient may have a QuantiFERON-TB Gold test or T-SPOT TB test (as available and if compliant with local TB guidelines). If the test results are not negative, the patient will be considered to have latent TB (for purposes of this study); or
- QuantiFERON-TB Gold test or T-SPOT TB test (as available and if compliant with local TB guidelines) may be used instead of the PPD test. If the test results are positive, the patient will be considered to have latent TB. If the test is not negative, the test may be repeated once within approximately 2 weeks of the initial value. If the repeat test results are again not negative, the patient will be considered to have latent TB (for purposes of this study).

Exception: Patients who have evidence of latent TB may be enrolled if he or she completes at least 4 weeks of appropriate treatment before randomization and agrees to complete the remainder of treatment while in the trial.

Exception: Patients with a history of 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 for PPD, QuantiFERON-TB Gold test, or T-SPOT TB test but must have a chest x-ray at screening.

[39] have a positive test for HBV infection, defined as:

a. positive for hepatitis B surface antigen, or

b. positive for hepatitis B core antibody (HBcAb) and positive HBV DNA.

Note: Patients who are HBcAb positive and HBV DNA negative may be enrolled in the study. Patients who meet these criteria at screening will be identified by the central laboratory and monitored during the study.

[40] have HCV infection (positive for anti-hepatitis C antibody with confirmed presence of HCV RNA.

Note: Patients who have documented anti-HCV treatment for a past HCV infection AND are HCV RNA negative may be enrolled in the study.

[41] have evidence of HIV infection and/or positive HIV antibodies.

[42] have screening laboratory test values, including thyroid-stimulating hormone (TSH), outside the reference range for the population or investigative site that, in the opinion of the investigator, pose an unacceptable risk for the patient's participation in the study.

Note: Patients who are receiving thyroxine as replacement therapy may participate in the study provided stable therapy has been administered for ≥12 weeks and TSH is within the laboratory's reference range. Patients who are receiving stable thyroxine replacement therapy who have TSH marginally outside the laboratory's normal reference range may participate if the treating physician has documented that the thyroxine replacement therapy is adequate for the patient.

[43] have any of the following specific abnormalities on screening laboratory tests:

a. AST or ALT ≥2 × upper limit of normal (ULN)

b. alkaline phosphatase (ALP) ≥2 × ULN

c. total bilirubin ≥1.5 × ULN

d. hemoglobin <10.0 g/dL (100.0 g/L)

e. total white blood cell count <2500 cells/μL (<2.50 × 10^3/μL or <2.50 GI/L)

f. neutropenia (absolute neutrophil count <1200 cells/μL) (<1.20 × 10^3/μL or <1.20 GI/L)

g. lymphopenia (lymphocyte count <750 cells/μL) (<0.75 × 10^3/μL or <0.75 GI/L)

h. thrombocytopenia (platelets <100,000/μL) (<100 × 10^3/μL or <100 GI/L)

i. estimated glomerular filtration rate <60 mL/min/1.73 m^2 (Chronic Kidney Disease Epidemiology Collaboration equation Creatinine 2009 equation).

Note: For cases with any of the aforementioned laboratory abnormalities (Exclusion Criteria [42] and [43]), the tests may be repeated during screening, and values resulting from repeat testing may be accepted for enrollment eligibility if they meet the eligibility criterion.

6.3. Lifestyle Restrictions

Not applicable.

6.4. Screen Failures

Patients who are entered into the study but do not meet the eligibility criteria for participation in this study (screen failure) may be rescreened a maximum of 2 times. If patients are rescreened, rescreening cannot occur until at least 4 weeks after the date of their previous screen failure. When rescreening is performed, the individual must sign a new ICF and will be assigned a new identification number. Additionally, all necessary screening procedures must be conducted at rescreen to ensure all eligibility criteria are met.

7. Treatments

7.1. Treatments Administered

This study involves a comparison of placebo, baricitinib 1-mg, and baricitinib 2-mg administered orally QD. Table JAIW.3 shows the treatment regimens.

Table JAIW.3. Treatment Regimens

Regimen	Investigational Product Supplied	Dose
Baricitinib 2–mg QD	Baricitinib 2-mg tablets Placebo to match 1-mg tablets	2 tablets per day
Baricitinib 1–mg QD	Baricitinib 1-mg tablets Placebo to match 2-mg tablets	2 tablets per day
Placebo QD	Placebo to match 2-mg tablets Placebo to match 1-mg tablets	2 tablets per day

Abbreviation: QD = once daily.

The investigator or his or her designee is responsible for the following:

- explaining the correct use of the investigational agent(s) to the patient
- verifying that instructions are followed properly
- maintaining accurate records of investigational product dispensing and collection
- returning all unused medication to Lilly or its designee at the end of the study, unless the sponsor and sites have agreed all unused medication is to be destroyed by the site, as allowed by local law

7.1.1. Packaging and Labeling

The sponsor (or its designee) will provide the following investigational products:

- tablets containing 2-mg of baricitinib (Bottle A)
- tablets containing 1-mg of baricitinib (Bottle B)
- placebo tablets to match baricitinib 2-mg tablets (Bottle A)
- placebo tablets to match baricitinib 1-mg tablets (Bottle B)

Baricitinib 1-mg and 2-mg tablets and packaging will be identical in appearance to the respective placebo tablets to maintain the blind.

Investigational product tablets will be provided in bottles. Clinical trial materials will be labeled according to the country's regulatory requirements, and the double-blind label will include the designation A or B on the front panel. **Patients will be instructed to take 1 tablet from bottle A and 1 tablet from bottle B each day for blinded dosing.**

7.2. Method of Treatment Assignment

Patients who meet all criteria for enrollment will be randomized in a 1:1:1 ratio (placebo, baricitinib 1-mg, or baricitinib 2-mg) to double-blind treatment at Visit 2 (Week 0). Randomization will be stratified by geographic region (North America, rest-of-world) if the planned country allocation justifies and disease severity at baseline (IGA 3 versus 4). Assignment to treatment groups will be determined by a computer-generated random sequence using an interactive web-response system (IWRS). The IWRS will be used to assign bottles, each containing double-blind investigational product tablets to each patient, starting at Visit 2 (Week 0), and at visits per the Schedule of Activities (Section 2), up to and including Visit 14 (Week 88). Site personnel will confirm that they have located the correct bottles by entering a confirmation number found on the bottle into the IWRS.

7.2.1. Selection and Timing of Doses

The investigational product (1 tablet from Bottle A and 1 tablet from Bottle B) should be taken QD without regard to food and, if possible, at approximately the same time every day, usually at the start of the patient's day, to aid patient compliance.

7.3. Blinding

This is a double-blind study. To preserve the blinding of the study, a minimum number of Lilly personnel will see the randomization table and treatment assignments before the study is complete. All study assessments will be performed by study personnel who are blinded to the patient's treatment group. Except in clinical circumstances where unblinding is required, the patients, investigators, Lilly study team, and any personnel interacting directly with patients or investigative sites will remain blinded to baricitinib and placebo assignment. Every effort should be made to preserve the blind unless there is a compelling reason that knowledge of the specific treatment would alter the medical care of the patient.

In case of an emergency, the investigator has the sole responsibility for determining if unblinding of a patient's treatment assignment is warranted for medical management of the event. The patient's safety must always be the first consideration in making such a determination. If a patient's treatment assignment is unblinded, Lilly must be notified immediately. If the investigator decides that unblinding is warranted, it is the responsibility of the investigator to promptly document the decision and rationale and notify Lilly as soon as possible.

Emergency unblinding for AEs may be performed through the IWRS. This option may be used ONLY if the patient's well-being requires knowledge of the patient's treatment assignment. All unblinding events are recorded and reported by the IWRS. If an investigator, site personnel performing assessments, or patient is unblinded, the patient must be discontinued from the study. In cases where there are ethical reasons to have the patient remain in the study, the investigator must obtain specific approval from a Lilly clinical research physician for the patient to continue in the study.

Processes to maintain blinding during the interim analysis conducted by the DMC are described in Section 10.3.6.1.

7.4. Dosage Modification

Not applicable.

7.5. Preparation/Handling/Storage/Accountability

All investigational product (used and partially used) will be returned to the sponsor or destroyed at site level with the sponsor's written approval. In some cases, sites may destroy the material if, during the investigative site selection, the evaluator has verified and documented that the site has appropriate facilities and written procedures to dispose of clinical trial materials.

Follow storage and handling instructions on the investigational product packaging.

7.6. Treatment Compliance

Patient compliance with study medication will be assessed at each visit during the treatment period (Visit 3 through Visit 15) by counting returned tablets.

A patient will be considered significantly noncompliant if he or she misses more than 20% of the prescribed doses of investigational product during the study, unless the patient's investigational product is withheld by the investigator for safety reasons. Similarly, a patient will be considered significantly noncompliant if he or she is judged by the investigator to have intentionally or repeatedly taken 20% more than the prescribed amount of medication during the study.

Patients will be counseled by study staff on the importance of taking the investigational product as prescribed, as appropriate.

Patients' compliance will be further defined in the statistical analysis plan (SAP).

7.7. Concomitant Therapy

All concomitant medication, whether prescription or over the counter, used at baseline and/or during the course of the study must be recorded on the Concomitant Medication electronic case report form (eCRF). Patients will be instructed to consult the investigator or other appropriate study personnel at the site before taking any new medications or supplements during the study. For AD therapies permitted as part of rescue therapy, see Section 7.7.3.

7.7.1. Prohibited Medications and Procedures

Prohibited Medications and Procedures Not Requiring Interruption of Investigational Product

The following therapies will not be allowed during the study, unless otherwise specified. Some therapies may be used after completion of Visit 8 (Week 16), as outlined below.

- TCS, TCNIs (for example, tacrolimus and pimecrolimus), or topical PDE-4 inhibitor (that is, crisaborole) are not allowed, except when given as rescue therapy as described in Section 7.7.3. Low-potency TCS is allowed after Week 16 (see Section 7.7.2).

- Topical antihistamines or sedating, systemic antihistamines including but not limited to alimemazine, chlorphenamine, clemastine, cyproheptadine, diphenhydramine, hydroxyzine, ketotifen, and promethazine are not allowed before Week 16 (Visit 8).
- Leukotriene inhibitors (for example, montelukast [Singulair®], zafirlukast [Accolate], and zileuton [Zyflo®]) are not allowed before Week 16 (Visit 8).
- Allergen immunotherapy is not allowed in the study.
- Phototherapy, including PUVA (psoralen and ultraviolet A), ultraviolet B, tanning booth, and excimer laser, is not allowed in the study.
- Bleach baths are not allowed in the study.

Prohibited Medications Requiring Temporary Interruption of Investigational Product

The following therapies will not be allowed during the course of the study and, if taken by or administered to the patient, temporary interruption of investigational product is required.

- Live vaccines (including Bacillus Calmette-Guérin or herpes zoster), (see Exclusion Criterion [27])
    - For Bacillus Calmette-Guérin vaccination, investigational product should be temporarily interrupted for 12 weeks.
    - For herpes zoster vaccination, investigational product should be temporarily interrupted for 4 weeks.
- Probenecid: If a patient is inadvertently started on probenecid, the investigational product should be temporarily interrupted and can be resumed after patient has discontinued probenecid. If a patient is not able to discontinue probenecid, then the investigational product should be permanently discontinued.
- Systemic corticosteroids may be used for the treatment of an AE (for example, worsening of existing condition, such as asthma flare) after Visit 8 (Week 16). Investigational product may be restarted if systemic corticosteroids were used for a short duration (<30 days). If used for >30 days, sponsor approval to restart investigational product is required.

Prohibited Medications Requiring Permanent Discontinuation of Investigational Product

- Systemic corticosteroids are not allowed before Visit 8 (Week 16), and if used it will be considered rescue therapy, and investigational product should be discontinued (see Section 7.7.3). Systemic corticosteroids may not be used to treat AD, before or after Week 16.
- Any systemic therapy, investigational or commercial (approved or off-label use), used for the treatment of AD or symptoms of AD (except for antihistamines, as specified above) is not allowed.
- Other JAK inhibitors (for example, tofacitinib and ruxolitinib) are not allowed.

- Systemic immunosuppressive/immunomodulatory substances, including but not limited to cyclosporine, methotrexate, mycophenolate mofetil, interferon γ, azathioprine, or biologic agents are not allowed.

7.7.2. Permitted Medications and Procedures

Treatment with concomitant AD therapies during the study is permitted only as described below.

- Daily use of emollients is required as background treatment. Moisturizers with additives such as antipruritics or antiseptics are not permitted. If daily applications are missed, it will not be considered a protocol violation.
    - Patients should not apply emollients on the day of their study visit before study procedures to allow adequate assessment of skin dryness.
- For those patients on stable dosing of prescription sleep medications at entry, downward dose adjustments or discontinuation of treatment may occur during the study.
- Nonsedating antihistamines including but not limited to acrivastine, bilastine, cetirizine, desloratadine, fexofenadine, levocetirizine, loratadine, mizolastine, and rupatadine are allowed.
- Intra-articular or soft tissue (bursa, tendons, and ligaments) corticosteroid injection: No more than 1 intra-articular or soft tissue (bursa, tendons, and ligaments) corticosteroid injection is allowed up until Week 16 (Visit 8). After Week 16, such injections are permitted.
- Intranasal or inhaled steroid use is allowed.
- Topical anesthetics and topical and systemic anti-infective medications are allowed.
- Nonlive seasonal vaccinations and/or emergency vaccination, such as rabies or tetanus vaccinations, are allowed.
- Low-potency TCS (for example, hydrocortisone 2.5% ointment) is permitted after Week 16 (Visit 8).

Any changes of these concomitant medications must be recorded in the Concomitant Therapy of Special Interest eCRF.

Treatment with concomitant therapies for other medical conditions such as diabetes and hypertension is permitted during the study.

7.7.3. Rescue Therapy

Rescue therapy is only available during the first 16 weeks of treatment.

Criteria for Rescue Therapy Initiation

Investigators should attempt to manage patients with emollients; however, investigators are allowed to rescue patients who are experiencing unacceptable or worsening symptoms of AD at

any time before primary endpoint assessment at Visit 8 (Week 16). Before rescue, it is recommended that increased frequency of emollient use is attempted to at least twice a day or more in an effort to control symptoms. The rationale for rescue will be documented. All rescued patients will be considered nonresponders at Week 16 and will be discontinued from the study. Patients may be eligible to enroll in the open-label extension.

After Week 16 patients are allowed to use low-potency TCS if needed as described below.

Choice of Rescue Therapy Treatment before Week 16

- Triamcinolone 0.1% cream and/or hydrocortisone 2.5% ointment. Where possible, both of these treatments will be supplied by the sponsor. In the event where providing 1 or both of these topical formulations is not possible, an alternate, equivalent-potency TCS cream and/or ointment may be provided by the sponsor. TCS use, supplied by the sponsor, should be recorded via weight of returned tubes as indicated in the Schedule of Activities (Section 2).
    - In the event that the sponsor is unable to supply TCS, commercially available triamcinolone 0.1% cream and/or hydrocortisone 2.5% ointment may be supplied by the sites. Where providing triamcinolone 0.1% cream and/or hydrocortisone 2.5% ointment is not possible, an equivalent-potency TCS cream and/or ointment that is in line with local practices can be supplied. Refer to Appendix 8 for guidance on potency equivalence.
    - If the TCS supplied by the sponsor is not considered suitable for an individual patient, an equivalent-potency TCS cream and/or ointment that is in line with local practices can be supplied by the sites. Refer to Appendix 8 for guidance on potency equivalence.
- Investigators may also elect to use TCNIs and/or crisaborole where approved, with or without TCS. If TCNIs are prescribed, use should be limited to problem areas only (for example, face, neck, skin folds, genital areas).
- On the days of study visits, topical therapy should not be applied before the patient has undergone all study procedures and clinical evaluations to allow adequate assessment of skin dryness.
- Patients rescued to topical therapy will continue to take investigational product, and use of rescue therapy will be documented in the eCRF.

During the first 16 weeks of the treatment period, in patients who do not improve sufficiently with the provided rescue topical therapy after 7 days, a higher-potency TCS may be used (see Appendix 8), and investigational product may continue. It is recommended that if a patient reaches “clear” to “almost clear” skin after topical rescue, then medium- and/or high-potency TCS and TCNI should be stopped, and low-potency TCS (for example, hydrocortisone 2.5% ointment) should be used QD for an additional 7 days, then stopped. If lesions return, patients can be retreated with TCS with or without TCNIs and/or crisaborole as before, at the discretion of the investigator.

If topical rescue therapy as described above fails to sufficiently control AD symptoms, then oral systemic medications may be used as rescue (for example, corticosteroids, cyclosporine, methotrexate); however, investigational product will be required to be permanently discontinued, and the patient will be discontinued from the study after assessment of the primary endpoint at Week 16.

Investigators should make every attempt to conduct efficacy and safety assessments immediately before administering any rescue treatment. An unscheduled visit can be used for this purpose if necessary. All patients requiring rescue during the first 16 weeks of the study will be considered nonresponders for the primary endpoint at Visit 8 (Week 16), will have to be discontinued from the study at Week 16, and may be eligible to enroll into the open-label Study JAIX.

Rescue Therapy after Week 16

No rescue therapies are allowed after Week 16.

Only patients who achieved an IGA score of 0 or 1 at Week 16 will remain in the study beyond Week 16. After Week 16, if patients experience worsening of their AD the use of low-potency TCS (for example, hydrocortisone 2.5% ointment) is permitted as concomitant therapy. If after starting low-potency TCS a patient reaches clear or almost clear skin, TCS should be discontinued. If lesions return, patients can be retreated with low-potency TCS.

Patients requiring more than low-potency TCS will be considered nonresponders and will be discontinued from the study. The use of TCNIs or crisaborole is not allowed after Week 16.

7.8. Treatment after the End of the Study

7.8.1. Continued Access

After the conclusion of the study, continued access to baricitinib will not be provided to patients. Patients will be referred to their local treatment centers for AD therapy as clinically indicated.

8. Discontinuation Criteria

8.1. Discontinuation from Study Treatment

8.1.1. Temporary Interruption from Investigational Product

In some circumstances, it may be necessary to temporarily interrupt treatment as a result of AEs or abnormal laboratory values that may have an unclear relationship to investigational product. For example, investigational product should be temporarily interrupted if the patient experiences a cardiovascular AE considered to be related to study treatment, is graded as moderate (Grade 2 according to Common Terminology Criteria for Adverse Events Version 3.0), and that does not resolve promptly with supportive care. Except in cases of emergency, it is recommended that the investigator consult with Lilly (or its designee) before temporarily interrupting therapy for reasons other than those defined in Table JAIW.4.

For the abnormal laboratory findings and clinical events (regardless of relatedness) listed in Table JAIW.4, specific guidance is provided for temporarily interrupting treatment and determining when treatment may be restarted. Retest frequency and timing of follow-up laboratory tests to monitor the abnormal finding is at the discretion of the investigator. Investigational product that was temporarily interrupted because of an AE or abnormal laboratory value not specifically covered in Table JAIW.4 may be restarted at the discretion of the investigator.

Table JAIW.4. Criteria for Temporary Interruption of Investigational Product

Hold Investigational Product if the Following Laboratory Test Results or Clinical Events Occur:	Investigational Product May Be Resumed When:
WBC count <2000 cells/µL (<2.00 × 10^3/µL or <2.00 GI/L)	WBC count ≥2500 cells/µL (≥2.50 × 10^3/µL or ≥2.50 GI/L)
ANC <1000 cells/µL (<1.00 × 10^3/µL or <1.00 GI/L)	ANC ≥1200 cells/µL (≥1.20 × 10^3/µL or ≥1.20 GI/L)
Lymphocyte count <500 cells/µL (<0.50 × 10^3/µL or <0.50 GI/L)	Lymphocyte count ≥750 cells/µL (≥0.75 × 10^3/µL or ≥0.75 GI/L)
Platelet count <75,000/µL (<75 × 10^3/µL or <75 GI/L)	Platelet count ≥100,000/µL (≥100 × 10^3/µL or ≥100 GI/L)
eGFR <50 mL/min/1.73 m^2 (from serum creatinine)	eGFR ≥60 mL/min/1.73 m^2
ALT or AST >5 × ULN	ALT and AST return to <2 × ULN and IP is not considered to be the cause of enzyme elevation
Hemoglobin <8 g/dL (<80.0 g/L)	Hemoglobin ≥10 g/dL (≥100.0 g/L)
Symptomatic herpes zoster	All skin lesions have crusted and are resolving
Infection that, in the opinion of the investigator, merits the IP being interrupted.	Resolution of infection

Abbreviations: ALT = alanine aminotransferase; ANC = absolute neutrophil count; AST = aspartate aminotransferase; eGFR = estimated glomerular filtration rate; IP = investigational product; ULN = upper limit of normal; WBC = white blood cell.

For specific guidance on temporary interruption of investigational product after use of a prohibited medication, please refer to Section 7.7.1.

Lastly, investigational product should be temporarily interrupted for suicidal ideation or any suicide-related behaviors as assessed by the following patient responses on the C-SSRS:

- A "yes" answer to Question 4 (Active Suicidal Ideation with Some Intent to Act, Without Specific Plan) **or**
- A "yes" answer to Question 5 (Active Suicidal Ideation with Specific Plan and Intent) on the "Suicidal Ideation" portion of the C-SSRS **or**
- A "yes" answer to any of the suicide-related behaviors (actual attempt, interrupted attempt, aborted attempt, preparatory act or behavior) on the "Suicidal Behavior" portion of the C-SSRS.

Note: Before resumption of investigational product, it is recommended that a patient be assessed by a psychiatrist or appropriately trained professional to assist in deciding whether the patient should remain on investigational product and ultimately continue participation in the study. A patient does not necessarily have to have investigational product interrupted if they have self-injurious behavior that would be classified as non-suicidal self-injurious behavior.

8.1.2. Permanent Discontinuation from Investigational Product

Investigational product should be permanently discontinued if the patient requests to discontinue investigational product.

Discontinuation of the investigational product for abnormal liver tests **should be** considered by the investigator when a patient meets 1 of the following conditions after consultation with the Lilly-designated medical monitor:

- ALT or AST >8 × ULN
- ALT or AST >5 × ULN for >2 weeks
- ALT or AST >3 × ULN and total bilirubin level (TBL) >2 × ULN or international normalized ratio >1.5
- ALT or AST >3 × ULN with the appearance of fatigue, nausea, vomiting, right upper-quadrant pain or tenderness, fever, and/or rash
- ALP >3 × ULN
- ALP >2.5 × ULN and TBL >2 × ULN
- ALP >2.5 × ULN with the appearance of fatigue, nausea, vomiting, right quadrant pain or tenderness, fever, and/or rash

Note: Patients who are discontinued from investigational product because of a hepatic event or liver test abnormality should have additional hepatic safety data collected via the hepatic safety eCRF.

Investigational product should be permanently discontinued if any of the following laboratory abnormalities are observed:

- white blood cell count <1000 cells/µL (1.00 × 10^3/µL or 1.00 GI/L)
- absolute neutrophil count <500 cells/µL (0.50 × 10^3/µL or 0.50 GI/L)
- lymphocyte count <200 cells/µL (0.20 × 10^3/µL or 0.20 GI/L)
- hemoglobin <6.5 g/dL (<65.0 g/L)

Note: Temporary interruption rules (see Section 8.1.1) must be followed where applicable. For laboratory values that meet permanent discontinuation thresholds, investigational product should be discontinued. However, if in the opinion of the investigator the laboratory abnormality is due to intercurrent illness such as cholelithiasis or another identified factor, laboratory tests may be repeated. Only when the laboratory value meets resumption thresholds (Table JAIW.4) following the resolution of the intercurrent illness or other identified factor may the investigator restart investigational product after consultation with the Lilly-designated medical monitor.

In addition, patients will be discontinued from investigational product in the following circumstances:

- pregnancy
- malignancy (except for successfully treated basal or squamous cell skin carcinoma)
- HBV DNA is detected with a value above limit of quantitation or 2 sequential tests return a value below the limit of quantitation (see Section 9.4.8)
- certain prohibited medications are taken per Section 7.7.1
- development of a VTE

  Note: Patients who develop a VTE may have additional follow-up and testing recommended (see Section 9.4.10. and Appendix 6).

If a patient develops multiple risk factors for a VTE during the conduct of the study as described in Exclusion Criterion [23], the investigator may consider study discontinuation if he/she believes the risk outweighs the benefits of continuing therapy. It is recommended that the investigator consult with Lilly (or its designee) before discontinuing therapy for this reason.

If a patient discontinues investigational product for any reason, the patient is encouraged to remain in the study through Week 16 (Visit 8) and follow the regular visit schedule to provide the primary efficacy and safety data. Patients discontinuing from the investigational product prematurely for any reason should complete AE and other follow-up procedures per Section 2 (Schedule of Activities), Section 9.2 (Adverse Events), and Section 9.4 (Safety) of this protocol.

8.1.3. Discontinuation of Inadvertently Enrolled Patients

If the sponsor or investigator identifies a patient who did not meet enrollment criteria and was inadvertently enrolled, then the patient should be discontinued from study treatment unless there are extenuating circumstances that make it medically necessary for the patient to continue on study treatment. If the investigator and the sponsor clinical research physician agree it is medically appropriate to continue, the investigator must obtain documented approval from the sponsor clinical research physician to allow the inadvertently enrolled patient to continue in the study with or without treatment with investigational product. Safety follow up is as outlined in Section 2 (Schedule of Activities), Section 9.2 (Adverse Events), and Section 9.4 (Safety) of this protocol.

8.2. Discontinuation from the Study

Patients may choose to withdraw from the study for any reason at any time, and the reason for early withdrawal will be documented.

Some possible reasons that may lead to permanent discontinuation include the following:

- enrollment in any other clinical study involving an investigational product or enrollment in any other type of medical research judged not to be scientifically or medically compatible with this study
- participation in the study needs to be stopped for medical, safety, regulatory, or other reasons consistent with applicable laws, regulations, and good clinical practice
- investigator decision
    - The investigator decides that the patient should be discontinued from the study.
    - If the patient, for any reason, requires treatment with another therapeutic agent (not allowed as part of rescue therapy [Section 7.7.3]) that has been demonstrated to be effective for treatment of the study indication, discontinuation from the study occurs before introduction of the new agent.
- patient decision
    - The patient requests to be withdrawn from the study.

Patients discontinuing from the study prematurely for any reason should complete AE and other safety follow-up per Section 2 (Schedule of Activities), Section 9.2 (Adverse Events), and Section 9.4 (Safety) of this protocol.

8.3. Lost to Follow-Up

A patient will be considered lost to follow-up if he or she repeatedly fails to return for scheduled visits and is unable to be contacted by the study site. Site personnel are expected to make diligent attempts to contact patients who fail to return for a scheduled visit or were otherwise unable to be followed up by the site.

9. Study Assessments and Procedures

Section 2 lists the Schedule of Activities, with the study procedures and their timing (including tolerance limits for timing).

Appendix 2 and Appendix 4 list the laboratory tests that will be performed for this study.

Unless otherwise stated in the subsections below, all samples collected for specified laboratory tests will be destroyed within 60 days of receipt of confirmed test results. Certain samples may be retained for a longer period, if necessary, to comply with applicable laws, regulations, or laboratory certification standards.

9.1. Efficacy Assessments

9.1.1. Primary Efficacy Assessments

Eczema Area and Severity Index scores

The EASI assesses extent of disease at 4 body regions and measures 4 clinical signs: (1) erythema, (2) induration/papulation, (3) excoriation, and (4) lichenification, each on a scale of 0 to 3. The EASI confers a maximum score of 72. The EASI evaluates 2 dimensions of AD: disease extent and clinical signs (Hanifin et al. 2001).

9.1.2. Secondary Efficacy Assessments

9.1.2.1. Validated Investigator's Global Assessment for Atopic Dermatitis (vIGA-AD)

The IGA used in this study, the vIGA-AD (referred to as the IGA throughout the protocol), measures the IGA of the patient's overall severity of their AD, on the basis of a static, numeric 5-point scale from 0 (clear skin) to 4 (severe disease). The score is based on an overall assessment of the degree of erythema, papulation/induration, oozing/crusting, and lichenification.

9.1.2.2. Eczema Area and Severity Index scores

The EASI assesses extent of disease at 4 body regions and measures 4 clinical signs: (1) erythema, (2) induration/papulation, (3) excoriation, and (4) lichenification, each on a scale of 0 to 3. The EASI confers a maximum score of 72. The EASI evaluates 2 dimensions of AD: disease extent and clinical signs (Hanifin et al. 2001).

BSA affected by AD will be derived from data collected as part of the EASI assessment.

9.1.2.3. SCORing Atopic Dermatitis

The SCORing Atopic Dermatitis (SCORAD index uses the rule of nines to assess disease extent and evaluates 6 clinical characteristics to determine disease severity: (1) erythema, (2) edema/papulation, (3) oozing/crusts, (4) excoriation, (5) lichenification, and (6) dryness. The SCORAD index also assesses subjective symptoms of pruritus and sleep loss. These 3 aspects (extent of disease, disease severity, and subjective symptoms) combine to give a maximum possible score of 103 (Stalder et al. 1993; Kunz et al. 1997; Schram et al. 2012).

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

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).

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

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

See Appendix 4 and Appendix 5 for a description of hepatic laboratory values that warrant patient exclusion from the study, temporary or permanent discontinuation of investigational product, or additional safety collection via the hepatic eCRF.

9.4.10. VTE Assessment

If a patient develops the signs and symptoms of a deep vein thrombosis or pulmonary embolism, appropriate local laboratory tests and imaging should be performed, as necessary, for diagnosis of the event. For confirmed cases, additional laboratory testing should be performed as outlined in Appendix 6. All suspected VTE events will be independently adjudicated by a blinded Clinical Event Committee (see Section 10.3.6.3).

9.4.11. Safety Monitoring

Lilly will periodically review evolving aggregate safety data within the study by appropriate methods.

In the event that safety monitoring uncovers an issue that needs to be addressed by unblinding at the group level, only members of the DMC (an advisory group for this study formed to protect the integrity of data [refer to Interim Analyses section, Section 10.3.6]) can conduct additional analyses of the safety data.

The Lilly clinical research physician will monitor safety data throughout the course of the study. Lilly will review SAEs within time frames mandated by company procedures. The Lilly clinical research physician will, as is appropriate, consult with the functionally independent Global Patient Safety therapeutic area physician or clinical scientist and periodically review trends in safety data and laboratory analytes. Any concerning trends in frequency or severity noted by an investigator and/or Lilly (or designee) may require further evaluation.

All deaths and SAE reports will be reviewed in a blinded manner by Lilly during the clinical trial. These reports will be reviewed to ensure completeness and accuracy but will not be unblinded to Lilly during the clinical trial. If a death or a clinical AE is deemed serious, unexpected, and possibly related to investigational product, only Lilly Global Patient Safety will be unblinded for regulatory reporting and safety monitoring purposes. These measures will preserve the integrity of the data collected during this trial and minimize any potential for bias while providing for appropriate safety monitoring.

Investigators will monitor vital signs and carefully review findings that may be associated with cardiovascular and VTEs. AE reports and vital signs will be collected at each study visit. The cardiovascular monitoring plan includes the following:

- regular monitoring of lipid levels

- 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 will be identified by the investigative site or through medical review and will be sent to a blinded Clinical Event Committee for adjudication at regular intervals.

9.5. Pharmacokinetics

Not applicable.

9.6. Pharmacodynamics

Not applicable.

9.7. Pharmacogenetics

9.7.1. Blood Samples for Pharmacogenetic Research

A blood sample will be collected for pharmacogenetic analysis as specified in the Schedule of Activities (Section 2) where local regulations allow.

There is growing evidence that genetic variation may impact a patient's response to therapy. Variable response to therapy may be due to genetic determinants that impact drug absorption, distribution, metabolism, excretion, the mechanism of action of the drug, the disease etiology, and/or the molecular subtype of the disease being treated. In the event of an unexpected AE, the samples may be genotyped, and analysis may be performed to evaluate a genetic association with response to baricitinib. These investigations may be limited to targeted exome sequencing approach of known targets involved in drug metabolism or, if appropriate, genome-wide association studies may be performed to identify regions of the genome associated with the variability observed in drug response. Samples will only be used for investigations related to disease and drug or class of drugs under study in the context of this clinical program.

Samples will not be used to conduct unspecified disease or population genetic research either now or in the future. Samples will be used to investigate variable response to baricitinib and to investigate genetic variants thought to play a role in AD or other inflammatory skin diseases. Assessment of variable response may include evaluation of AEs or differences in efficacy.

All samples will be coded with the patient number. These samples and any data generated can be linked back to the patient only by the investigator site personnel.

Samples will be retained for a maximum of 15 years after the last patient visit for the study, or for a shorter period if local regulations and/or (ethical review boards/investigational review boards) impose shorter time limits, at a facility selected by Lilly. This retention period enables use of new technologies, response to regulatory questions, and investigation of variable response that may not be observed until later in the development of baricitinib or after baricitinib becomes commercially available.

Molecular technologies are expected to improve during the 15-year storage period and therefore cannot be specifically named. However, existing approaches include whole genome or exome sequencing, genome-wide association studies, and candidate gene studies. Regardless of technology utilized, genotyping data generated will be used only for the specific research scope described in this section.

9.8. Biomarkers

Biomarker research is performed to address questions of relevance to drug disposition, target engagement, pharmacodynamics, mechanism of action, variability of patient response (including safety), and clinical outcome. Sample collection is incorporated into clinical studies to enable examination of these questions through measurement of biomolecules including DNA, RNA, proteins, lipids, and other cellular elements.

Blood samples for nonpharmacogenetic biomarker research will be collected at the times specified in the Schedule of Activities (Section 2) where local regulations allow.

Samples will be used for research on the drug target, disease process, variable response to baricitinib, pathways associated with AD, mechanism of action of baricitinib, and/or research method or in validating diagnostic tools or assay(s) related to AD.

All samples will be coded with the patient number. These samples and any data generated can be linked back to the patient only by the investigator site personnel.

Samples will be retained for a maximum of 15 years after the last patient visit for the study, or for a shorter period if local regulations and ethical review boards impose shorter time limits, at a facility selected by Lilly. This retention period enables use of new technologies, response to regulatory questions, and investigation of variable response that may not be observed until later in the development of baricitinib or after baricitinib becomes commercially available.

9.9. Medical Resource Utilization and Health Economics

Health Economics will be evaluated in this study utilizing the EQ-5D-5L and WPAI-AD (see Section 9.1.3). Medical Resource Utilization parameters will not be evaluated in this study.

10. Statistical Considerations

10.1. Sample Size Determination

Study JAIW will aim to enroll approximately 450 patients aged ≥18 years. The proposed sample size will ensure approximately 68% power to detect any difference between the baricitinib 2-mg and placebo treatment groups, assuming a 10% placebo and 20% baricitinib 2-mg response rate for the primary endpoint EASI75 using a Chi-squared test with a 2-sided α level of 0.05. The assumptions are based on what was observed in the Phase 3 monotherapy Studies JAHL and JAHM.

Sample size and power estimates were obtained from nQuery® Advisor 7.0.

10.2. Populations for Analyses

Unless otherwise specified, the efficacy and health outcome analyses will be conducted on the intent-to-treat population, defined as all randomized patients, even the patients who do not receive the correct treatment or otherwise did not follow the protocol. Patients will be analyzed according to the treatment to which they were assigned. Significant protocol violations will be described in the SAP.

Unless otherwise specified, safety analyses will be performed on all randomized patients who receive at least 1 dose of investigational product and who did not discontinue from the study for the reason "Lost to Follow-up" at the first postbaseline visit.

Further details of other populations will be described in the SAP. Patients will be analyzed according to the dosing regimen to which they were assigned in the Double-Blinded Treatment Period.

10.3. Statistical Analyses

10.3.1. General Statistical Considerations

Statistical analysis of this study will be the responsibility of Lilly or its designee. A detailed SAP describing the statistical methodologies will be developed by Lilly or its designee.

Any change to the data analysis methods described in the protocol will require an amendment ONLY if it changes a principal feature of the protocol. Any other change to the data analysis methods described in the protocol and the justification for making the change will be described in the clinical study report. Additional exploratory analyses of the data will be conducted as deemed appropriate.

All tests of treatment effects will be conducted at a 2-sided α level of 0.05, unless otherwise stated. Treatment comparisons of discrete efficacy variables between baricitinib and placebo will be made using a logistic regression analysis with IGA baseline disease severity, baseline score and treatment group in the model. Region may be added as an additional factor if there is a sufficient number of patients for rest-of-world and stratification has not resulted in empty strata. This will be finalized in the SAP. If appropriate, treatment by region interaction may be added

to the model of the primary and key secondary variables as a sensitivity analysis. If this interaction is significant at a 2-sided 0.1 level, further inspection will be used to assess whether the interaction is quantitative (that is, the treatment effect is consistent in direction but not size of effect) or qualitative (the treatment is beneficial for some but not all regions). The percentages, difference in percentages, and 95% CI of the difference in percentages will be reported. The p-value from the Fisher's exact test will also be produced.

When evaluating continuous measures over time, a restricted maximum likelihood-based MMRM will be used. The model will include treatment, baseline disease severity (IGA), baseline score, visit, and treatment-by-visit interaction as fixed categorical effects and baseline score and baseline score-by-visit interaction as fixed continuous effects. Geographic region may be added as an additional factor if appropriate. An unstructured (co)variance structure will be used to model the between- and within-patient errors. If this analysis fails to converge, other structures will be tested. The Kenward–Roger method will be used to estimate the degrees of freedom. Type III sums of squares for the LSMs will be used for the statistical comparison; 95% CI will also be reported. Contrasts will be set up within the model to test treatment groups at specific time points of interest. Further details on the use of MMRM will be described in the SAP.

Treatment comparisons of continuous efficacy and health outcome variables may also be made using ANCOVA with baseline disease severity (IGA), treatment group, and baseline score in the model. Geographic region may be added as an additional factor if appropriate. Type III tests for LSM will be used for statistical comparison between treatment groups. The LSM difference, standard error, p-value, and 95% CI may also be reported. The method used to handle missing data will be specified in the SAP.

Fisher's exact test will be used for the AEs, discontinuation, and other categorical safety data for between-treatment group comparisons. Continuous vital signs, body weight, and other continuous safety variables including laboratory variables will be analyzed by an ANCOVA with treatment and baseline value in the model. Shift tables for categorical safety analyses (for example, "high" or "low" laboratory results) will also be produced.

Missing Data Imputation:

1. NRI: All patients who discontinue the study or the study treatment at any time for any reason will be defined as nonresponders for the NRI analysis for categorical variables such as IGA 0/1 or EASI 50/75/90 after discontinuation and onward. Patients who receive rescue therapy will be analyzed as nonresponders after rescue and onward. An additional analysis will be performed that includes all available data whether rescue medication was given or not.
2. Continuous variables such as EASI and SCORAD scores will be assumed to be missing after rescue or discontinuation, and then an MMRM analysis will be performed. An additional analysis will be performed that includes all available data whether rescue medication was given or not.

3. Last observed carried forward: An additional analysis will be performed that uses the last observed value on or before discontinuation/dropout or rescue therapy. This will then be analyzed using a logistic model for categorical variables or ANCOVA for continuous variables as described above.

Additional sensitivity analyses for the primary and key secondary endpoints such as tipping point analyses as well as placebo multiple imputation will be done. Details of these methods are in Appendix 9, and any changes to these details will be finalized in the SAP.

Adjustment for Multiple Comparisons:

The following is a list of primary and key secondary hypotheses to be tested:

Primary Null Hypothesis:

$H_{2,1,0}$: Proportion of baricitinib 2-mg patients achieving EASI75 is less than or equal to the proportion of placebo patients achieving EASI75 at Week 16

Key Secondary Null Hypotheses:

$H_{2,2,0}$: Proportion of baricitinib 2-mg patients achieving IGA of 0 or 1 and ≥2-point improvement from baseline at Week 16 is less than or equal to the proportion of placebo patients achieving IGA of 0 or 1 and ≥2-point improvement from baseline at Week 16

$H_{2,3,0}$: Proportion of baricitinib 2-mg patients achieving EASI90 is less than or equal to the proportion of placebo patients achieving EASI90 at Week 16

$H_{2,4,0}$: Mean percent change from baseline in EASI score for baricitinib 2-mg is less than or equal to the mean percent change from baseline in EASI score for placebo at Week 16

$H_{2,5,0}$: Proportion of baricitinib 2-mg patients achieving SCORAD75 is less than or equal to the proportion of placebo patients achieving SCORAD75 at Week 16

$H_{2,6,0}$: Proportion of baricitinib 2-mg patients achieving a 4-point improvement in Itch NRS is less than or equal to the proportion of placebo patients achieving a 4-point improvement in Itch NRS at Week 16

$H_{2,7,0}$: Proportion of baricitinib 2-mg patients achieving a 4-point improvement in Itch NRS is less than or equal to the proportion of placebo patients achieving a 4-point improvement in Itch NRS at Week 4

$H_{2,8,0}$: Proportion of baricitinib 2-mg patients achieving a 4-point improvement in Itch NRS is less than or equal to the proportion of placebo patients achieving a 4-point improvement in Itch NRS at Week 2

$H_{2,9,0}$: Proportion of baricitinib 2-mg patients achieving a 4-point improvement in Itch NRS is less than or equal to the proportion of placebo patients achieving a 4-point improvement in Itch NRS at Week 1

$H_{2,10,0}$: Mean change from baseline in the score of Item 2 of the ADSS for baricitinib 2-mg is less than or equal to the mean change from baseline in the score of Item 2 of the ADSS for placebo at Week 16

$H_{2,11,0}$: Mean change from baseline in the score of Item 2 of the ADSS for baricitinib 2-mg is less than or equal to the mean change from baseline in the score of Item 2 of the ADSS for placebo at Week 1

$H_{2,12,0}$: Mean change from baseline in Skin Pain NRS for baricitinib 2-mg is less than or equal to the mean change from baseline in Skin Pain NRS for placebo at Week 16
$H_{1,1,0}$: Proportion of baricitinib 1-mg patients achieving EASI75 is less than or equal to the proportion of placebo patients achieving EASI75 at Week 16
$H_{1,2,0}$: Proportion of baricitinib 1-mg patients achieving IGA of 0 or 1 and ≥2-point improvement from baseline at Week 16 is less than or equal to the proportion of placebo patients achieving IGA of 0 or 1 and ≥2-point improvement from baseline at Week 16
$H_{1,3,0}$: Proportion of baricitinib 1-mg patients achieving EASI90 is less than or equal to the proportion of placebo patients achieving EASI90 at Week 16
$H_{1,4,0}$: Mean percent change from baseline in EASI score for baricitinib 1-mg is less than or equal to the mean percent change from baseline in EASI score for placebo at Week 16
$H_{1,5,0}$: Proportion of baricitinib 1-mg patients achieving SCORAD75 is less than or equal to the proportion of placebo patients achieving SCORAD75 at Week 16
$H_{1,6,0}$: Proportion of baricitinib 1-mg patients achieving a 4-point improvement in Itch NRS is less than or equal to the proportion of placebo patients achieving a 4-point improvement in Itch NRS at Week 16
$H_{1,7,0}$: Proportion of baricitinib 1-mg patients achieving a 4-point improvement in Itch NRS is less than or equal to the proportion of placebo patients achieving a 4-point improvement in Itch NRS at Week 4
$H_{1,8,0}$: Proportion of baricitinib 1-mg patients achieving a 4-point improvement in Itch NRS is less than or equal to the proportion of placebo patients achieving a 4-point improvement in Itch NRS at Week 2
$H_{1,9,0}$: Proportion of baricitinib 1-mg patients achieving a 4-point improvement in Itch NRS is less than or equal to the proportion of placebo patients achieving a 4-point improvement in Itch NRS at Week 1
$H_{1,10,0}$: Mean change from baseline in the score of Item 2 of the ADSS for baricitinib 1-mg is less than or equal to the mean change from baseline in the total score of the ADSS for placebo at Week 16
$H_{1,11,0}$: Mean change from baseline in the score of Item 2 of the ADSS for baricitinib 1-mg is less than or equal to the mean change from baseline in the total score of the ADSS for placebo at Week 1
$H_{1,12,0}$: Mean change from baseline in Skin Pain NRS for baricitinib 1-mg is less than or equal to the mean change from baseline in Skin Pain NRS for placebo at Week 16

Multiplicity adjusted analyses will be performed on these primary and key secondary null hypotheses to control the overall family-wise Type I error rate at a 2-sided α level of 0.05. The graphical multiple testing procedure described in Bretz et al. (2011), which is a closed testing procedure, will be used; therefore, it strongly controls the family-wise error rate across all endpoints (Alosh et al. 2014). Figure JAIW.2 depicts the graphical testing scheme (including testing order, interrelationships, Type I error allocation, and the associated propagation) and is further described in Appendix 10. Further details of the specific graphical testing scheme

(including testing order, interrelationships, Type I error allocation, and the associated propagation) will be prespecified and finalized in the SAP. These details may be revised when newer information is obtained on the endpoints that are being tested. However, the graphical testing scheme will be finalized before primary database lock, which will occur after the last patient in the study completes Visit 8 (Week 16) or is permanently discontinued from the study.

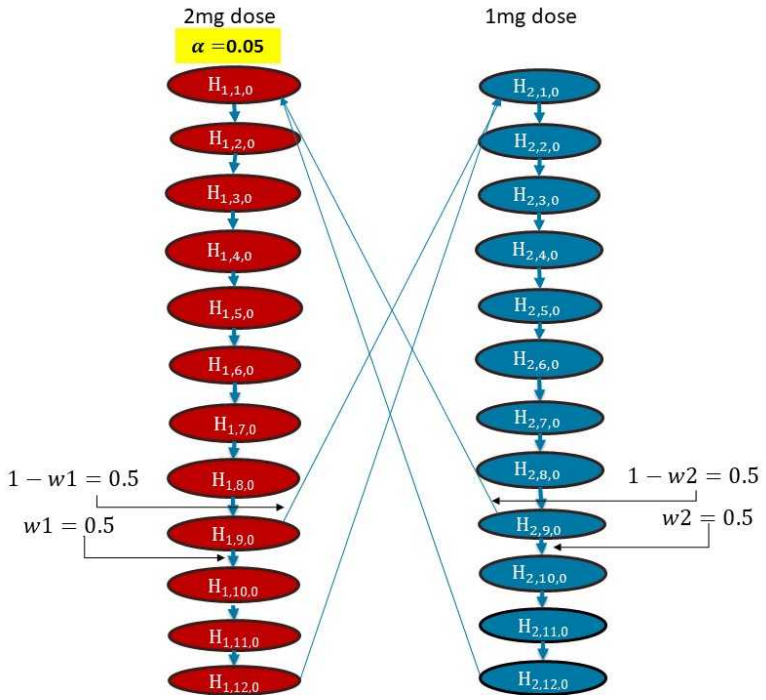


Figure JAIW.2. Study I4V-MC-JAIW scheme structure.

Abbreviations: H = null hypothesis.

10.3.2. Treatment Group Comparability

10.3.2.1. Patient Disposition

All patients who discontinue from the study or the study treatment will be identified, along with their reason for discontinuation. Reasons for discontinuation from the study will be summarized by treatment group.

10.3.2.2. Patient Characteristics

Demographic and baseline characteristics will be summarized descriptively by treatment group. Descriptive statistics including number of patients, mean, standard deviation, median, minimum, and maximum will be provided for continuous measures, and frequency counts and percentages will be tabulated for categorical measures. No formal statistical comparisons will be made among treatment groups unless otherwise stated.

10.3.2.3. Concomitant Therapy

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

10.3.2.4. Treatment Compliance

Treatment compliance with the randomly assigned study medication will be evaluated at every clinic visit through the counts of returned investigational product tablets. A patient will be considered significantly noncompliant if he or she misses more than 20% of the prescribed doses during the study, unless the patients investigational product is withheld by the investigator for safety reasons (that is, compliance <80%). Similarly, a patient will be considered significantly noncompliant if he or she is judged by the investigator to have intentionally or repeatedly taken more than the prescribed amount of medication (that is, compliance ≥120%).

10.3.3. Efficacy Analyses

10.3.3.1. Primary Analyses

The primary efficacy measure is the categorical outcome of EASI75 at Week 16 for baricitinib 2-mg. EASI75 is defined as having an improvement of at least 75% from baseline. Primary analysis will be conducted using a logistic regression as described above with treatment and stratification variables, which include baseline disease severity (IGA), baseline score and region, as appropriate in the model. NRI for missing data as described above will be used.

Additional analysis of the primary efficacy outcome will include analyzing the outcome as observed, that is, whether rescue medication was given.

10.3.3.2. Secondary Analyses

The following secondary categorical outcomes will be analyzed in a similar manner as the primary, that is, using the same logistic regression model. This applies to both baricitinib 2-mg vs placebo, and baricitinib 1-mg vs placebo unless stated otherwise. NRI will be used for these analyses unless otherwise noted.

- IGA score of 0 or 1 (clear or almost clear skin) and ≥2-point improvement from baseline at Week 16. Besides NRI, this outcome will also be analyzed using observed cases, that is, whether rescue medication was given.
- EASI75 at Week 16 for baricitinib 1-mg vs placebo. EASI75 is defined as having an improvement of at least 75% from baseline.
- EASI90 at Week 16. EASI90 is defined as having an improvement of at least 90% from baseline.
- SCORAD75 at Week 16. SCORAD75 is defined as having an improvement of at least 75% from baseline.
- SCORAD90 at Week 16. SCORAD90 is defined as having an improvement of at least 90% from baseline.
- 4-point improvement in Itch NRS at 1 week, 2 weeks, 4 weeks, and 16 weeks.

The following continuous measures will be analyzed with the MMRM model described above unless otherwise noted. Contrasts within the MMRM model will be used to assess treatment differences for time points of interest as specified above in the list of objectives.

- Mean change from baseline in the following outcome measures:
    - ADSS Item 2 score
    - EASI score
    - SCORAD score
    - BSA
    - Itch NRS
    - POEM total score
    - PGI-S-AD
    - HADS
    - DLQI total score
    - WPAI-AD
    - EQ-5D-5L

The EASI total score and SCORAD total score will also be analyzed as observed, that is, not assuming missing values after rescue medication is given.

10.3.4. Safety Analyses

All safety data will be descriptively summarized by treatment groups and analyzed using the safety population.

TEAEs are defined as AEs that first occurred or worsened in severity after the first dose of study treatment. The number of TEAEs as well as the number and percentage of patients who experienced at least 1 TEAE will be summarized using Medical Dictionary for Regulatory Activities for each system organ class (or a body system) and each preferred term by treatment group. SAEs and AEs that lead to discontinuation of investigational product will also be summarized by treatment group. Fisher's exact test will be used to perform comparisons between each baricitinib dose and the placebo group.

All clinical laboratory results will be descriptively summarized by treatment group. Individual results that are outside of normal reference ranges will be flagged in data listings. Quantitative clinical hematology, chemistry, and urinalysis variables obtained at the baseline to postbaseline visits will be summarized as changes from baseline by treatment group and analyzed using ANCOVA with treatment and baseline value in the model. Categorical variables, including the incidence of abnormal values and incidence of AEs of special interest, will be summarized by frequency and percentage of patients in corresponding categories. Shift tables will be presented for selected measures.

Observed values and changes from baseline (predose or screening if missing) for vital signs and physical characteristics will be descriptively summarized by treatment group and time point. Change from baseline to postbaseline in vital signs and body weight will be analyzed using ANCOVA with treatment and baseline value in the model.

The incidence and average duration of investigational product interruptions will be summarized and compared descriptively among treatment groups. Various techniques may be used to estimate the effects of investigational product interruptions on safety measures. Further analyses may be performed and will be planned in the SAP.

Data collected after initiation of rescue therapy will be summarized as appropriate.

10.3.5. Other Analyses

10.3.5.1. Health Outcome Measures

The health outcome measures will be analyzed using methods described for continuous or categorical data as described for efficacy measures in Section 10.3.3.

10.3.5.2. Subgroup Analyses

To assess whether the treatment effect is similar across subgroups for the primary efficacy outcome, a logistic model will be used and will include treatment, stratification variables (IGA (3,4) at baseline, baseline disease severity and region) as appropriate, the subgroup variable (for example, sex), and the subgroup-by-treatment interaction. If the interaction is statistically significant at α level of 0.10, the nature of the interaction will be explored, that is, within each subgroup the treatment effect will be estimated. Similarly, for the continuous variables of EASI, the MMRM model will include additional variables for subgroup and the subgroup-by-treatment interaction.

Subgroups to be evaluated will include geographic region (as appropriate), baseline severity, sex, age, race, previous therapy, etc. Further definitions for the levels of the subgroup variables, the analysis methodology, and any additional subgroup analyses will be defined in the SAP. Because this study is not powered for subgroup analyses, all subgroup analyses will be treated as exploratory.

10.3.6. Interim Analyses

10.3.6.1. Data Monitoring Committee

A DMC will monitor the overall safety of this trial. The DMC will consist of members external to Lilly. This DMC will follow the rules defined in the DMC charter, focusing on potential and identified risks for this molecule and for this class of compounds. DMC membership will include, at a minimum, specialists with expertise in dermatology, statistics, and other appropriate specialties.

The DMC will be authorized to review unblinded results of analyses by treatment group before 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 α is spent. Details of

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.

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

and no or limited prior disease-modifying antirheumatic drug treatment. *Arthritis Rheumatol*. 2017;69(3):506-517.

Fridman JS, Scherle PA, Collins R, Burn TC, Li Y, Li J, Covington MB, Thomas B, Collier P, Favata MF, Wen X, Shi J, McGee R, Haley PJ, Shepard S, Rodgers JD, Yeleswaram S, Hollis G, Newton RC, Metcalf B, Friedman SM, Vaddi K. Selective inhibition of JAK1 and JAK2 is efficacious in rodent models of arthritis: preclinical characterization of INCB028050. *J Immunol*. 2010;184(9):5298-5307.

Futamura M, Leshem YA, Thomas KS, Nankervis H, Williams HC, Simpson EL. A systematic review of Investigator Global Assessment (IGA) in atopic dermatitis (AD) trials: Many options, no standards. *J Am Acad Dermatol*. 2016;74(2):288-294.

Genovese MC, Kremer J, Zamani O, Ludivico C, Krogulec M, Xie L, Beattie SD, Koch AE, Cardillo TE, Rooney TP, Macias WL, de Bono S, Schlichting DE, Smolen JS. Baricitinib in patients with refractory rheumatoid arthritis. *N Engl J Med*. 2016;374(13):1243-1252.

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

Herdman M, Gudex C, Lloyd A, Janssen MF, Kind P, Parkin D, Bonsel G, Badia X. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res*. 2011;20(10):1727-1736.

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

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

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

Krakowski AC, Eichenfield LF, Dohil MA. Management of atopic dermatitis in the pediatric population. *Pediatrics*. 2008;122(4):812-824.

Kunz B, Oranje A, Labrèze, Stadler JF, Ring J, Taïeb A. Clinical validation and guidelines for the SCORAD Index: consensus report of the European task force on atopic dermatitis. *Dermatology*. 1997;195(1):10-19.

Langenbruch A, Radtke M, Franzke N, Ring J, Foelster-Holst R, Augustin M. Quality of health care of atopic eczema in Germany: results of the national health care study AtopicHealth. *J Eur Acad Dermatol Venereol*. 2014;28(6):719-726.

Langley RG, Feldman SR, Nyirady J, van de Kerkhof P, Papavassilis C. The 5-point Investigator's Global Assessment (IGA) Scale: A modified tool for evaluating plaque psoriasis severity in clinical trials. *J Dermatolog Treat*. 2015;26(1):23-31.

Li M, Messaddeq N, Teletin M, Pasquali JL, Metzger D, Chambon P. Retinoid X receptor ablation in adult mouse keratinocytes generates an atopic dermatitis triggered by thymic stromal lymphopoietin. *Proc Natl Acad Sci U S A*. 2005;102(41):14795-14800.

Mallinckrodt CH, Tamura RN, Tanaka Y. Recent developments in improving signal detection and reducing placebo response in psychiatric clinical trials. *J Psychiatric Res*. 2011;45(9):1202-1207.

Moniaga CS, Jeong SK, Egawa G, Nakajima S, Hara-Chikuma M, Jeon JE, Lee SH, Hibino T, Miyachi Y, Kabashima K. Protease activity enhances production of thymic stromal lymphopoietin and basophil accumulation in flaky tail mice. *Am J Pathol*. 2013;182(3):841-851.

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

Nomura T, Kabashima K. Advances in atopic dermatitis in 2015. *J Allergy Clin Immunol*. 2016;138(6):1548-1555.

Palmer CN, Irvine AD, Terron-Kwiatkowski A, Zhao Y, Liao H, Lee SP, Goudie DR, Sandilands A, Campbell LE, Smith FJ, O'Regan GM, Watson RM, Cecil JE, Bale SJ, Compton JG, DiGiovanna JJ, Fleckman P, Lewis-Jones S, Arseculeratne G, Sergeant A, Munro CS, El Houate B, McElreavey K, Halkjaer LB, Bisgaard H, Mukhopadhyay S, McLean WH. Common loss-of-function variants of the epidermal barrier protein filaggrin are a major predisposing factor for atopic dermatitis. *Nat Genet.* 2006;38(4):441-446.

Papp KA, Menter MA, Raman M, Disch D, Schlichting DE, Gaich C, Macias W, Zhang X, Janes JM. A randomized phase 2b trial of baricitinib, an oral Janus kinase (JAK) 1/JAK2 inhibitor, in patients with moderate-to-severe psoriasis. *Br J Dermatol*. 2016;174(6):1266-1276.

Reilly MC, Zbrozek AS, Dukes EM. The validity and reproducibility of a work productivity and activity impairment instrument. *Pharmacoeconomics*. 1993;4(5):353-365.

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

Simpson EL. Comorbidity in atopic dermatitis. *Curr Dermatol Rep*. 2012;1(1):29-38.

Snaith RP. The Hospital Anxiety and Depression Scale. *Health Qual Life Outcomes*. 2003;1:29.

Stalder JF, Taïeb A. Severity scoring of atopic dermatitis: the SCORAD index. Consensus report of the European task force on atopic dermatitis. *Dermatology*. 1993;186(1):23-31.

Tadicherla S, Ross K, Shenefelt PD, Fenske NA. Topical corticosteroids in dermatology. *J Drugs Dermatol.* 2009;8(12):1093-1105.

Taylor PC, Keystone EC, van der Heijde D, Weinblatt ME, del Carmen Morales L, Gonzaga JR, Yakushin S, Ishii T, Emoto K, Beattie S, Arora V, Gaich CL, Rooney T, Schlichting D, Macias WL, de Bono S, Tanaka Y. Baricitinib versus placebo or adalimumab in rheumatoid arthritis. *N Engl J Med.* 2017;376(7):652-662.

Tuttle KR, Brosius FC, Adler SG, Kretzler M, Mehta RL, Tumlin JA, Liu J, Silk ME, Cardillo TE, Duffin KL, Haas JV, Macias WL, Janes JM. Baricitinib in diabetic kidney disease: Results from a Phase 2, multicenter, randomized, double-blind, placebo-controlled study. Paper presented at: Annual Meeting of the American Diabetes Association; 5–9 June 2015; Boston, MA.

Vaddi K, Luchi M. JAK inhibition for the treatment of rheumatoid arthritis: a new era in oral DMARD therapy. *Expert Opin Investig Drugs*. 2012;21(7):961-973.

White D, Leach C, Sims R, Atkinson M, Cottrell D. Validation of the Hospital Anxiety and Depression Scale for use with adolescents. *Br J Psychiatry*. 1999;175(5):452-454.

[WHO] World Health Organization. Model Prescribing Information: Drugs used in skin diseases. 1997; Geneva.

Wilson SR, Thé L, Batia LM, Beattie K, Katibah GE, McClain SP, Pellegrino M, Estandian DM, Bautista DM. The epithelial cell-derived atopic dermatitis cytokine TSLP activates neurons to induce itch. *Cell*. 2013;155(2):285-295.

Yano C, Saeki H, Ishiji T, Ishiuji Y, Sato J, Tofuku Y, Nakagawa H. Impact of disease severity on sleep quality in Japanese patients with atopic dermatitis. *J Dermatol Sci*. 2013;72(2):195-197.

Ziegler SF, Roan F, Bell BD, Stoklasek TA, Kitajima M, Han H. The biology of thymic stromal lymphopoietin (TSLP). *Adv Pharmacol*. 2013;66:129-155.

Zigmond AS, Snaith RP. The Hospital Anxiety and Depression Scale. *Acta Psychiatr Scand*. 1983;67(6):361-370.

12. Appendices

Appendix 1. Abbreviations and Definitions

Term	Definition
AD	atopic dermatitis
ADSS	Atopic Dermatitis Sleep Scale
AE	adverse event: Any untoward medical occurrence in a patient or clinical investigation subject administered a pharmaceutical product that does not necessarily have a causal relationship with this treatment. An adverse event can therefore be any unfavorable and unintended sign (including an abnormal laboratory finding), symptom, or disease temporally associated with the use of a medicinal (investigational) product, whether or not related to the medicinal (investigational) product.
ALP	alkaline phosphatase
ALT	alanine aminotransferase
ANCOVA	analysis of covariance
AST	aspartate aminotransferase
blinding/masking	A single-blind study is one in which the investigator and/or his staff are aware of the treatment but the patient is not, or vice versa, or when the sponsor is aware of the treatment but the investigator and/his staff and the patient are not. A double-blind study is one in which neither the patient nor any of the investigator or sponsor staff who are involved in the treatment or clinical evaluation of the subjects are aware of the treatment received.
BMI	body mass index
BSA	body surface area
C-SSRS	Columbia Suicide Severity Rating Scale
CSR	clinical study report
DLQI	Dermatology Life Quality Index
DMC	data monitoring committee
EASI	Eczema Area and Severity Index
ECG	electrocardiogram
eCOA	electronic clinical outcome assessment
eCRF	electronic case report form
enroll	The act of assigning a patient to a treatment. Patients who are enrolled in the trial are those who have been assigned to a treatment.
Enter	Patients entered into a trial are those who sign the informed consent form directly or through their legally acceptable representatives.
ePRO	electronic patient-reported outcome
EQ-5D-5L	European Quality of Life–5 Dimensions–5 Levels
ETV	early termination visit
GCP	good clinical practice
HADS	Hospital Anxiety Depression Scale
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.

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

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.

[e] eGFR for serum creatinine calculated by the central laboratory using the Chronic Kidney Disease Epidemiology Collaboration Creatinine 2009 equation.

[f] Test required at Visit 1 only to determine eligibility of patient for the study.

[g] A positive hepatitis C antibody result will be confirmed with an alternate hepatitis C method.

[h] For all women of childbearing potential, a serum pregnancy test will be performed at Visit 1 and a local urine pregnancy test will be performed at Visit 2 and at all subsequent study visits after Visit 3. If required per local regulations and/or institutional guidelines, pregnancy testing can occur at other times during the study treatment period.

[i] To confirm postmenopausal status for women ≥40 and <60 years of age who have had a cessation of menses, an FSH test will be performed. Nonchildbearing potential is defined as an FSH ≥40 mIU/mL and a cessation of menses for at least 12 months.

[j] The QuantiFERON®-TB Gold test is the preferred alternative to the PPD test for the evaluation of TB infection, and it may be used instead of the PPD test or T-SPOT® TB test and may be read locally. If the QuantiFERON-TB Gold test is indeterminate, 1 retest is allowed. If the retest is indeterminate, then the patient is excluded from the study.

[k] HBV DNA testing will be done in those patients who are HBcAb+ at screening.

Appendix 3. Study Governance Considerations

Appendix 3.1. Regulatory and Ethical Considerations, Including the Informed Consent Process

Appendix 3.1.1. Informed Consent

The investigator is responsible for ensuring the following:

- that the patient understands the potential risks and benefits of participating in the study.
- that informed consent is given by each patient. This includes obtaining the appropriate signatures and dates on the informed consent form (ICF) prior to the performance of any protocol procedures and prior to the administration of investigational product.
- answering any questions the patient may have throughout the study and sharing in a timely manner any new information that may be relevant to the patient's willingness to continue his or her participation in the trial.

Appendix 3.1.2. Ethical Review

The investigator must give assurance that the ethical review board (ERB) was properly constituted and convened as required by International Council for Harmonisation (ICH) guidelines and other applicable laws and regulations.

Documentation of ERB approval of the protocol and the ICF must be provided to Eli Lilly and Company (Lilly) before the study may begin at the investigative site(s). Lilly or its representatives must approve the ICF, including any changes made by the ERBs, before it is used at the investigative site(s). All ICFs must be compliant with the ICH guideline on Good Clinical Practice (GCP).

The study site's ERB(s) should be provided with the following:

- the current investigator's brochure and updates during the course of the study
- ICF
- relevant curricula vitae.

Appendix 3.1.3. Regulatory Considerations

This study will be conducted in accordance with:

- consensus ethics principles derived from international ethics guidelines, including the Declaration of Helsinki and Council for International Organizations of Medical Sciences (CIOMS) International Ethical Guidelines
- applicable ICH GCP Guidelines

- applicable laws and regulations.

Some of the obligations of the sponsor will be assigned to a third party.

Appendix 3.1.4. Investigator Information

Physicians with a specialty in dermatology will participate as investigators in this clinical trial.

Appendix 3.1.5. Protocol Signatures

The sponsor's responsible medical officer will approve the protocol, confirming that, to the best of his or her knowledge, the protocol accurately describes the planned design and conduct of the study.

After reading the protocol, each principal investigator will sign the protocol signature page and send a copy of the signed page to a Lilly representative.

Appendix 3.1.6. Final Report Signature

Lilly will select a qualified investigator(s) from among investigators participating in the design, conduct, and/or analysis of the study to serve as the clinical study report (CSR) coordinating investigator. If this investigator is unable to fulfill this function, another investigator will be chosen by Lilly to serve as the CSR coordinating investigator.

The CSR coordinating investigator will sign the final CSR for this study, indicating agreement that, to the best of his or her knowledge, the report accurately describes the conduct and results of the study.

The sponsor's responsible medical officer and statistician will approve the final CSR for this study, confirming that, to the best of his or her knowledge, the report accurately describes the conduct and results of the study.

Appendix 3.2. Data Quality Assurance

To ensure accurate, complete, and reliable data, Lilly or its representatives will do the following:

- provide instructional material to the study sites, as appropriate
- provide sponsor start-up training to instruct the investigators and study coordinators. This training will give instruction on the protocol, the completion of the CRFs, and study procedures.
- make periodic visits to the study site
- be available for consultation and stay in contact with the study site personnel by mail, telephone, and/or fax
- review and evaluate CRF data and use standard computer edits to detect errors in data collection
- conduct a quality review of the database.

In addition, Lilly or its representatives will periodically check a sample of the patient data recorded against source documents at the study site. The study may be audited by Lilly or its representatives and/or regulatory agencies at any time. Investigators will be given notice before an audit occurs.

The investigator will keep records of all original source data. This might include laboratory tests, medical records, and clinical notes. If requested, the investigator will provide the sponsor, applicable regulatory agencies, and applicable ERBs with direct access to original source documents.

Appendix 3.2.1. Data Capture System

An electronic data capture system will be used in this study. The site maintains a separate source for the data entered by the site into the sponsor-provided electronic data capture system.

Electronic patient-reported outcome (ePRO) measures (for example, a rating scale) and electronic clinical outcome assessments (eCOAs) are entered into an ePRO/eCOA instrument at the time that the information is obtained. In these instances where there is no prior written or electronic source data at the site, the ePRO/eCOA instrument record will serve as the source.

If ePRO/eCOA records are stored at a third-party site, investigator sites will have continuous access to the source documents during the study and will receive an archival copy at the end of the study for retention.

Any data for which the ePRO/eCOA instrument record will serve to collect source data will be identified and documented by each site in that site's study file.

Case report form data will be encoded and stored in InForm. Data managed by a central vendor, such as laboratory test data, will be stored electronically in the central vendor's database system. Data will subsequently be transferred from the central vendor to the Lilly data warehouse.

Data managed by a central vendor, such as laboratory test data, will be stored electronically in the central vendor's database system. Data will subsequently be transferred from the central vendor to the Lilly data warehouse.

Data from complaint forms submitted to Lilly will be encoded and stored in the global product complaint management system.

Appendix 3.3. Study and Site Closure

Appendix 3.3.1. Discontinuation of Study Sites

Study site participation may be discontinued if Lilly, the investigator, or the ERB of the study site judges it necessary for medical, safety, regulatory, or other reasons consistent with applicable laws, regulations, and GCP.

Appendix 3.3.2. Discontinuation of the Study

The study will be discontinued if Lilly judges it necessary for medical, safety, regulatory, or other reasons consistent with applicable laws, regulations, and GCP.

Appendix 4. Hepatic Monitoring Tests for Treatment-Emergent Abnormality

Selected tests may be obtained in the event of a treatment-emergent hepatic abnormality and may be required in follow-up with patients in consultation with the Lilly, its designee, or the clinical research physician.

Hepatic Monitoring Tests

Hepatic Hematology[a]	**Haptoglobin[a]**
Hemoglobin	
Hematocrit	**Hepatic Coagulation[a]**
RBC	Prothrombin Time
WBC	Prothrombin Time, INR
Neutrophils, segmented	
Lymphocytes	**Hepatic Serologies[a,b]**
Monocytes	Hepatitis A antibody, total
Eosinophils	Hepatitis A antibody, IgM
Basophils	Hepatitis B surface antigen
Platelets	Hepatitis B surface antibody
	Hepatitis B Core antibody
Hepatic Chemistry[a]	Hepatitis C antibody
Total bilirubin	Hepatitis E antibody, IgG
Direct bilirubin	Hepatitis E antibody, IgM
Alkaline phosphatase	
ALT	**Anti-nuclear antibody[a]**
AST	
GGT	**Alkaline Phosphatase Isoenzymes[a]**
CPK	
	Anti-smooth muscle antibody (or anti-actin antibody)[a]

Abbreviations: ALT = alanine aminotransferase; AST = aspirate aminotransferase; CPK = creatinine phosphokinase; GGT = gamma-glutamyl transferase; Ig = immunoglobulin; INR = international normalized ratio; RBC = red blood cell; WBC = white blood cell.

[a] Assayed by Lilly-designated or local laboratory.

[b] Reflex/confirmation dependent on regulatory requirements and/or testing availability.

Appendix 5. Liver Function Testing and Hepatic Safety Monitoring

Liver Function Testing and Hepatic Safety Monitoring

Analyte	Exclusion Criteria	Additional Hepatic Testing	Hepatic eCRF Reporting	Temporary Interruption of IP	Permanent Discontinuation of IP after Consultation with the Lilly-Designated Medical Monitor
Protocol Section	Section 6.2	Section 9.4.9	Section 9.4.9	Section 8.1.1	Section 8.1.2
ALT/AST	≥2x ULN	ALT only ≥3x ULN	ALT only ≥5x ULN on ≥2 consecutive tests	≥5x ULN	• >8x ULN • >5x ULN for >2 weeks • >3x ULN AND TBL >2x ULN or INR >1.5 • >3x ULN with symptoms[a]
ALP	≥2x ULN	≥2x ULN	≥2x ULN on ≥2 consecutive tests	No applicable criteria	• >3x ULN • >2.5x ULN AND TBL >2x ULN • >2.5x ULN with symptoms[a]
TBL	≥1.5x ULN	≥2x ULN	≥2x ULN (excluding Gilbert's syndrome)	No applicable criteria	• ALT or AST >3x ULN AND TBL >2x ULN • ALP >2.5x ULN AND TBL >2x ULN

Abbreviations: ALP = alkaline phosphatase; ALT = alanine aminotransferase; AST = aspartate aminotransferase; INR = international normalized ratio; IP = investigational product; TBL = total bilirubin level; ULN = upper level of normal.

[a] Fatigue, nausea, vomiting, right upper-quadrant pain or tenderness, fever, and/or rash.

Appendix 6. Monitoring Tests for Confirmed VTE

Selected tests may be obtained in the event of a confirmed venous thromboembolic event (VTE) and may be required in follow-up with patients in consultation with Eli Lilly and Company, its designee, or the clinical research physician. The choice and optimal timing of these tests will be directed by the patient's management and may require ongoing follow-up after study discontinuation.

Protein C Functional
Protein S Clottable
Antithrombin III
APC Resistance
PT
APTT
Fibrinogen
Cardiolipin Antibodies
PT Gene
Factor VIII C Assay
Hexagonal Phase Phospholipid Neutralization
C-Reactive Protein
PTT Incubated Mixing
Dilute Russell Viper Venom
Platelet Neutralization
Factor V Leiden
MTHFR
Thrombin Time
Reptilase
Fibrinogen Antigen
Protein C Immunologic
Protein S Immunologic
Heparin fXa Inhibition

Abbreviations: APC = activated protein C; APTT = activated partial thromboplastin time; MTHFR = methylene tetrahydrofolate reductase; PT = prothrombin time; PTT = partial thromboplastin time.

Appendix 7. American Academy of Dermatology: Criteria for the Diagnosis and Assessment of Atopic Dermatitis

Features to be considered in diagnosis of patients with atopic dermatitis:

Essential Features—Must be present:

- pruritus
- eczema (acute, subacute, chronic)
    - typical morphology and age-specific patterns*
    - chronic or relapsing history

*Patterns include:

1) facial, neck, and extensor involvement in infants and children
2) current or previous flexural lesions in any age group
3) sparing of the groin and axillary regions

Important Features—Seen in most cases, adding support to the diagnosis:

- early age of onset
- atopy
    - personal and/or family history
    - Immunoglobulin E reactivity
- xerosis

Associated Features—These clinical associations help to suggest the diagnosis of atopic dermatitis but are too nonspecific to be used for defining or detecting atopic dermatitis for research and epidemiologic studies:

- atypical vascular responses (for example, facial pallor, white dermographism, delayed blanch response)
- keratosis pilaris/pityriasis alba/hyperlinear palms/ichthyosis
- ocular/periorbital changes
- other regional findings (for example, perioral changes/periauricular lesions)
- perifollicular accentuation/lichenification/prurigo lesions

Exclusionary Features—It should be noted that a diagnosis of atopic dermatitis depends on excluding conditions, such as:

- scabies
- seborrheic dermatitis
- contact dermatitis (irritant or allergic)
- ichthyoses
- cutaneous T-cell lymphoma
- psoriasis
- photosensitivity dermatoses
- immune deficiency diseases
- erythroderma of other causes

Source: Eichenfield et al. 2014.

Appendix 8. Classification of Potency for Topical Corticosteroids

Potency	Class	Topical Corticosteroid	Formulation
Ultra high	I	Clobetasol propionate	Cream 0.05%
		Diflorasone diacetate	Ointment 0.05%
High	II	Amcinonide	Ointment 0.1%
		Betamethasone dipropionate	Ointment 0.05%
		Desoximetasone	Cream or ointment 0.025%
		Fluocinonide	Cream, ointment or gel 0.05%
		Halcinonide	Cream 0.1%
	III	Betamethasone dipropionate	Cream 0.05%
		Betamethasone valerate	Ointment 0.1%
		Diflorasone diacetate	Cream 0.05%
		Triamcinolone acetonide	Ointment 0.1%
Moderate	IV	Desoximetasone	Cream 0.05%
		Fluocinolone acetonide	Ointment 0.025%
		Fludroxycortide	Ointment 0.05%
		Hydrocortisone valerate	Ointment 0.2%
		Triamcinolone acetonide	Cream 0.1%
	V	Betamethasone dipropionate	Lotion 0.02%
		Betamethasone valerate	Cream 0.1%
		Fluocinolone acetonide	Cream 0.025%
		Fludroxycortide	Cream 0.05%
		Hydrocortisone butyrate	Cream 0.1%
		Hydrocortisone valerate	Cream 0.2%
		Triamcinolone acetonide	Lotion 0.1%
Low	VI	Betamethasone valerate	Lotion 0.05%
		Desonide	Cream 0.05%
		Fluocinolone acetonide	Solution 0.01%
	VII	Dexamethasone sodium phosphate	Cream 0.1%
		Hydrocortisone	Lotion, cream, or ointment 2.5%
		Hydrocortisone acetate	Cream 1%
		Methylprednisolone acetate	Cream 0.25%

Source: WHO (1997) and Tadicherla et al. 2009.

Appendix 9. Multiple Imputation

Placebo Multiple Imputation

The placebo multiple imputation (pMI) method will be used as a sensitivity analysis for the analysis of the primary efficacy endpoint, EASI75, as well as the key secondary endpoints at Week 16 (Visit 8). In the main analysis, all patients who either permanently discontinue study treatment or discontinue from the study for any reason at any time or who are rescued will be considered missing after discontinuation or rescue. The secondary ("as observed") analysis will be conducted for the primary endpoint, IGA (0,1) with ≥2-point improvement, and 4-point improvement from baseline in Itch Numeric Rating Scale (NRS) at Week 16. In this analysis, all patients who either permanently discontinue study treatment or discontinue from the study for any reason at any time will be considered to be missing after discontinuation.

The pMI assumes that the statistical behavior of drug- and placebo- treated patients after discontinuing study medication becomes that of placebo-treated patients. Multiple imputations are used to replace missing outcomes (IGA, EASI, SCORing Atopic Dermatitis (SCORAD), Itch NRS, Atopic Dermatitis Sleep Scale (ADSS) Item 2, and Skin Pain NRS scores) for drug- and placebo-treated patients who discontinued using multiple draws from the posterior predictive distribution estimated from the placebo arm. The binary outcomes of IGA (0,1), EASI75, EASI90, SCORAD75, and 4-point improvement from baseline in Itch NRS will then be derived from the imputed data.

Data are processed sequentially by repeatedly calling SAS® PROC MI to impute missing outcomes at visits t=1,.., T.

1. *Initialization:* Set t=0 (baseline visit)
2. *Iteration:* Set t=t+1. Create a data set combining records from drug- and placebo-treated patients with columns for covariates **X** and outcomes at visits 1,..,t with outcomes for all drug-treated patients set to missing at visit t and set to observed or imputed values at visits 1,..,t-1.
3. *Imputation:* Run Bayesian regression in SAS PROC MI on this data to impute missing values for visit t using previous outcomes for visits 1 to t-1 and baseline covariates. Note that only placebo data will be used to estimate the imputation model since no outcome is available for drug-treated patients at visit t.
4. Replace imputed data for all drug-treated patients at visit t with their observed values, whenever available up to permanent study drug discontinuation and/or rescue (if censoring on rescue). If t < T then go to Step 2, otherwise proceed to Step 5.
5. Repeat steps 1-4, m times with different seed values to create m imputed complete data sets.

Analysis: For each completed data set use the model as would have been applied had the data been complete for the continuous outcome. For the primary and secondary key efficacy endpoints of IGA (0,1), EASI75, EASI90, SCORAD75, and 4-point improvement from baseline in Itch NRS, the binary outcomes will be derived from the imputed data for each patient before fitting into the analysis model. A logistic regression model will be applied.

The number of imputed data sets will be m=100 and a 6-digit seed value will be prespecified for each analysis. Within the program, the seed will be used to generate the m seeds needed for imputation. The initial seed values are given below:

Analysis	Seed value
Proportion of patients achieving IGA of 0 or 1 with a ≥2−point improvement from baseline at Week 16, with data up to rescue	123450
Proportion of patients achieving EASI75 at 16 weeks, with data up to rescue	123451
Proportion of patients achieving EASI90 at 16 weeks, with data up to rescue	123451
Mean percent change from baseline in EASI score at 16 weeks, with data up to rescue	123451
Proportion of patients achieving SCORAD75 at 16 week, with data up to rescue	123452
Proportions of patients achieving a 4−point improvement from baseline in Itch NRS at Week 16, with data up to rescue	123453
Mean change from baseline in Skin Pain NRS at Week 16, with data up to rescue	123454
Mean change from baseline in the score of Item 2 of the ADSS at Week 16, with data up to rescue	123455
Proportion of patients achieving IGA of 0 or 1 with a ≥2−point improvement from baseline at Week 16, including data after rescue	123456
Proportion of patients achieving EASI75 at 16 weeks, including data after rescue	123457
Proportions of patients achieving a 4−point improvement from baseline in Itch NRS at Week 16, including data after rescue	123458

The final inference on treatment difference is conducted from the multiple datasets using Rubin's combining rules, as implemented in SAS PROC MI ANALYZE.

Thus, in the effectiveness context, pMI assumes no pharmacological benefit of the drug after dropout but is a more reasonable approach than modified last observed carried forward because unlike modified last observed carried forward it accounts for uncertainty of imputation, and therefore does not underestimate standard errors, and it limits bias by taking into account study/placebo effects. In the efficacy context pMI is a specific form of a missing not at random analysis expected to yield a conservative estimate of efficacy.

Tipping Point Analyses

To investigate the missing data mechanism, sensitivity analysis using multiple imputation (MI) under the missing not at random assumption will be provided for the following primary and major secondary objectives:

- EASI75 at Week 16, baricitinib 2-mg compared to placebo.
- IGA (0,1) with ≥2-point improvement at Week 16, baricitinib 2-mg compared to placebo,

- Itch NRS 4-point improvement from baseline to Week 16, baricitinib 2-mg compared to placebo.

All available data will be included in the analyses, excluding data that occurs after rescue or after permanent discontinuation of study drug. All patients in the intent-to-treat population are included.

Within each analysis, a most extreme case will be considered, in which all missing data for patients randomized to baricitinib 1-mg or 2-mg will be imputed using the worst possible result and all missing data for patients randomized to placebo will be imputed with the best possible result. Treatment differences will be analyzed using logistic regression or analysis of covariance (ANCOVA) (Section 10.3.1) as appropriate.

For continuous variables, the following process will be used to determine the tipping point:

1. To handle intermittent missing visit data, a Markov chain Monte Carlo method (SAS Proc MI with MCMC option) will be used to create a monotone missing pattern.
2. A set of Bayesian regressions (using SAS Proc MI with MONOTONE option) will be used for the imputation of monotone dropouts. Starting from the first visit with at least 1 missing value, the regression models will be fit sequentially with treatment as a fixed effect and values from the previous visits as covariates.
3. A delta score is added to all imputed scores at the primary time point for patients in the baricitinib treatment groups, thus worsening the imputed value. The delta score is capped for patients based on the range of the outcome measure being analyzed.
4. Treatment differences between baricitinib and placebo are analyzed for each imputed dataset using ANCOVA. Results across the imputed datasets are aggregated using SAS Proc MI ANALYZE to compute a p-value for the treatment comparisons for the given delta value.
5. Steps 3 and 4 are repeated, and the delta value added to the imputed baricitinib scores is gradually increased. The tipping point is identified as the delta value which leads to a loss of statistical significance (aggregated p-value >0.05) when evaluating baricitinib relative to the placebo group.

As a reference, for each delta value used in Steps 3-5, a fixed selection of delta values (ranging from slightly negative to slightly positive) will be added to imputed values in the placebo group, and Step 4 will be performed for the combination. This will result in a 2-d table, with the columns representing the delta values added to the imputed placebo responses, and the rows representing the delta values added to the imputed baricitinib responses. Separate 2-d tables will compare each baricitinib dose group to placebo.

A similar process will be used for the categorical variables:

1. Missing responses in the baricitinib groups will be imputed with a range of low response probabilities, including probabilities of 0, 0.1, and 0.2.

2. For missing responses in the placebo group, a range of responses probabilities (for example, probability=0, 0.2 … 1) will be used to impute the missing values. Multiple imputed datasets will be generated for each response probability.

3. Treatment differences between baricitinib and placebo will be analyzed for each imputed dataset using logistic regression. Results across the imputed datasets will be aggregated using SAS Proc MI ANALYZE to compute a p-value for the treatment comparisons for the given response probability. If the probability values do not allow for any variation between the multiple imputed datasets (for example, all missing responses in the placebo and baricitinib groups are imputed as responders and nonresponders, respectively), then the p-value from the single imputed dataset will be used.

The tipping point is identified as the response probability value within the placebo group that leads to a loss of statistical significance when evaluating baricitinib relative to placebo.

For tipping point analyses the number of imputed data sets will be m=100 and the seed values to start the pseudorandom number generator of SAS Proc MI (same values for MCMC option and for MONOTONE option) will be:

Analysis	Seed value
Proportion of patients achieving IGA of 0, 1 or 2 at Week 16, with data up to rescue	123461
Proportion of patients achieving EASI75 at 16 weeks, with data up to rescue	123462
Proportions of patients achieving a 4−point improvement from baseline in Itch NRS at Week 16, including data after rescue	123463

Appendix 10. Adjustment for Multiple Comparisons

The primary hypothesis $H_{2,1,0}$ will be tested at a 2-sided α=0.05. If the initial null hypothesis is not rejected, no further testing will be conducted as the α for that test is considered "spent" and cannot be passed to other endpoints. If $H_{2,1,0}$ is rejected, then α will be propagated to $H_{2,2,0}$. The testing process will continue with α propagated according to the weights on the corresponding edges displayed in Figure JAIW.2, as long as each hypothesis in the sequence can be rejected at its allocated α level. Each time a hypothesis is rejected, the graph will be updated to reflect the reallocation of α, which is considered "recycled" by Alosh et al. 2014. This iterative process of updating the graph and reallocating α will be repeated until all hypotheses have been tested or when no remaining hypotheses can be rejected at their corresponding α levels.

There will be no adjustment for multiple comparisons for any other analyses. Any changes to the details of the adjustment for multiple comparisons will be finalized in the statistical analysis plan prior to primary database lock.

Appendix 11. Investigator Global Assessment

Validated Investigator Global Assessment scale for Atopic Dermatitis

vIGA-AD™

Instructions:

The IGA score is selected using the descriptors below that best describe the overall appearance of the lesions at a given time point. It is not necessary that all characteristics under Morphological Description be present.

Score	Morphological Description
0 – Clear	No inflammatory signs of atopic dermatitis (no erythema, no induration/papulation, no lichenification, no oozing/crusting). Post-inflammatory hyperpigmentation and/or hypopigmentation may be present.
1 – Almost clear	Barely perceptible erythema, barely perceptible induration/papulation, and/or minimal lichenification. No oozing or crusting.
2 – Mild	Slight but definite erythema (pink), slight but definite induration/papulation, and/or slight but definite lichenification. No oozing or crusting.
3 – Moderate	Clearly perceptible erythema (dull red), clearly perceptible induration/papulation, and/or clearly perceptible lichenification. Oozing and crusting may be present.
4 – Severe	Marked erythema (deep or bright red), marked induration/papulation, and/or marked lichenification. Disease is widespread in extent. Oozing or crusting may be present.

Notes:

1. In indeterminate cases, please use extent to differentiate between scores.

For example:

- *Patient with marked erythema (deep or bright red), marked papulation and/or marked lichenification that is limited in extent, will be considered "3 – Moderate".*

2. Excoriations should not be considered when assessing disease severity.

Copyright ©2017 Eli Lilly and Company – Used with the permission of Eli Lilly and Company under a Creative Commons Attribution-NoDerivatives 4.0 International License - https://creativecommons.org/licenses/by-nd/4.0/

Appendix 12. Protocol Amendment I4V-MC-JAIW(c) Summary - A Multicenter, Randomized, Double-Blind, Placebo-Controlled, Phase 3 Study to Evaluate the Efficacy and Safety of Baricitinib in Adult Patients with Moderate to Severe Atopic Dermatitis

Overview

Protocol I4V-MC-JAIW, A Multicenter, Randomized, Double Blind, Placebo Controlled, Phase 3 Study to Evaluate the Efficacy and Safety of Baricitinib in Adult Patients with Moderate to Severe Atopic Dermatitis, has been amended. The new protocol is indicated by amendment (c) and will be used to conduct the study in place of any preceding version of the protocol.

Amendment Summary for Protocol I4V-MC-JAIW Amendment (c)

Section #	Description of Change	Brief Rationale
Section 1, 4, 5.1.2, 5.1.3, 9.1.1, 9.1.2.1, 9.1.2.2, 9.1.2.3, 9.1.2.4, 10.3.3.1, 10.3.3.2, 10.3.5.2 and Appendix 9	Primary efficacy endpoint was changed	IGA (0,1) at week 16 was assigned as a key secondary endpoint while EASI75 at week 16 became the primary endpoint. Both EASI 75 and IGA 0/1 are regulatory accepted endpoints for clinical trials in Atopic Dermatitis. EASI 75 was already included in the study as a key secondary endpoint. Based on the available data from completed BREEZE studies, there is greater power to test EASI 75 compared to IGA 0/1, thus exchanging the order of testing of IGA 0/1 and EASI 75 will optimize the multiplicity testing approach of the study.
	Baricitinib 1-mg was removed as a component of the primary objective .	Based on available data from completed BREEZE studies, baricitinib 1-mg was removed as a co-primary and will be considered with the secondary analyses only.
	Criteria for patient eligibility for enrollment into Study JAIX and language for related sections were updated.	To allow all patients participating in Study JAIW the opportunity to receive IP (2-mg QD), the study design has been updated to allow “responders” from Study JAIW to enter open-label Study JAIX after

Section #	Description of Change	Brief Rationale
		completing the full treatment period of Study JAIW. This change allows all patients enrolled in Study JAIW to be assessed for entering Study JAIX at some point at or after JAIW Visit 8 and will generate more safety and efficacy data at the 2-mg dose.
	Additional exploratory endpoints were included.	Based on results from completed studies with similar patient populations, additional exploratory endpoints for PROs have been included.
Section 5.4 and Section 11	Provided information regarding the EASI scale and updated the References list.	Incorporated empirical evidence of EASI scale reliability to support its' inclusion as the primary efficacy endpoint for this study.
Section 10.1, 10.3.1 and Appendix 10	Updated statistical model and estimated statistical power calculations.	Given the new primary endpoint for the study, the statement and the rationale for power calculation needed to be updated.
	Hypotheses graphical testing figure was updated.	With the removal of baricitinib 1-mg as a co-primary endpoint, null hypothesis testing needed to be revised to reflect the change in alpha.
	References to the primary endpoint of IGA were removed	References to IGA as the primary endpoint were removed to reflect the change in primary endpoint.

Revised Protocol Sections

Note:	Deletions have been identified by ~~strikethroughs~~. Additions have been identified by the use of underscore.

1. Synopsis

Objectives/Endpoints:

Objectives	Endpoints
Primary *This is a prespecified objective that will be adjusted for multiplicity.*	
To test the hypothesis that baricitinib ~~1-mg QD or~~ 2-mg QD is superior to placebo in the treatment of patients with moderate to severe 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
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

Statistical Analysis:

Treatment comparisons of discrete efficacy variables will be made using a logistic regression analysis with treatment, ~~and~~ baseline disease severity (IGA), and baseline score in the model. Region may be added to the model if patient numbers allow. The proportions and 95% 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, baseline disease severity (IGA), 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.

4. Objectives and Endpoints

Table JAIW.2 shows the objectives and endpoints of the study.

Table JAIW.2. Objectives and Endpoints

Objectives	Endpoints
Primary *This is a prespecified objective that will be adjusted for multiplicity.*	
To test the hypothesis that baricitinib ~~1 mg QD or~~ 2-mg QD is superior to placebo in the treatment of patients with moderate to severe 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
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)
Exploratory Objectives/Endpoints	
• 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, and HADS.	

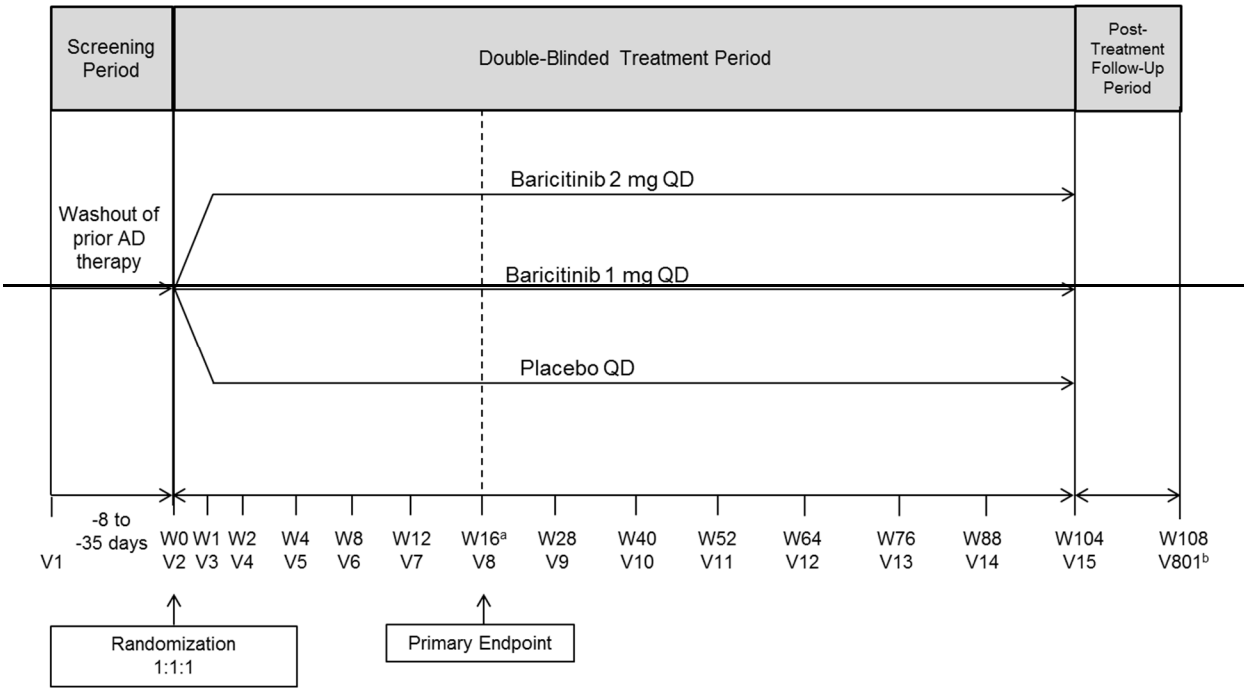
5.1.2 Period 2: Double-Blind Placebo-Controlled Treatment, Weeks 0 to 104

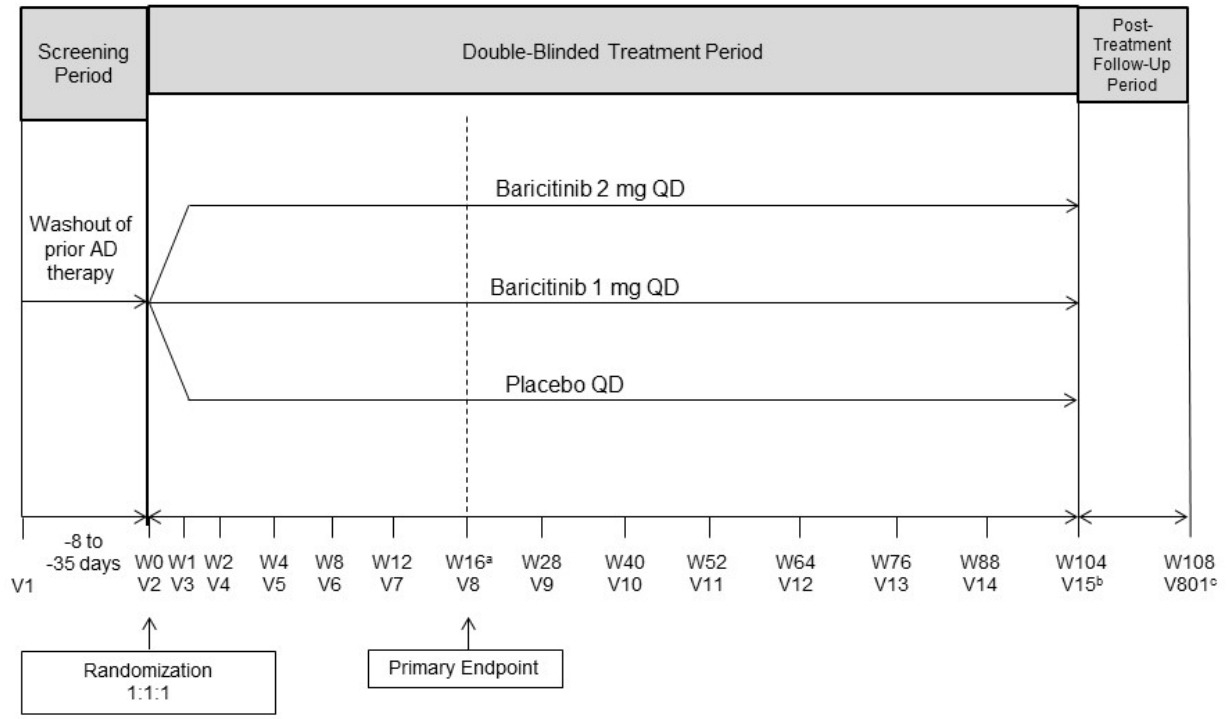
At Week 16, all patients who meet ~~the primary endpoint~~ ~~(~~IGA 0 or 1~~)~~ and who have not required rescue therapy before Week 16 will be allowed to continue on this study. All other patients will be discontinued from this study and may be eligible to enroll in a separate open-label study (Study 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 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.


Screening Period
Double-Blinded Treatment Period
Post-Treatment Follow-Up Period
Washout of prior AD therapy
Baricitinib 2 mg QD
Baricitinib 1 mg QD
Placebo QD
-8 to -35 days
V1
W0 V2
W1 V3
W2 V4
W4 V5
W8 V6
W12 V7
W16[a] V8
W28 V9
W40 V10
W52 V11
W64 V12
W76 V13
W88 V14
W104 V15
W108 V801[b]
Randomization 1:1:1
Primary Endpoint


Screening Period
Double-Blinded Treatment Period
Post-Treatment Follow-Up Period
Washout of prior AD therapy
Baricitinib 2 mg QD
Baricitinib 1 mg QD
Placebo QD
-8 to -35 days
V1
W0 V2
W1 V3
W2 V4
W4 V5
W8 V6
W12 V7
W16[a] V8
W28 V9
W40 V10
W52 V11
W64 V12
W76 V13
W88 V14
W104 V15[b]
W108 V801[c]
Randomization 1:1:1
Primary Endpoint

Abbreviations: AD = atopic dermatitis; QD = once daily; V = visit; W = week.

[a] At Week 16, all patients who ~~meet the primary endpoint~~ achieve an ~~(~~IGA 0 or 1~~)~~ and who have not required rescue therapy before Week 16 will be allowed to continue in this study. All other patients will be discontinued from this study and may be eligible to enroll in the separate open-label Study JAIX.

[b] Patients who complete this study will be eligible for assessment to enroll in open-label Study JAIX.

[c] Occurs approximately 28 days after the last dose of investigational product.

Figure JAIW.1. Illustration of study design for Clinical Protocol I4V-MC-JAIW.

5.4 Scientific Rationale for Study Design

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 ~~meet the primary endpoint of~~ achieve IGA of 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.

9.1.1 Primary Efficacy Assessments

Eczema Area and Severity Index scores: The EASI assesses extent of disease at 4 body regions and measures 4 clinical signs: (1) erythema, (2) induration/papulation, (3) excoriation, and (4) lichenification, each on a scale of 0 to 3. The EASI confers a maximum score of 72. The EASI evaluates 2 dimensions of AD: disease extent and clinical signs (Hanifin et al. 2001).

~~**Validated Investigator's Global Assessment for Atopic Dermatitis (vIGA-AD):** The IGA used in this study, the vIGA-AD (referred to as the IGA throughout the protocol), measures the IGA of the patient's overall severity of their AD, on the basis of a static, numeric 5-point scale from 0 (clear skin) to 4 (severe disease). The score is based on an overall assessment of the degree of erythema, papulation/induration, oozing/crusting, and lichenification.~~

9.1.2.1 ~~Eczema Area and Severity Index scores~~Validated Investigator's Global Assessment for Atopic Dermatitis (vIGA-AD)

The IGA used in this study, the vIGA-AD (referred to as the IGA throughout the protocol), measures the IGA of the patient's overall severity of their AD, on the basis of a static, numeric 5-point scale from 0 (clear skin) to 4 (severe disease). The score is based on an overall assessment of the degree of erythema, papulation/induration, oozing/crusting, and lichenification.

~~The EASI assesses extent of disease at 4 body regions and measures 4 clinical signs: (1) erythema, (2) induration/papulation, (3) excoriation, and (4) lichenification, each on a scale of 0 to 3. The EASI confers a maximum score of 72. The EASI evaluates 2 dimensions of AD: disease extent and clinical signs (Hanifin et al. 2001).~~

~~BSA affected by AD will be derived from data collected as part of the EASI assessment.~~

9.1.2.~~2~~3 SCORing Atopic Dermatitis

9.1.2.~~3~~4 Hospital Anxiety Depression Scale

10.1 Sample Size Determination

Study JAIW will aim to enroll approximately 450 patients aged ≥18 years. The proposed sample size will ensure ~~at least 90%~~ approximately 68% power to detect any difference between the baricitinib 2-mg and placebo treatment groups, assuming a 10% placebo and 2~~5~~0% baricitinib 2-mg response rate for the primary endpoint EASI75 using a Chi-squared test with a 2-sided α level of 0.0~~2~~5. The assumptions are based on what was observed in the Phase ~~2~~3 monotherapy Stud~~y~~ies JAH~~G~~L and JAHM.

Sample size and power estimates were obtained from nQuery® Advisor 7.0 ~~and through simulations following adjustments for multiple comparisons~~.

10.3.1 General Statistical Considerations

All tests of treatment effects will be conducted at a 2-sided α level of 0.05, unless otherwise stated. Treatment comparisons of discrete efficacy variables between baricitinib and placebo will be made using a logistic regression analysis with IGA baseline disease severity, baseline score, and treatment group in the model. Region may be added as an additional factor if there is a sufficient number of patients for rest-of-world and stratification has not resulted in empty strata. This will be finalized in the SAP. If appropriate, treatment by region interaction may be added to the model of the primary and key secondary variables as a sensitivity analysis. If this interaction is significant at a 2-sided 0.1 level, further inspection will be used to assess whether the interaction is quantitative (that is, the treatment effect is consistent in direction but not size of effect) or qualitative (the treatment is beneficial for some but not all regions). The percentages, difference in percentages, and 95% CI of the difference in percentages will be reported. The p-value from the Fisher's exact test will also be produced.

When evaluating continuous measures over time, a restricted maximum likelihood-based MMRM will be used. The model will include treatment, baseline disease severity (IGA), baseline score, visit, and treatment-by-visit interaction as fixed categorical effects and baseline

score and baseline score-by-visit interaction as fixed continuous effects. Geographic region may be added as an additional factor if appropriate. An unstructured (co)variance structure will be used to model the between- and within-patient errors. If this analysis fails to converge, other structures will be tested. The Kenward–Roger method will be used to estimate the degrees of freedom. Type III sums of squares for the LSMs will be used for the statistical comparison; 95% CI will also be reported. Contrasts will be set up within the model to test treatment groups at specific time points of interest. Further details on the use of MMRM will be described in the SAP.

Treatment comparisons of continuous efficacy and health outcome variables may also be made using ANCOVA with baseline disease severity (IGA), treatment group, and baseline ~~value~~score in the model. Geographic region may be added as an additional factor if appropriate. Type III tests for LSM will be used for statistical comparison between treatment groups. The LSM difference, standard error, p-value, and 95% CI may also be reported. The method used to handle missing data will be specified in the SAP.

Primary Null Hypothesis:

$H_{2,1,0}$: Proportion of baricitinib 2-mg patients achieving EASI75 is less than or equal to the proportion of placebo patients achieving EASI75 at Week 16

~~$H_{2,1,0}$: Proportion of baricitinib 2-mg patients achieving IGA of 0 or 1 and ≥2-point improvement from baseline at Week 16 is less than or equal to the proportion of placebo patients achieving IGA of 0 or 1 and ≥2-point improvement from baseline at Week 16~~

~~$H_{1,1,0}$: Proportion of baricitinib 1-mg patients achieving IGA of 0 or 1 and ≥2-point improvement from baseline at Week 16 is less than or equal to the proportion of placebo patients achieving IGA of 0 or 1 and ≥2-point improvement from baseline at Week 16~~

Key Secondary Null Hypotheses:

~~$H_{2,2,0}$: Proportion of baricitinib 2-mg patients achieving EASI75 is less than or equal to the proportion of placebo patients achieving EASI75 at Week 16~~

~~$H_{1,2,0}$:~~ Proportion of baricitinib 1-mg patients achieving EASI75 is less than or equal to the proportion of placebo patients achieving EASI75 at Week 16

$H_{2,2,0}$: Proportion of baricitinib 2-mg patients achieving IGA of 0 or 1 and >2-point improvement from baseline at Week 16 is less than or equal to the proportion of placebo patients achieving IGA of 0 or 1 and ≥2-point improvement from baseline at Week 16

$H_{2,10,0}$: Mean change from baseline in the score of Item 2 of the ADSS for baricitinib 2-mg is less than or equal to the mean change from baseline in the score of Item 2 of the ADSS for placebo at Week 16

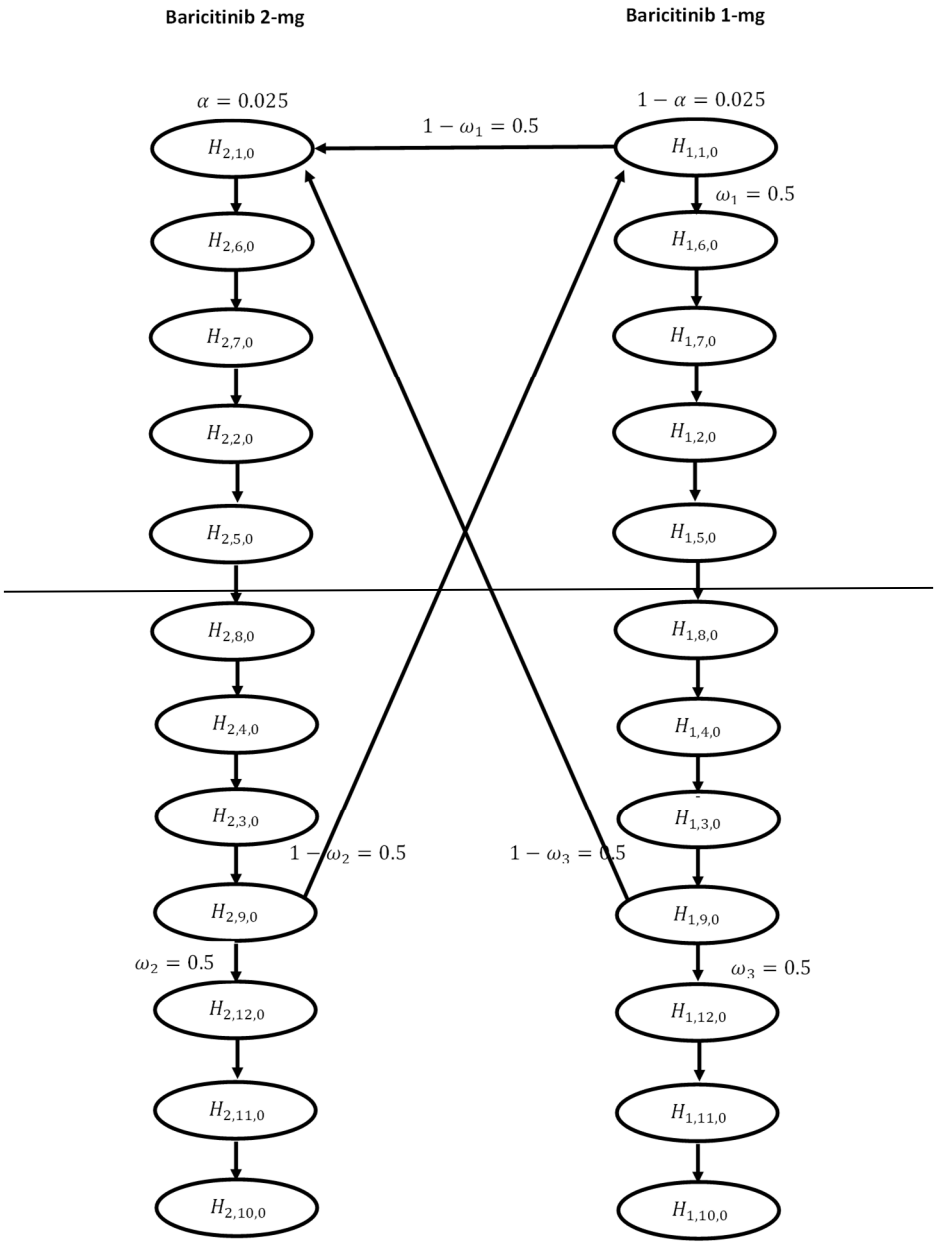
$H_{2,11,0}$: Mean change from baseline in the score of Item 2 of the ADSS for baricitinib 2-mg is less than or equal to the mean change from baseline in the score of Item 2 of the ADSS for placebo at Week 1~~6~~

$H_{1,1,0}$: Proportion of baricitinib 1-mg patients achieving EASI75 is less than or equal to the proportion of placebo patients achieving EASI75 at Week 16

$H_{1,2,0}$: Proportion of baricitinib 1-mg patients achieving IGA of 0 or 1 and >2-point improvement from baseline at Week 16 is less than or equal to the proportion of placebo patients achieving IGA of 0 or 1 and >2-point improvement from baseline at Week 16

$H_{1,10,0}$: Mean change from baseline in the score of Item 2 of the ADSS for baricitinib 1-mg is less than or equal to the mean change from baseline in the total score of the ADSS for placebo at Week 16

$H_{1,11,0}$: Mean change from baseline in the score of Item 2 of the ADSS for baricitinib 1-mg is less than or equal to the mean change from baseline in the total score of the ADSS for placebo at Week 16


Baricitinib 2-mg
Baricitinib 1-mg
$\alpha = 0.025$
$1 - \alpha = 0.025$
$1 - \omega_1 = 0.5$
$\omega_1 = 0.5$
$H_{2,1,0}$
$H_{2,6,0}$
$H_{2,7,0}$
$H_{2,2,0}$
$H_{2,5,0}$
$H_{2,8,0}$
$H_{2,4,0}$
$H_{2,3,0}$
$H_{2,9,0}$
$H_{2,12,0}$
$H_{2,11,0}$
$H_{2,10,0}$
$H_{1,1,0}$
$H_{1,6,0}$
$H_{1,7,0}$
$H_{1,2,0}$
$H_{1,5,0}$
$H_{1,8,0}$
$H_{1,4,0}$
$H_{1,3,0}$
$H_{1,9,0}$
$H_{1,12,0}$
$H_{1,11,0}$
$H_{1,10,0}$
$1 - \omega_2 = 0.5$
$1 - \omega_3 = 0.5$
$\omega_2 = 0.5$
$\omega_3 = 0.5$

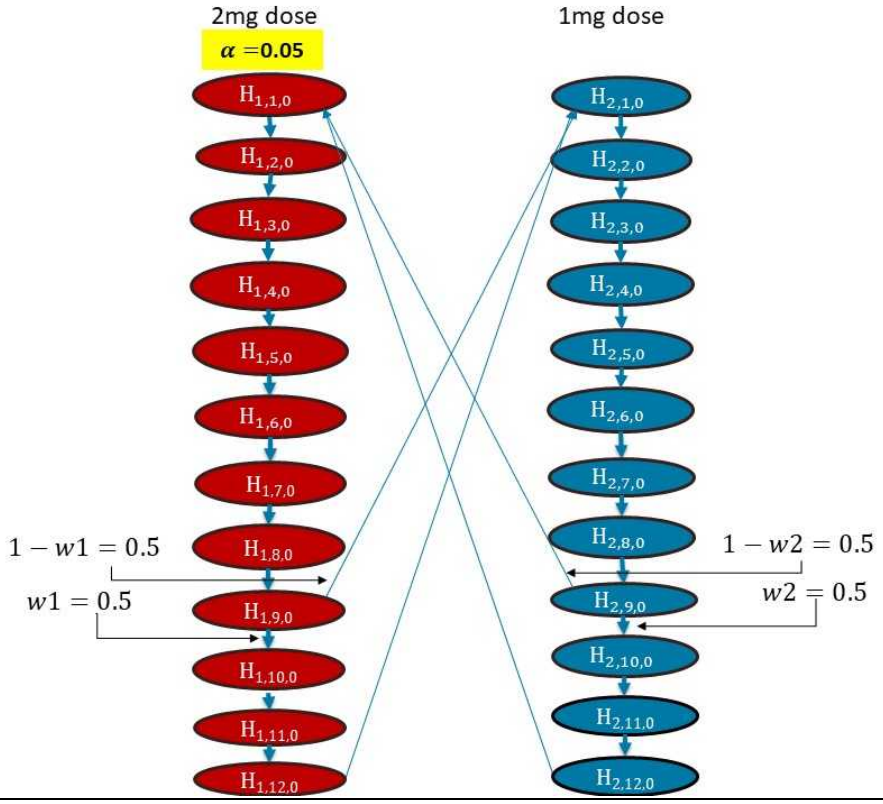


Figure JAIW.2. Study I4V-MC-JAIW scheme structure.

10.3.3.1 Primary Analyses

The primary efficacy measure is the ~~binary outcome of response defined as IGA score of 0 or 1 (clear or almost clear skin) and ≥2-point improvement from baseline~~ categorical outcome of EASI75 at Week 16 for baricitinib 2-mg. EASI75 is defined as having an improvement of at least 75% from baseline. Primary analysis will be conducted using a logistic regression as described above with treatment and stratification variables, which include ~~(~~IGA baseline disease severity, baseline score and region, as appropriate~~)~~ in the model. NRI for missing data as described above will be used.

Additional analysis of the primary efficacy outcome will include analyzing the outcome as observed, that is, whether rescue medication was given.

10.3.3.2 Secondary Analyses

The following secondary categorical outcomes will be analyzed in a similar manner as the primary, that is, using the same logistic regression model. This applies to both baricitinib 2-mg vs placebo, and baricitinib 1-mg vs placebo unless stated otherwise. NRI will be used for these analyses unless otherwise noted.

- ~~EASI75 at Week 16. EASI75 is defined as having an improvement of at least 75% from baseline~~. IGA score of 0 or 1 (clear or almost clear skin) and ≥2-point improvement from baseline at Week 16. Besides NRI, this outcome will also be analyzed using observed cases, that is, whether rescue medication was given.
- EASI75 at Week 16 for baricitinib 1-mg vs placebo. EASI75 is defined as having an improvement of at least 75% from baseline.

10.3.5.2 Subgroup Analyses

To assess whether the treatment effect is similar across subgroups for the primary efficacy outcome, a logistic model will be used and will include treatment, stratification variables (IGA (3,4) at baseline, baseline disease severity and region) as appropriate, the subgroup variable (for example, sex), and the subgroup-by-treatment interaction. If the interaction is statistically significant at α level of 0.10, the nature of the interaction will be explored, that is, within each subgroup the treatment effect will be estimated. Similarly, for the continuous variables of EASI, the MMRM model will include additional variables for subgroup and the subgroup-by-treatment interaction.

11. References

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.

Appendix 9. Multiple Imputation

The placebo multiple imputation (pMI) method will be used as a sensitivity analysis for the analysis of the primary efficacy endpoint, ~~IGA (0,1) with ≥2-point improvement~~ EASI75 as well as the key secondary endpoints at Week 16 (Visit 8). In the main analysis, all patients who either permanently discontinue study treatment or discontinue from the study for any reason at any time or who are rescued will be considered missing after discontinuation or rescue. The secondary ("as observed") analysis will be conducted for the primary endpoint, ~~Eczema Area and Severity Index (EASI75)~~ IGA (0,1) with ≥2-point improvement, and 4-point improvement from baseline in Itch Numeric Rating Scale (NRS) at Week 16 ~~and will be referred to as the "pMI as observed" analysis~~. In this analysis, ~~A~~all patients who either permanently discontinue study treatment or discontinue from the study for any reason at any time will be considered to be missing after discontinuation.

The pMI assumes that the statistical behavior of drug- and placebo- treated patients after discontinuing study medication becomes that of placebo-treated patients ~~which is the rationale for "pMI as observed". The main analysis will be conducted to keep in line with the other analyses and to assess the impact of pMI with other imputation methods~~. Multiple imputations are used to replace missing outcomes (IGA, EASI, SCORing Atopic Dermatitis (SCORAD), Itch NRS, Atopic Dermatitis Sleep Scale (ADSS) Item 2, and Skin Pain NRS scores) for drug- and placebo-treated patients who discontinued using multiple draws from the posterior predictive distribution estimated from the placebo arm. The binary outcomes of IGA (0,1), EASI75, EASI90, SCORAD75, and 4-point improvement from baseline in Itch NRS will then be derived from the imputed data.

Tipping Point Analyses

To investigate the missing data mechanism, sensitivity analysis using multiple imputation (MI) under the missing not at random assumption will be provided for the following primary and major secondary objectives:

- EASI75 at Week 16, baricitinib 2-mg compared to placebo.
- IGA (0,1) with ≥2-point improvement at Week 16, baricitinib 2-mg compared to placebo,
- ~~EASI mean percent change from baseline to Week 16, baricitinib 2-mg compared to placebo,~~
- Itch NRS 4-point improvement from baseline to Week 16, baricitinib 2-mg compared to placebo.

Appendix 10. Adjustment for Multiple Comparisons

The primary hypothesiesis $H_{2,1,0}$ ~~and $H_{1,1,0}$~~ will ~~each~~ be tested at a 2-sided $\alpha=0.0$~~2~~5. If the initial null hypothesiesis ~~are~~ is not rejected, no further testing will be conducted as the α for that test is considered "spent" and cannot be passed to other endpoints. If $H_{2,1,0}$ is rejected, then α will be propagated to $H_{2,2,0}$. ~~In addition, if $H_{1,1,0}$ is rejected, then α will be split equally and propagated to $H_{2,1,0}$ and $H_{1,2,0}$.~~ The testing process will continue ~~in both branches (starting with $H_{2,2,0}$ and $H_{1,2,0}$ respectively),~~ with α propagated according to the weights on the corresponding edges displayed in Figure JAIW.2, as long as each hypothesis in the sequence can be rejected at its allocated α level. Each time a hypothesis is rejected, the graph will be updated to reflect the reallocation of α, which is considered "recycled" by Alosh et al. 2014. This iterative process of updating the graph and reallocating α will be repeated until all hypotheses have been tested or when no remaining hypotheses can be rejected at their corresponding α levels.

Leo Document ID = a913b391-57dd-4a5a-aed2-1b334f4cfa69

Approver: PPD

Approval Date & Time: 10-Oct-2019 09:29:01 GMT

Signature meaning: Approved

Approver: PPD

Approval Date & Time: 10-Oct-2019 09:58:01 GMT

Signature meaning: Approved