Protocol I4V-MC-JAHZ(a)
A Randomized, Double-Blind, Placebo-Controlled, Parallel-Group, Phase 3 Study of Baricitinib in Patients with Systemic Lupus Erythematosus

EUDRA CTA 2017-005026-37

Confidential Information

Baricitinib (LY3009104)

Eli Lilly and Company
Indianapolis, Indiana USA 46285

Protocol Electronically Signed and Approved by Lilly: 25 April 2018
Amendment (a) Electronically Signed and Approved by Lilly
on approval date provided below.

Approval Date: 07-Dec-2018 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 Randomized, Double-Blind, Placebo-Controlled, Parallel-Group, Phase 3 Study of Baricitinib in Patients with systemic lupus erythematosus (SLE).

Rationale:
Systemic lupus erythematosus is a chronic, often debilitating, multisystem, autoimmune disease that is characterized by the presence of autoreactive B cells and elevated autoantibodies, which directly damage the body's cells and tissues. Systemic lupus erythematosus can affect multiple organ systems simultaneously or sequentially, and follows a highly variable clinical course where periods of relatively stable disease are followed by flares and/or periods of persistently active disease; all of which can ultimately lead to irreversible damage to tissues and organ systems.

Baricitinib is an oral, reversible, selective inhibitor of Janus kinase (JAK)1 and JAK2 (Fridman et al. 2010). This activity profile suggests that baricitinib may inhibit cytokines implicated in SLE, most notably type I interferon (IFN; JAK1/tyrosine kinase [TYK]2), interleukin (IL-6; JAK1/JAK2/TYK2), and type II IFNγ, as well as other cytokines that may have a role in SLE, including IL-23 (JAK2/TYK2), granulocyte–-macrophage colony stimulating factor (JAK2/JAK2) and IL-12 (JAK2/TYK2). In a recently completed Phase 2 study (I4V-MC-JAHH [JAHH]), baricitinib demonstrated clinical efficacy in patients with SLE. Baricitinib plus standard of care was superior to placebo plus standard of care in the proportion of patients achieving remission of rash and/or arthritis as defined by the Systemic Lupus Erythematosus Disease Activity Index 2000 (SLEDAI-2K), as well as the proportion of patients achieving a Systemic Lupus Erythematosus Responder Index-4 (SRI-4) response at Week 24.

Given the efficacy of baricitinib demonstrated in clinical trials for treating autoimmune/autoinflammatory diseases involving joints, skin, and kidney (including SLE), the acceptable safety profile of baricitinib observed through the current stage of development, and a continuing unmet medical need in patients with SLE, there is a compelling rationale for the initiation of a Phase 3 program to evaluate baricitinib in treatment of SLE.

Objective(s)/Endpoints:

Objectives	Endpoints
Primary	
To evaluate the effect of baricitinib 4-mg QD and background standard-of-care (SoC) therapy compared with placebo and SoC on SLE disease activity.	Proportion of patients achieving an SRI-4 response at Week 52, defined as: • Reduction of ≥4 points from baseline in SLEDAI-2K score; and • No new British Isles Lupus Assessment Group (BILAG) A or no more than 1 new BILAG B disease activity score; and • No worsening (defined as an increase of ≥0.3 points [10 mm] from baseline) in the Physician's Global Assessment of Disease Activity.
Major Secondary	
• To evaluate the effect of baricitinib 4-mg QD plus SoC compared to placebo plus SoC on SLE disease activity.	• Proportion of patients achieving an SRI-4 response at Week 24. • Proportion of patients achieving a lupus low disease activity state (LLDAS) response at Week 52
• To evaluate the corticosteroid sparing effect of baricitinib 4-mg QD plus SoC compared to placebo plus SoC.	• Proportion of patients receiving >7.5 mg prednisone (or equivalent) at baseline able to decrease dose by ≥25% to a prednisone equivalent dose of ≤7.5 mg/day maintained between Week 40 and Week 52.
• To evaluate the effect of baricitinib 4-mg QD plus SoC compared to placebo plus SoC on SLE flares.	• Time to first severe flare over 52 weeks.
• To evaluate the effect of baricitinib 4-mg QD plus SoC compared to placebo plus SoC on patient-reported outcomes (PROs).	• Change from baseline in Worst Pain NRS at Week 52. • Change from baseline in FACIT-Fatigue total score at Week 52.
• To evaluate the effect of baricitinib 2-mg QD plus SoC compared to placebo plus SoC on SLE disease activity.	• Proportion of patients achieving an SRI-4 response at Week 52. • Proportion of patients achieving an SRI-4 response at Week 24. • Proportion of patients achieving a lupus low disease activity state (LLDAS) response at Week 52
• To evaluate the corticosteroid sparing effect of baricitinib 2-mg QD plus SoC compared to placebo plus SoC.	• Proportion of patients receiving >7.5 mg prednisone (or equivalent) at baseline able to decrease dose by ≥25% to a prednisone equivalent dose of ≤7.5 mg/day maintained between Week 40 and Week 52.
• To evaluate the effect of baricitinib 2-mg QD plus SoC compared to placebo plus SoC on SLE flares.	• Time to first severe flare over 52 weeks.
• To evaluate the effect of baricitinib 2-mg QD plus SoC compared to placebo plus SoC on patient-reported outcomes (PROs).	• Change from baseline in Worst Pain NRS at Week 52. • Change from baseline in FACIT-Fatigue total score at Week 52

Abbreviations: BILAG = British Isles Lupus Assessment Group; FACIT-Fatigue = Functional Assessment of Chronic Illness Therapy-Fatigue; LLDAS = Lupus Low Disease Activity State; NRS = Numeric Rating Scale; QD = once daily; SLE = Systemic Lupus Erythematosus; SLEDAI-2K = Systemic Lupus Erythematosus Disease Activity Index 2000; SoC = standard of care; SRI-4 = Systemic Lupus Erythematosus Responder Index-4.

Summary of Study Design:

Study I4V-MC-JAHZ (JAHZ) is a Phase 3, multicenter, randomized, double-blind, parallel-group, placebo-controlled, outpatient, 52-week study evaluating the efficacy and safety of baricitinib 4-mg and 2-mg in patients with SLE receiving standard therapy. Background standard therapies for SLE include corticosteroids, antimalarials, and immunosuppressants.

Treatment Arms and Duration:

At Week 0, patients will be randomized to 1 of 3 treatment groups: placebo once daily (QD), baricitinib 2-mg QD, or baricitinib 4-mg QD. The study duration will be up to 62 weeks (Screening Period: up to 6 weeks prior to randomization; Double-Blinded Treatment Period: 52 weeks; Follow-up Period: approximately 4 weeks after the last dose of investigational product).

Number of Patients:

The study will include approximately 750 patients with SLE who will be randomized 1:1:1 to receive placebo QD, baricitinib 2-mg QD, or baricitinib 4-mg QD (250 patients in each treatment group).

Statistical Analysis:

Unless otherwise specified, the efficacy and health outcome analyses will be conducted on the modified intent-to-treat (mITT) population, which includes randomized patients who receive at least 1 dose of investigational product. Safety analyses will be conducted on those patients who receive at least 1 dose of investigational product who do not discontinue at the first postbaseline visit for the reason “Lost to Follow-up.”

Treatment comparisons of categorical efficacy and health outcomes variables will be made using a logistic regression analysis with Firth correction (Firth 1993; Heinze and Schemper 2001). Baseline disease activity, baseline corticosteroid dose, region, and treatment group will be explanatory variables in the model. The p-value and 95% confidence interval (CI) for the odds ratio from the logistic regression model will be used for primary statistical inference. Patients who discontinue treatment early, increased use of corticosteroids above the baseline dose other than the permitted burst, or have initiation of or an increase above the baseline dose in immunosuppressant or antimalarial treatment any time after baseline will be defined as nonresponders and the data will be imputed using the nonresponder imputation (NRI) method.

Treatment comparisons of continuous efficacy and health outcomes variables will be made using a restricted maximum likelihood-based mixed model for repeated measures (MMRM) analysis. The model will include treatment, baseline disease activity, baseline corticosteroid dose, region,

visit, treatment-by-visit interaction as fixed categorical effects, and baseline value and baseline value-by-visit interaction as fixed continuous effects to estimate change from baseline across postbaseline visits. Type III tests for the least squares (LS) means will be used for the statistical comparisons.

Time to first flare will be analyzed using a Cox proportional hazards model with treatment group, baseline disease activity, baseline corticosteroid dose, and region fitted as explanatory variables. Hazard ratios, 95% CI and p-values will be presented for treatment comparisons to placebo.

2. Schedule of Activities

Table JAHZ.1. Schedule of Activities

	Screening	Double-Blinded Treatment																Post-Treatment Follow-up
Visit Number	**V1**	**V2[a]**	**V3**	**V4**	**V5**	**V6**	**V7**	**V8**	**V9**	**V10**	**V11**	**V12**	**V13**	**V14**	**V15**	**V16[b]**	**ET[c]**	**V801[d]**
Study Week		**0**	**1**	**4**	**8**	**12**	**16**	**20**	**24**	**28**	**32**	**36**	**40**	**44**	**48**	**52**	**any**	**56 or last dose + 4 weeks**
Study Day	**-42 to -3**	**1**	**8 ± 3**	**29 ± 4**	**57 ± 4**	**85 ± 4**	**113 ± 4**	**141 ± 4**	**169 ± 4**	**197 ± 4**	**225 ± 4**	**253 ± 4**	**281 ± 4**	**309 ± 4**	**337 ± 4**	**365 ± 4**	**any**	**last dose + 28 ± 5 days**
Procedures and Assessments																		
Informed consent	X																	
Demographics	X																	
Medical history	X																	
Physical examination[e]	X																	
Symptom-directed physical examination[e]		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Height		X																
Weight	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Waist circumference		X				X			X			X				X	X	
Vital signs (BP[f], pulse, temperature)	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Substance use (caffeine, alcohol, tobacco)	X																	
ECG[g]	X																	
Administer TB test[h]	X																	
Read PPD test (if applicable)[i]	X																	
Chest x-ray[j]	X																	
Confirm SLE Criteria[k]	X																	
Previous SLE therapy	X	X																

	Screening	Double-Blinded Treatment																Post-Treatment Follow-up
Visit Number	**V1**	**V2[a]**	**V3**	**V4**	**V5**	**V6**	**V7**	**V8**	**V9**	**V10**	**V11**	**V12**	**V13**	**V14**	**V15**	**V16[b]**	**ET[c]**	**V801[d]**
Study Week		**0**	**1**	**4**	**8**	**12**	**16**	**20**	**24**	**28**	**32**	**36**	**40**	**44**	**48**	**52**	**any**	**56 or last dose + 4 weeks**
Study Day	**-42 to -3**	**1**	**8 ± 3**	**29 ± 4**	**57 ± 4**	**85 ± 4**	**113 ± 4**	**141 ± 4**	**169 ± 4**	**197 ± 4**	**225 ± 4**	**253 ± 4**	**281 ± 4**	**309 ± 4**	**337 ± 4**	**365 ± 4**	**any**	**last dose + 28 ± 5 days**
Review inclusion/exclusion criteria	X	X																
Preexisting conditions	X	X																
Adverse events		X	X	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	X	X
C-SSRS[l]/Self-Harm Supplement Form	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Self-Harm Follow-up Form[m]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
QIDS-SR16	X	X				X			X			X				X	X	X
Randomization[n]		X[n]																
Log in IWRS	X	X[n]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Investigational product dispensed		X	X	X	X	X	X	X	X	X	X	X	X	X	X			
Investigational product returned and compliance assessed			X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
ePRO (patient diary) dispensed/returned[o,p]	X[o]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X[o]	X[o]	X[o]
Physician-Completed Scales[q]																		
SLEDAI-2K	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
BILAG-2004	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Physician's Global Assessment of Disease Activity	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
SLEDAI Flare Index		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

	Screening	Double-Blinded Treatment																Post-Treatment Follow-up
Visit Number	**V1**	**V2[a]**	**V3**	**V4**	**V5**	**V6**	**V7**	**V8**	**V9**	**V10**	**V11**	**V12**	**V13**	**V14**	**V15**	**V16[b]**	**ET[c]**	**V801[d]**
Study Week		**0**	**1**	**4**	**8**	**12**	**16**	**20**	**24**	**28**	**32**	**36**	**40**	**44**	**48**	**52**	**any**	**56 or last dose + 4 weeks**
Study Day	**-42 to -3**	**1**	**8 ± 3**	**29 ± 4**	**57 ± 4**	**85 ± 4**	**113 ± 4**	**141 ± 4**	**169 ± 4**	**197 ± 4**	**225 ± 4**	**253 ± 4**	**281 ± 4**	**309 ± 4**	**337 ± 4**	**365 ± 4**	**any**	**last dose + 28 ± 5 days**
SLICC/ACR Damage Index		X														X	X	
CLASI		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Tender/swollen joint count (28 joints)	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Patient Questionnaires																		
Worst Pain NRS[p]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Worst Joint Pain NRS[p]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Worst Fatigue NRS[p]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Joint Stiffness Duration[p]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Joint Stiffness Severity NRS[p]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Patient's Global Impression of Severity (PGI-S)[p]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Patient's Global Impression of Change (PGI-C)[r]									X							X	X	
SF-36v2[r]		X				X			X			X				X	X	
FACIT-Fatigue[r]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
WPAI-Lupus[r]		X		X		X			X			X				X	X	
EQ-5D-5L[r]		X		X		X		X	X	X		X			X	X	X	
Laboratory Tests																		
TSH	X																	
HIV	X																	

	Screening	Double-Blinded Treatment																Post-Treatment Follow-up
Visit Number	**V1**	**V2[a]**	**V3**	**V4**	**V5**	**V6**	**V7**	**V8**	**V9**	**V10**	**V11**	**V12**	**V13**	**V14**	**V15**	**V16[b]**	**ET[c]**	**V801[d]**
Study Week		**0**	**1**	**4**	**8**	**12**	**16**	**20**	**24**	**28**	**32**	**36**	**40**	**44**	**48**	**52**	**any**	**56 or last dose + 4 weeks**
Study Day	**-42 to -3**	**1**	**8 ± 3**	**29 ± 4**	**57 ± 4**	**85 ± 4**	**113 ± 4**	**141 ± 4**	**169 ± 4**	**197 ± 4**	**225 ± 4**	**253 ± 4**	**281 ± 4**	**309 ± 4**	**337 ± 4**	**365 ± 4**	**any**	**last dose + 28 ± 5 days**
Hepatitis C virus (HCV) antibody testing[s]	X																	
Hepatitis B testing (HBsAg, HBcAb, HBsAb)	X																	
HBV DNA[t]	X					X			X			X			X	X	X	X
Serum pregnancy test[u]	X																	
Urine pregnancy test[u]		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
FSH[v]	X																	
Clinical chemistry[w]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Fasting lipid panel[x]		X				X			X			X				X		
Hematology	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Iron studies (iron, TIBC, and ferritin)	X					X			X			X				X		
Urinalysis[bb]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Urine Creatinine and Protein, Ratio	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
ANA	X															X		
anti-dsDNA	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
anti-Smith	X																	
Autoantibodies (anti-Smith, anti-RNP, anti-SSA/Ro, anti-SSB/La)		X				X			X			X				X		

	Screening	Double-Blinded Treatment																Post-Treatment Follow-up
Visit Number	**V1**	**V2[a]**	**V3**	**V4**	**V5**	**V6**	**V7**	**V8**	**V9**	**V10**	**V11**	**V12**	**V13**	**V14**	**V15**	**V16[b]**	**ET[c]**	**V801[d]**
Study Week		**0**	**1**	**4**	**8**	**12**	**16**	**20**	**24**	**28**	**32**	**36**	**40**	**44**	**48**	**52**	**any**	**56 or last dose + 4 weeks**
Study Day	**-42 to -3**	**1**	**8 ± 3**	**29 ± 4**	**57 ± 4**	**85 ± 4**	**113 ± 4**	**141 ± 4**	**169 ± 4**	**197 ± 4**	**225 ± 4**	**253 ± 4**	**281 ± 4**	**309 ± 4**	**337 ± 4**	**365 ± 4**	**any**	**last dose + 28 ± 5 days**
Anti-cardiolipin antibodies IgG, IgA, IgM		X				X			X			X				X	X	X
Anti-beta-2 glycoprotein-I (IgG, IgM)		X				X			X			X				X	X	X
Lupus anticoagulant (dRVVT)		X				X			X			X				X	X	X
Complement (C3 and C4)	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Immunoglobulins (IgG, IgA, IgM)		X		X		X			X			X				X	X	X
Lymphocyte subsets (T, B, NK, and T cell subsets)		X		X		X			X			X				X	X	X
Exploratory storage samples (serum and plasma)		X		X		X			X							X		
Exploratory storage samples (urine)		X		X		X			X							X		
Exploratory samples for storage (RNA)		X		X		X			X							X		
Baricitinib plasma concentrations (PK sample)		X[z]		X[z]	X[z]	X[z]	X[z]										X[z]	
Pharmacogenetic (DNA) collection[aa]		X														X	X	

Abbreviations: ANA = antinuclear antibody; BILAG = British Isles Lupus Assessment Group; BP = blood pressure; CLASI = Cutaneous Lupus Erythematosus Disease Area and Severity Index; C-SSRS = Columbia-Suicide Severity Rating Scale; dRVVT = Dilute Russell's Viper Venom Time; dsDNA = double-stranded deoxyribonucleic acid; ECG = electrocardiogram; eGFR = estimated glomerular filtration rate; ePRO = electronic patient-reported outcomes; EQ-5D = EuroQol-5Dimensions; ET = early termination; FACIT-Fatigue = Functional Assessment of Chronic Illness Therapy-Fatigue; FSH = follicle‑stimulating hormone; HBcAb = hepatitis B core antibody; HBsAb = hepatitis B surface antibody; HBsAg = hepatitis B surface antigen; HBV DNA = hepatitis B virus deoxyribonucleic acid; HIV = human immunodeficiency virus; Ig = immunoglobulins; IWRS = interactive web-response system; NRS = Numerical Rating Scale; PK = pharmacokinetics; PGI-C = Patient's Global Impression of Change; PPD = purified protein derivative; QIDS-SR16 = Quick Inventory of Depressive Symptomatology; RNA = ribonucleic acid; RNP = ribonucleoprotein; SF-36 = Short-Form 36-item Health Survey; SLE = systemic lupus erythematosus; SLEDAI-2K = Systemic Lupus Erythematosus Disease Activity Index 2000; SLICC/ACR = Systemic Lupus Erythematosus International Collaborating Clinics/American College of Rheumatology; SSA/Ro = Sjögren's-Syndrome-related antigen A; SSB/La = Sjögren syndrome type B antigen/Lupus La protein; TB = tuberculosis; TIBC = total iron binding capacity; TSH = thyroid-stimulating hormone; V = visit; WPAI-Lupus = Work Productivity and Activity Impairment-Lupus.

a At Visit 2, all baseline assessments must be performed and baseline laboratory samples must be drawn PRIOR to administration of the first dose of investigational product for randomized patients.

b Patients who complete the treatment period through Visit 16 may be eligible to participate in the long-term extension study (I4V-MC-JAIM [JAIM]). For patients enrolled in JAIM, Visit 1 of Study JAIM should occur on the same day as Visit 16 of Study I4V-MC-JAHZ (JAHZ).

c Those patients who discontinue from the study prior to Visit 16 should complete the ET visit and proceed to post-treatment follow-up.

d Patients should return for Visit 801, a post-treatment follow-up visit, 28±5 days after the last dose of investigational product. Patients who complete the treatment period through Visit 16 and enter the long-term extension (LTE; JAIM) will not complete Visit 801 in Study JAHZ, as they will complete post-treatment follow-up at the conclusion of JAIM. Patients who discontinue investigational product (IP) but remain in the study for at least 28±5 days without IP can combine their Visit 16/ET with their Visit 801 (post-treatment follow-up visit).

e One complete physical examination (excluding pelvic and rectal examinations) will be performed at Visit 1. All subsequent physical examinations may be symptom-directed. A complete physical examination may be repeated at the investigator's discretion any time.

f At each time point, 3 replicate readings should be made at approximately 30- to 60-second intervals. Blood pressure is recorded as the average of these three readings. A single pulse measurement should be made simultaneously with at least one of the readings at each time point (Section 9.4.2. Vital Signs).

g 12-lead ECGs (single) will be performed locally and will be locally (machine) read.

h TB test(s) include PPD, QuantiFERON®-TB Gold, and T SPOT®. See Exclusion Criterion [29] for details 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 must be performed 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, have no clinical features of active TB, and have a screening chest x-ray with no evidence of active TB are not required to undergo the protocol-specific TB testing (i.e., PPD, QuantiFERON®-TB Gold, or T SPOT®) but must have a chest x-ray at screening (i.e.: chest imaging performed within the past 6 months will not be accepted).

i For those patients with a PPD placed at Visit 1, the PPD must be read 2 to 3 days (48 to 72 hours) post-Visit 1. Purified protein derivative does not need to be read at the site, but must be read by a trained medical professional and results must be presented to the site at Visit 2.

j A chest x-ray (posterior-anterior view) will be performed at screening unless chest imaging (excluding ultrasound) has been performed within the past 6 months and the imaging and reports are available for review by the investigator.

k Prior to randomization, there must be documentation of a patient having met at least 4 of 11 Revised Criteria for Classification of Systemic Lupus Erythematosus according to the 1997 Update of the 1982 ACR criteria for classification of SLE (Tan et al. 1982; Hochberg et al. 1997). See Inclusion Criterion [3].

l Suicidal ideation and behavior subscales excerpt; adapted for the assessment of 11 preferred ideation and behavior categories. C-SSRS should be completed after collection of unsolicited adverse events.

m The Self-Harm Follow-up Form is only required if triggered by the Self-Harm Supplement Form.

n Patient eligibility must be reviewed and confirmed by an eligibility review committee prior to randomization. At Visit 2, patients will be randomized with stratification in IWRS, by total SLEDAI-2K score from Screening, corticosteroid dose at baseline, and region. The Screening eGFR result must also be entered into IWRS for appropriate dose adjustment to occur if necessary.

o ePRO devices (daily patient diaries) should be dispensed to all patients at screening and will need to be collected from screen-failed patients. Patients are encouraged to bring the ePRO device to every visit; however, if they do not bring the device, this will not be a protocol deviation. For patients NOT entering the LTE, the patient diary will be returned at Visit 801 or the patient's last visit. For patients entering the LTE, the patient diary will be returned at Visit 16.

p Worst Pain NRS, Worst Joint Pain NRS, Worst Fatigue NRS, Joint Stiffness Duration, Joint Stiffness Severity NRS, and PGI-S will be collected on an electronic patient diary daily (ePRO device) starting at Visit 1 (screening) through Visit 16 (or Visit 801 for those not entering the LTE Study JAIM) and should be completed prior to any clinical assessments on days when study visits occur.

q SLE assessments must be completed by a physician and will be documented on an eCOA (electronic clinical outcome assessment) device.

r Patient-reported questionnaires will be administered via an on-site eCOA device and should be administered prior to any clinical assessments.

s 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.

t For patients who are positive for HBcAb, a follow-up test for HBV DNA will be performed automatically. 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 of events (see Section 9.4.8 for details of HBV DNA monitoring).

u For all women of child-bearing potential, a serum pregnancy test (central laboratory) will be performed at Visit 1. Urine pregnancy tests (local laboratory) will be performed at Visit 2 prior to randomization and at all subsequent study visits.

v For female patients ≥50 and <55 years of age who have had a cessation of menses for at least 6 and less than 12 months, an FSH test will be performed to confirm non-child-bearing potential (FSH ≥40 mIU/mL).

w Clinical chemistry will include the following value calculated by the central laboratory from serum creatinine: estimated glomerular filtration rate (eGFR; calculated using the Modification of Diet in Renal Disease [MDRD] isotope dilution mass spectrometry traceable method).

x Fasting lipid profile. Patients should not eat or drink anything except water for 12 hours prior to sample collection. If a patient attends these visits in a nonfasting state, the sample should still be collected. This will not be considered a protocol violation.

z See Section 9.5 for details.

aa Patients choosing not to provide consent for whole blood samples to be collected for pharmacogenetic research will still be eligible to participate in the study.

bb Urine should be collected as a clean catch sample.

3. Introduction

3.1. Study Rationale

Baricitinib is an oral, reversible, selective inhibitor of Janus kinase (JAK)1 and JAK2 (Fridman et al. 2010). This activity profile suggests that baricitinib may inhibit cytokines implicated in systemic lupus erythematosus (SLE), most notably type I interferon (IFN JAK1/TYK2), interleukin (IL)-6 (JAK1/JAK2/TYK2), and type II IFN-γ, as well as other cytokines that may have a role in SLE, including IL-23 (JAK2/TYK2), granulocyte-macrophage colony stimulating factor (JAK2/JAK2) and IL-12 (JAK2/TYK2). The potential impact of baricitinib on the IFN pathway is relevant to SLE, as clinical and preclinical studies have established that this pathway is involved in the pathogenesis of SLE (Hoffman et al. 2017). Suppressing the IFN signature through the use of a monoclonal antibody that targets the type I IFN receptor has been shown to be effective in SLE (Furie et al. 2017). Baseline gene expression in SLE, examined as part of the ILLUMINATE Phase 3 trials, demonstrated upregulation of IL–6-related genes, IL-2 receptor, IL-12, and oncostatin M in patients with SLE compared to control subjects, all of which signal through the JAK/signal transducers and activators of transcription (STAT) pathways and would not be anticipated to be directly modulated by therapeutic strategies which only target the IFN pathway (Hoffman et al. 2017). Thus, baricitinib is anticipated to be effective as it modulates not only IFN but also other potential SLE-relevant pathways, providing a biologic rationale for its evaluation in this disease.

Baricitinib has demonstrated clinical efficacy and safety in patients with autoimmune (SLE, rheumatoid arthritis [RA; Taylor et al. 2015; Fleischmann et al. 2016; Genovese et al. 2016; Dougados et al. 2017]) and autoinflammatory diseases (Chronic atypical neutrophilic dermatosis with lipodystrophy and elevated temperature [CANDLE], Stimulator of Interferon Genes (STING)-Associated Vasculopathy With Onset During Infancy [SAVI], Juvenile Dermatomyositis [JDM], and Aicardi-Goutières Syndrome [AGS]) (Montealegre Sanchez et al. 2018). A Phase 2 study of baricitinib in patients with SLE demonstrated that baricitinib 4-mg plus standard of care (SoC) was superior to placebo plus SoC in the proportion of patients achieving remission of rash and/or arthritis as defined by the Systemic Lupus Erythematosus Disease Activity Index 2000 (SLEDAI-2K) at Week 24 (67.3% baricitinib 4-mg, 58.1% baricitinib 2-mg, 53.3% placebo, $p<0.05$ for baricitinib 4-mg versus placebo). Patients treated with baricitinib 4-mg also had a higher Systemic Lupus Erythematosus Responder Index-4 (SRI-4) response rate at Week 24 compared to patients treated with placebo (64.4% baricitinib 4-mg, 51.4% baricitinib 2-mg, 47.6% placebo, $p<0.05$ for baricitinib 4-mg versus placebo) (Eli Lilly and Company data on file, 2017). Additionally, completed Phase 2 studies of baricitinib also demonstrated efficacy in patients with moderate-to-severe plaque psoriasis (Papp et al. 2016), diabetic kidney disease (Tuttle et al. 2015), and atopic dermatitis (Guttman-Yassky et al. 2018).

Given the efficacy of baricitinib demonstrated in clinical trials for treating autoimmune/autoinflammatory diseases involving joints, skin, and kidney (including SLE), the acceptable safety profile of baricitinib observed through the current stage of development, and a

continuing unmet medical need in patients with SLE, there is a compelling rationale for initiation of a Phase 3 program to evaluate baricitinib in the treatment of SLE.

3.2. Background

Systemic lupus erythematosus is a chronic, often debilitating, multisystem, autoimmune disease that is characterized by the presence of autoreactive B cells and elevated autoantibodies, which directly damage the body's cells and tissues. Systemic lupus erythematosus can affect multiple organ systems simultaneously or sequentially, and follows a highly variable clinical course where periods of relatively stable disease are followed by flares and/or periods of persistently active disease; all of which can ultimately lead to irreversible damage to tissues and organ systems.

Systemic lupus erythematosus is predominately a disease of women (approximately 9:1 female to male ratio), which can begin at any age but most commonly begins in adolescence or early adulthood (Yu et al. 2017). It affects 20 to 150 in 1,000,000 people in the US (UpToDate® 2018 [WWW]) and is more common in African-Americans (Lim et al. 2014; Somers et al. 2014), with as many as 1 in 537 African-American women afflicted with SLE (Somers et al. 2014). Additionally, SLE appears to be more severe in African-Americans, Asian-Americans, and Latinos compared to Caucasians (Kaslow 1982; Alarcón et al. 2001).

Clinically, SLE presents with varying signs and symptoms, including fever, arthralgia/arthritis, skin rash, alopecia, pleuritis, pericarditis, nephritis, vasculitis, stroke, seizure, leukopenia, thrombocytopenia, anemia, photosensitivity, and the presence of autoantibodies reactive with nuclear antigens. Fatigue is the most prevalent symptom reported among patients with SLE (Avina-Zubieta et al. 2007). Pain that interferes with daily living activities is also commonly reported (Ozel and Argon 2015). Skin and joint disease are also among the most prevalent features of the illness. Age, USA African race/ethnicity, SLEDAI-2K score, steroid use, and hypertension were associated with transition from no damage to damage, and increase(s) in preexisting damage (Bruce et al. 2015). Over 60% of patients with SLE will develop clinically detectable organ damage within 2 to 7 years of diagnosis, as measured by the Systemic Lupus Erythematosus International Collaborating Clinics (SLICC)/American College of Rheumatology (ACR) damage index (Cooper et al. 2007).

Improvements in earlier diagnosis, treatment regimens, and medical care over the past several decades have reduced mortality in SLE. However, patients continue to experience premature death, with cardiovascular disease being the leading cause. A recent meta-analysis of published data involving over 27,000 patients with SLE observed a 3-fold increase in the risk of death in patients with SLE compared with the general population (Yurkovich et al. 2014). Morbidity remains substantial as measured by various tools for features such as health-related quality of life, loss of work productivity, pain, and fatigue (Avina-Zubieta et al. 2007; Ozel and Argon 2015). Thus, there remains substantial unmet medical need for individuals who have SLE.

Standard of care (SoC) for SLE includes antimalarial agents, corticosteroids, nonsteroidal anti-inflammatory drugs (NSAIDs), immunosuppressive agents, and cytotoxic agents; however, there are relatively few drugs approved for the treatment of SLE. For example, in the US,

approved therapies for SLE include aspirin, antimalarials, corticosteroids, and belimumab. In general, treatment regimens are broadly similar around the world and are tailored to the severity of disease and the specific organs involved. Mild disease is often treated with low-dose corticosteroids, NSAIDs, and antimalarials; while serious, organ-threatening or life-threatening disease is typically treated with high-dose corticosteroids and immunosuppressive agents. In addition to their direct impact on disease, immunosuppressive agents are also utilized as so called "corticosteroid-sparing agents" to reduce chronic exposure to corticosteroids.

The current SoC therapies have broad effects on immune and inflammatory pathways, including host defense, and have been associated with short- and long-term morbidity. For example, long-term use of corticosteroids is associated with cataracts, osteoporosis, avascular necrosis, increased infection, cardiovascular events, hyperglycemia, and weight gain, while cyclophosphamide increases the risk of premature ovarian failure, serious infection, and cancer.

Although recent improvements in treatment regimens and medical care have reduced overall morbidity and mortality, many patients still have incompletely controlled disease and progress to end-stage organ involvement. In addition, the disease increases mortality and negatively impacts health-related quality of life. New treatment options with an acceptable safety profile that reduce disease activity, reduce flares, delay organ damage, and reduce the requirement for corticosteroids and cytotoxic agents are urgently needed for patients with SLE.

Accordingly, pharmacologic interventions that target specific pathways associated with the pathology of SLE may provide novel therapeutic approaches to disease management. One of the signaling pathways implicated in SLE disease activity is the type I IFN signaling pathway. Upregulation of genes associated with the activation of type I IFN signaling, referred to as a type I IFN signature, is observed in approximately 75% of patients with SLE (Hoffman et al. 2017). In SLE, a high type I IFN signature was associated with increased disease severity as measured by SLEDAI score, increased anti–double-stranded deoxyribonucleic acid (anti-dsDNA), decreased complement, and increased risk of severe flares (Hoffman et al. 2017). Another cytokine implicated in the pathogenesis of SLE is IL-6. Increased expression of IL-6 has been found in murine models of SLE and in patients with SLE, and inhibition of IL-6 signaling was associated with a decrease in disease activity (Linker-Israeli et al. 1999, Illei et al. 2010). Both type I IFNs and IL-6 signal through the JAK/STAT pathway; therefore, treatment of SLE with baricitinib or other JAK inhibitors is an area of intense interest.

3.3. Benefit/Risk Assessment

Phase 2 SLE Data

In the baricitinib Phase 2 SLE clinical Study I4V-MC-JAHH (JAHH), significant benefit for baricitinib 4-mg once daily (QD) treatment over placebo was demonstrated for the primary efficacy endpoint of resolution of arthritis and/or rash, the major secondary efficacy endpoint of SRI-4 response, and important supporting measures such as flare and lupus low disease activity score (LLDAS) after 24 weeks of treatment. Baricitinib 2-mg QD was numerically better than placebo across these measures although the results were not statistically significant.

The percentage of patients with at least 1 treatment-emergent adverse event (TEAE) during the 24-week treatment period was 64.8% for placebo, 71.4% for baricitinib 2-mg QD, and 73.1% for baricitinib 4-mg QD. The most common TEAEs (occurring in ≥5% of patients in either baricitinib dose group over 24 weeks) were viral upper respiratory tract infection, urinary tract infection, upper respiratory tract infection, pharyngitis, and headache. The proportions of patients who reported viral upper respiratory tract infections were higher for baricitinib-treated patients compared to placebo, but the proportions were similar for the baricitinib 2-mg and 4-mg dose groups. It is important to consider that in SLE, patients are at an increased risk for infections due to predisposing conditions related to impaired cellular and humoral immune functions (Danza and Ruiz Irastorza 2013), as well as concomitant use of immunosuppressive agents such as corticosteroids. In the Phase 2 study of baricitinib, there were more serious infections reported in the baricitinib 4-mg group versus the baricitinib 2-mg group or placebo. There were no cases of serious herpes zoster or opportunistic infection, and no reports of tuberculosis (TB). There were no malignancies or major adverse cardiovascular events (MACE) reported. One serious adverse event (SAE) of deep vein thrombosis (DVT) was reported in the baricitinib 4-mg group, in a patient with pre-existing antiphospholipid antibodies and pain in the affected limb (right calf) prior to study entry, who was taking a concomitant oral corticosteroid and celecoxib during the study.

Risks Identified in the Baricitinib Program

Serious infections, venous thromboembolism, hepatotoxicity, and fetal malformations were identified as important potential risks with baricitinib. Inclusion/Exclusion criteria in the protocol limit enrolment of SLE patients who are at risk for these important potential risks. Although infections were seen in about half of the SLE study population, the incidence rate of serious infection in patients exposed to baricitinib in the RA program was only 3.0 per 100 patient-years. During the controlled period (through 16 weeks), rates were similar in both baricitinib- and placebo-treated patients. The nonserious infections (upper respiratory tract infections, herpes zoster, and herpes simplex) associated with baricitinib in the RA program are readily diagnosed, manageable, and typically resolve without long-term sequelae. It is recommended that where indicated, herpes zoster vaccination be offered to patients prior to receiving baricitinib.

Increases in levels of alanine aminotransferase (ALT), aspartate aminotransferase (AST), and total bilirubin have been seen in patients with RA. Most of these increases improved with continued use or temporary discontinuation of baricitinib with no long-term effects. No cases of severe drug-induced liver injury were observed with baricitinib treatment.

Fetal malformations were reported in toxicology studies at higher doses than what is used in human patients. Only a small number of patients have become pregnant in baricitinib clinical trials, and there have been no reports of fetal malformations in these pregnancies.

Venous thromboembolic events (VTEs) have been determined to be an important potential risk for baricitinib. There was a numerical imbalance in reports of VTEs in the 24-week placebo-controlled period of the Phase 3 trials of patients with RA. Available evidence does not

establish a causal association. With long-term exposures, the exposure-adjusted incidence rate of VTE for baricitinib-treated patients with RA was similar to the background rates published in the literature for the target population. There was no pattern of increased or decreased risk during long-term exposures, and cases observed with baricitinib were confounded by one or more recognized risk factors for VTE. Venous thromboembolic event risk can be managed through risk mitigation strategies. Exclusion and discontinuation criteria have been added to the protocol to limit participation of patients who are at an increased risk of VTE.

Therefore, in the context of the cumulative knowledge, the benefit/risk balance for baricitinib in the treatment of adult patients with SLE is assessed to be favorable.

More information about the known and expected benefits, risks, SAEs, and reasonably anticipated adverse events (AEs) of baricitinib are to be found in the Investigator's Brochure (IB).

4. Objectives and Endpoints

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

Table JAHZ.2. Objectives and Endpoints

Objectives	Endpoints
Primary To evaluate the effect of baricitinib 4-mg QD and background standard-of-care (SoC) therapy compared to placebo and SoC on SLE disease activity.	Proportion of patients achieving an SRI-4 response at Week 52, defined as: • Reduction of ≥4 points from baseline in SLEDAI-2K score; and • No new British Isles Lupus Assessment Group (BILAG) A or no more than 1 new BILAG B disease activity score; and • No worsening (defined as an increase of ≥0.3 points [10 mm] from baseline) in the Physician's Global Assessment of Disease Activity.
Major Secondary (Multiplicity Controlled) • To evaluate the effect of baricitinib 4-mg QD plus SoC compared to placebo plus SoC on SLE disease activity.	• Proportion of patients achieving an SRI-4 response at Week 24. • Proportion of patients achieving a lupus low disease activity state (LLDAS) response at Week 52
• To evaluate the corticosteroid sparing effect of baricitinib 4-mg QD plus SoC compared to placebo plus SoC.	• Proportion of patients receiving >7.5 mg prednisone (or equivalent) at baseline able to decrease dose by ≥25% to a prednisone equivalent dose of ≤7.5 mg/day maintained between Week 40 and Week 52.
• To evaluate the effect of baricitinib 4-mg QD plus SoC compared to placebo plus SoC on SLE flares.	• Time to first severe flare over 52 weeks.
• To evaluate the effect of baricitinib 4-mg QD plus SoC compared to placebo plus SoC on patient-reported outcomes (PROs).	• Change from baseline in Worst Pain NRS at Week 52. • Change from baseline in FACIT-Fatigue total score at Week 52.
• To evaluate the effect of baricitinib 2-mg QD plus SoC compared to placebo plus SoC on SLE disease activity.	• Proportion of patients achieving an SRI-4 response at Week 52. • Proportion of patients achieving an SRI-4 response at Week 24. • Proportion of patients achieving a lupus low disease activity state (LLDAS) response at Week 52
• To evaluate the corticosteroid sparing effect of baricitinib 2-mg QD plus SoC compared to placebo plus SoC.	• Proportion of patients receiving >7.5 mg prednisone (or equivalent) at baseline able to decrease dose by ≥25% to a prednisone equivalent dose of ≤7.5 mg/day maintained between Week 40 and Week 52.
• To evaluate the effect of baricitinib 2-mg QD plus	• Time to first severe flare over 52 weeks.

SoC compared to placebo plus SoC on SLE flares. • To evaluate the effect of baricitinib 2-mg QD plus SoC compared to placebo plus SoC on patient-reported outcomes (PROs).	• Change from baseline in Worst Pain NRS at Week 52. • Change from baseline in FACIT-Fatigue total score at Week 52
Other Secondary (Non-Multiplicity Controlled) • To evaluate the effect of baricitinib 4-mg or 2-mg QD plus SoC compared to placebo plus SoC on SLE disease activity.	• Change from baseline in total SLEDAI-2K scores over time through Week 52. • Change from baseline in PGA score over time through Week 52. • Proportion of patients achieving an SRI-4 response over time through Week 52. • Proportion of patients achieving an SRI-5, -6, -7, and -8 response over time through Week 52. • Proportion of patients achieving reduction of ≥4 points from baseline in SLEDAI-2K score at Week 52. • Proportion of patients with no new British Isles Lupus Assessment Group (BILAG) A or no more than 1 new BILAG B disease activity score at Week 52. • Proportion of patients with no worsening (defined as an increase of ≥0.3 points [10 mm] from baseline) in the Physician's Global Assessment of Disease Activity at Week 52.
• To evaluate the corticosteroid sparing effect of baricitinib 4-mg or 2-mg QD plus SoC compared to placebo plus SoC.	• Change from baseline in prednisone dose at Week 52. • Proportion of patients taking corticosteroids at baseline able to discontinue use at Week 52.
• To evaluate the effect of baricitinib 4-mg or 2-mg QD plus SoC compared to placebo plus SoC on SLE flares.	• Time to first mild/moderate flare over 52 weeks. • Time to first flare (any severity) over 52 weeks. • Annualized mild/moderate flare rate • Annualized severe flare rate • Annualized flare rate (any severity)
• To evaluate the effect of baricitinib 4-mg QD or 2-mg QD plus SoC compared to placebo plus SoC on mucocutaneous manifestations of SLE.	• Proportion of patients with CLASI total activity score ≥ 10 at baseline with ≥50% reduction in CLASI total activity score at Week 52.
• To evaluate the effect of baricitinib 4-mg QD or 2-mg QD plus SoC compared to placeb plus SoC on musculoskeletal manifestations of SLE.	• Change from baseline in tender joint count at Week 52. • Change from baseline in swollen joint count at Week 52.
• To evaluate the effect of baricitinib 4-mg QD or 2-mg QD plus SoC compared to placebo plus SoC on individual organ system disease activity.	• Proportion of patients with improvement in each SLEDAI-2K organ system versus baseline at Week 52. • Proportion of patients with worsening in each SLEDAI-2K organ system versus baseline at Week 52.

• To evaluate the effect of baricitinib 4-mg QD or 2-mg QD plus SoC compared to placebo plus SoC on PROs. • To measure baricitinib PK exposure and assess the relationship between exposure and efficacy.	• Change from baseline in Worst Fatigue NRS at Week 52. • Change from baseline in Worst Joint Pain NRS at Week 52. • Change from baseline in Joint Stiffness Duration at Week 52. • Change from baseline in Joint Stiffness Severity NRS at Week 52. • Change from baseline in Patient Global Impression of Severity at Week 52. • Change from baseline in Patient Global Impression of Change at Week 52 • Change from baseline in mental component score (MCS), physical component score (PCS), and domain scores in the Short-Form 36-item health survey version 2 (SF-36v2) acute at Week 52. • Change from baseline in the EQ-5D-5L at Week 52. • Change from baseline in the WPAI-Lupus at Week 52. • Population pharmacokinetics of baricitinib in patients with SLE. • Proportions of patients achieving SRI-4 by exposure quartile.
Exploratory • To evaluate the effect of baricitinib 4-mg QD or 2-mg QD plus SoC compared to placebo plus SoC on damage. • To evaluate the effect of baricitinib 4-mg QD or 2-mg QD plus SoC compared to placebo plus SoC on serologic markers of SLE. • To evaluate the effect of baricitinib 4-mg QD or 2-mg QD plus SoC compared to placebo plus SoC on SLE disease activity in subgroups of interest.	• Change from baseline in SLICC/ACR damage index total score at Week 52. • Change from baseline in IFN signature over 52 weeks. • Change from baseline in anti-dsDNA level in patients with elevated anti-dsDNA at baseline over 52 weeks. • Change from baseline in C3 and/or C4 levels in patients with low C3 and/or C4 at baseline over 52 weeks. • Proportion of patients with high IFN signature achieving an SRI-4 response at Week 52.

Abbreviations: ACR = American College of Rheumatology; BILAG = British Isles Lupus Assessment Group; C3 = complement 3; C4 = complement 4; CLASI = Cutaneous Lupus Erythematosus Disease Area and Severity Index; dsDNA = double-stranded deoxyribonucleic acid; EQ-5D-5L = European Quality of Life–5 Dimensions–5 Levels; FACIT-Fatigue = Functional Assessment of Chronic Illness Therapy–Fatigue Scale; IFN = interferon; MCS = Mental Component Score; NRS = Numeric Rating Scale; PCS = physical component score; PGA = Physician Global Assessment; PK = pharmacokinetic; QD = once daily; SF-36v2 = Short-Form 36-item health survey version 2; SLE = systemic lupus erythematosus; SLEDAI-2K = Systemic Lupus Erythematosus Disease Activity Index 2000; SLICC = Systemic Lupus Erythematosus International Collaborating Clinics; SoC = standard of care; SRI = Systemic Lupus Erythematosus Responder Index; WPAI-Lupus = Work Productivity and Activity Impairment Questionnaire-Lupus.

5. Study Design

5.1. Overall Design

Study I4V-MC-JAHZ (JAHZ) is a Phase 3, multicenter, randomized, double-blind, parallel-group, placebo-controlled, outpatient, 52-week study evaluating the efficacy and safety of baricitinib 4-mg and 2-mg in patients with SLE receiving standard therapy.

Approximately 750 patients will be randomized at a 1:1:1 ratio to receive baricitinib 4-mg daily, baricitinib 2-mg daily, or placebo (250 patients per treatment group). Patients will be stratified at randomization according to disease activity (total SLEDAI-2K at screening <10; ≥10), corticosteroid dose at baseline (<10 mg/day; ≥10 mg/day prednisone or equivalent), and region (defined in the statistical analysis plan [SAP]).

All procedures to be conducted during the study, including timing of all procedures, are indicated in the Schedule of Activities (Section 2). Section 7.7 describes permitted concomitant therapy. Study governance considerations are described in detail in Appendix 3.

Figure JAHZ.1 illustrates the study design.

The study consists of 3 periods:

Screening Period: The screening period is up to 42 days and begins when the patient signs the informed consent form (ICF). In exceptional circumstances, the screening window can be extended after consultation with the sponsor. During the screening period, the site must confirm the patient meets all inclusion and no exclusion criteria for the study (see Section 6). Screening procedures will be performed according to the Schedule of Activities (Section 2). Patient eligibility will be reviewed and confirmed by an eligibility review committee prior to randomization. Eligible patients will be randomized at Week 0 (Visit 2).

Treatment Period:

At Visit 2 (Week 0, baseline) after all assessments are completed and baseline laboratory samples obtained according to the Schedule of Activities, randomized patients will take the first dose of investigational product at the clinic and pharmacokinetic (PK) samples will be drawn 15 minutes and 1 hour postdose.

Investigational product will be taken daily for 52 weeks. Clinical assessments and laboratory samples, including additional PK sampling, will be obtained at scheduled visits according to the Schedule of Activities (Section 2). Special timing considerations of administration of investigational product relative to PK sample collection at study visits are described in Section 9.5.

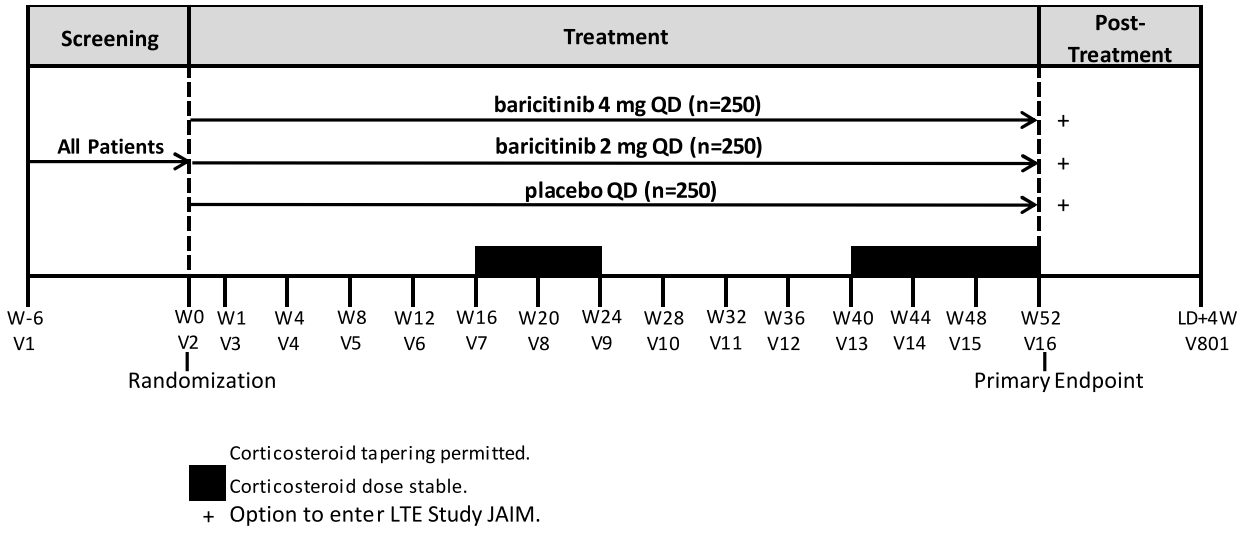
During the treatment period, in addition to randomized treatment, patients will also maintain their usual medication regimen for SLE, including background standard-of-care therapy consisting of corticosteroids, antimalarials, and/or immunosuppressants (see Section 6.1, Inclusion Criterion [8] and Section 7.7 for details).

Between Weeks 0 through 12 of the treatment period, patients with intolerable disease activity may receive a single corticosteroid burst if necessary. See Section 7.7.1 for details.

Between Weeks 0 through 16 and Weeks 24 through 40 of the treatment period, investigators are encouraged to taper corticosteroids according to clinical practice, with a goal of ≤7.5 mg/day prednisone (or equivalent) by Week 40 (Visit 13). In most cases, tapering should not be considered for at least 4 weeks after randomization, consistent with clinical practice. If a patient experiences an increase in disease activity during an attempted taper, the patient's corticosteroid dose may be increased once, by up to 10 mg/day, but may not exceed the baseline dose. Corticosteroid dose must be stable for 8 weeks prior to Week 24 (i.e., no changes between Weeks 16 and 24) and for 12 weeks prior to Week 52 (i.e., no changes between Weeks 40 and 52). This design minimizes confounding of assessments at Week 24 and Week 52.

Throughout the trial, investigators should continue to assess benefit/risk for patients to remain in the trial and should consider discontinuing patients if sufficient clinical benefit is not observed with protocol permitted treatments.

Follow-up Period: Patients who complete the treatment period of the study through Week 52 (Visit 16) are eligible to participate in the long-term extension (LTE), Study I4V-MC-JAIM (JAIM). Patients who do not participate in the LTE, as well as those who discontinue JAHZ study treatment early, will have a post-treatment follow-up visit (Visit 801) approximately 4 weeks after the last dose of investigational product.



Abbreviations: LD = last dose of investigational product; LTE = long-term extension; n = number of patients in each group; QD = once daily; V = visit; W = week

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

5.2. Number of Participants

Approximately 750 patients will be randomized for a total of 250 patients per treatment group.

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 be double-blinded and placebo-controlled to minimize bias. The selection of placebo as a comparator for this study is warranted as there are few approved therapies for SLE and all patients in the study will be receiving concomitant SoC, including corticosteroids, antimalarials, and/or immunosuppressants. In order to minimize the confounding effects of concomitant medications, doses of antimalarials and immunosuppressants should remain stable during the treatment period (see Section 7.7).

The population studied will include patients with active disease and varying degrees of SLE organ involvement; however, patients with severe active lupus nephritis and active central nervous system (CNS) SLE are excluded as they would require aggressive treatment with medications that could confound the safety and/or efficacy results of the study.

The primary efficacy endpoint of SRI-4 is consistent with FDA and EU guidance (FDA 2010; EMA 2015) and follows regulatory precedent (FDA 2012) for assessing efficacy in SLE registration studies. Furthermore, the SRI-4, as a composite index, detects overall improvements in disease activity (SLEDAI-2K component) while ensuring there is not worsening in other organ systems (BILAG and Physician's Global Assessment components).

Owing to the heterogeneity and natural course of SLE disease, a duration of 52 weeks is considered necessary to confirm efficacy, safety, and maintenance of response. This duration also allows for a corticosteroid burst up to Week 12, for patients who enter the study with increased disease activity. This design prevents premature discontinuation of patients from the study while minimizing potential confounding effects on efficacy assessments at Weeks 24 and 52 (the primary endpoint).

5.5. Justification for Dose

The 4-mg and 2-mg QD doses of baricitinib selected for this study are based on the Phase 2 SLE study, Study JAHH, and are additionally supported by PK, safety, and efficacy data for baricitinib in Phase 2 and Phase 3 RA studies and the Phase 2 psoriasis study.

In the Phase 2 SLE study, baricitinib 4-mg QD demonstrated statistically significant efficacy compared to placebo across several relevant measures, including remission of arthritis and/or rash (SLEDAI-2K), SRI-4, flare, and LLDAS. Baricitinib 2-mg QD was numerically better than placebo across these measures, although the results were not statistically significant. Dose response patterns favoring baricitinib 4-mg were evident across measures and improvement was observed for both doses in patient-reported outcomes including worst pain and worst joint pain. Pharmacokinetic/pharmacodynamic (PK/PD) modeling suggests the maximum response after dosing with 2-mg was not attained by 24 weeks. Therefore, there is potential that baricitinib

2-mg will also demonstrate efficacy at Week 52. No notable safety findings emerged compared to the results from baricitinib studies for other indications.

In Phases 2 and 3 RA studies, baricitinib 4-mg QD has demonstrated consistent efficacy; baricitinib 2-mg QD was also effective, but was less consistent than 4-mg QD across measures of efficacy and patient populations. Both doses were acceptably safe and well tolerated in the context of the efficacy observed and the benefit/risk profiles of other disease-modifying antirheumatic drugs. There was no substantial increase in efficacy noted with doses higher than 4-mg, where studied.

In the Phase 2 psoriasis study, baricitinib doses between 4-mg and 10-mg were associated with statistically significant reductions in measures of disease activity, with greater efficacy at the higher doses. The 2-mg dose did show numeric improvements in efficacy compared with placebo. The 8-mg and 10-mg doses were associated with a higher rate of AEs related to laboratory abnormalities (decreases in hemoglobin, neutrophils, and lymphocytes), while the 2-mg and 4-mg dose groups had a pattern of AEs similar to placebo.

Dose Adjustment for Renal Impairment

As detailed in the IB, baricitinib exposure increases with decreased renal function (Study I4V-MC-JADL [JADL]). Based on PK simulations, dose adjustment is not required for patients with estimated glomerular filtration rate (eGFR) ≥60 mL/min/1.73 m^2. Patients with eGFR <60 mL/min/1.73 m^2 at screening who are randomized to the 4-mg dose will receive a dose of 2-mg QD, which will ensure that exposures do not exceed those of the 4-mg QD dose in patients with eGFR ≥60 mL/min/1.73 m^2. For patients randomized to the 2-mg dose, there will be no dose reduction based on renal function.

See Section 7.2.2 for method of treatment assignment in patients with renal impairment.

6. Study Population

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

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

Patient eligibility will be reviewed and must be confirmed by an eligibility review committee prior to randomization.

6.1. Inclusion Criteria

Patients are eligible to be included in the study only if they meet all of the following criteria:

Type of Patient and Disease Characteristics

[1] Are at least 18 years of age.

[2] Have a clinical diagnosis of SLE at least 24 weeks prior to screening.

[3] Have documentation of having met at least 4 of 11 Revised Criteria for Classification of Systemic Lupus Erythematosus according to the 1997 Update of the 1982 ACR criteria for classification of SLE (Tan et al. 1982; Hochberg et al. 1997) prior to randomization.

[4] Have 1 or more of the following as assessed by the central lab during screening: a positive antinuclear antibody (ANA; titer ≥1:80), and/or a positive anti-dsDNA, and/or a positive anti-Smith (anti-Sm). Patients with an ANA <1:80 at screening with documentation of a historical ANA ≥1:80 may be eligible, as assessed by the eligibility review committee.

Note: The ANA, anti-dsDNA, and anti-Smith measurements may be repeated by the central lab once during the screening period, and the value resulting from repeat testing may be accepted for enrollment eligibility if it meets the eligibility criterion.

[5] Have a total SLEDAI-2K score ≥6 during screening, with at least 4 points attributed to clinical items (not including items requiring laboratory value assessment). SLEDAI-2K items requiring laboratory values should be assessed based on the results from the labs drawn during the screening period.

[6] Have a clinical SLEDAI-2K score ≥4 at Baseline (Visit 2); not including any items requiring laboratory value assessment.

[7] Have at least 1 BILAG A score or 2 BILAG B scores during the screening period. BILAG items requiring laboratory values should be assessed based on the results from the labs drawn during the screening period.

Prior/Concomitant Therapy

[8] Are receiving at least one of the following SoC medications for SLE:

a. A single antimalarial (such as hydroxychloroquine, chloroquine, quinacrine) at a stable therapeutic dose for at least 8 weeks prior to screening (Visit 1).

b. A single immunosuppressant (such as methotrexate [MTX], azathioprine, mycophenolate, tacrolimus, leflunomide, cyclosporine) at a stable therapeutic dose for at least 8 weeks prior to screening (Visit 1).

c. An oral corticosteroid, initiated at least 4 weeks prior to screening (Visit 1), at a stable dose ≤ 40 mg/day prednisone (or equivalent) for at least 2 weeks prior to screening (Visit 1) and through baseline (Visit 2). If the patient is not receiving an antimalarial or immunosuppressant, the dose of corticosteroid must be ≥ 7.5 mg/day prednisone (or equivalent).

Patient Characteristics

[9] Male or nonpregnant, nonbreastfeeding female patient

a. Patients of child-bearing 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 the opposite sex.

b. Total abstinence is defined as refraining from intercourse during the entirety of the study and for at least 1 week following the last dose of investigational product. Periodic abstinence such as calendar, ovulation, symptothermal, post-ovulation methods and withdrawal are not acceptable methods of contraception.

c. Otherwise, patients of childbearing potential must agree to use 2 effective methods of contraception, where at least 1 form is highly effective, for the entirety of the study and for at least 1 week following the last dose of investigational product.

d. The following contraception methods are considered acceptable (the patient should choose 2, and 1 must be highly effective [defined as less than 1% failure rate per year when used consistently and correctly]):

- Highly effective birth control methods:
  - Combined (estrogen and progestogen containing) hormonal contraception associated with inhibition of ovulation: oral, intravaginal, or transdermal
  - Progestogen- only containing hormonal contraception associated with inhibition of ovulation: oral, intravaginal, or transdermal
  - intrauterine device (IUD)/intrauterine hormone-releasing system (IUS)

- vasectomized male (with appropriate post-vasectomy documentation of the absence of sperm in the ejaculate).

- Effective birth control methods:
  - Male or female condom with spermicide. It should be noted that the use of male and female condoms as a double barrier method is not considered acceptable due to the high failure rate when these methods are combined.
  - Diaphragm with spermicide
  - Cervical sponge
  - Cervical cap with spermicide

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

Patients of non–child-bearing potential are not required to use birth control and they are defined as:

- Women who are infertile due to surgical sterilization (hysterectomy, bilateral oophorectomy, or tubal ligation)
- Post-menopausal – defined either as
  - A woman at least 50 years of age with an intact uterus, not on hormone therapy, who has either
    - Cessation of menses for at least 1 year
    - At least 6 months of spontaneous amenorrhea with follicle-stimulating hormone >40 mIU/mL
- Women aged 55 years or older who are not on hormone therapy, and who have had at least 6 months of spontaneous amenorrhea
- Women aged 55 years or older who have a diagnosis of menopause

Informed Consent

[10] Must read and understand the informed consent approved by Eli Lilly and Company (Lilly), or its designee, and the institutional review board (IRB)/ethics review board (ERB) governing the site, and provide written informed consent.

6.2. Exclusion Criteria

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

Medical Conditions

[11] Have severe active lupus nephritis defined clinically and/or by histologic evidence of proliferative glomerulonephritis on renal biopsy (if available) within the 24 weeks prior to screening, or urine protein/creatinine ratio >200 mg/mmol (as an estimate of approximate proteinuria >2 g/day) or eGFR (Modification of Diet in Renal Disease [MDRD]) <40 mL/min/1.73 m^2 at screening, or as determined by the eligibility review committee.

Note: The lab measurements related to lupus nephritis may be repeated once by the central lab during the screening period, and the values resulting from repeat testing may be accepted for enrollment eligibility if they meet the eligibility criterion.

[12] Have active CNS lupus as defined by ACR nomenclature for neuropsychiatric lupus syndromes and as captured by SLEDAI-2K (seizure, psychosis, organic brain syndrome, visual disturbance, cranial nerve disorder, lupus headache, and cerebrovascular accident).

[13] Have active fibromyalgia that, in the investigator's opinion, would make it difficult to appropriately assess SLE activity for the purposes of this study.

[14] Have been treated for or had an active occurrence of a systemic inflammatory condition other than SLE such as, but not limited to, RA, juvenile chronic arthritis, spondyloarthropathy, Crohn's disease, ulcerative colitis, or psoriatic arthritis within the 12 weeks prior to screening. Patients with secondary Sjögren's syndrome are not excluded.

[15] Have had any major surgery within 8 weeks prior to 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.

[16] 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.

[17] Have experienced any of the following within 12 weeks of screening: VTE (DVT/pulmonary embolism [PE]), myocardial infarction (MI), unstable ischemic heart disease, stroke, or New York Heart Association Stage III/IV heart failure.

[18] Have a history of recurrent (≥ 2) VTE (DVT/PE).

[19] 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.

[20] Have a history of lymphoproliferative disease; have signs or symptoms suggestive of possible lymphoproliferative disease, including lymphadenopathy or splenomegaly (other than primarily due to SLE); have active primary or recurrent malignant disease; or have been in remission from clinically significant malignancy for <5 years prior to randomization.

The following may be exempted:

a. Patients with cervical carcinoma in situ that has been resected 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 completely resected with no evidence of recurrence for at least 3 years may participate in the study.

[21] Have a current or recent (<4 weeks prior to randomization) clinically serious viral, bacterial, fungal, or parasitic infection or any other active or recent infection that in the opinion of the investigator, would pose an unacceptable risk to the patient if participating in the study.

Note: For example, a recent viral upper respiratory tract infection or uncomplicated urinary tract infection need not be considered clinically serious.

[22] Have symptomatic herpes simplex at the time of randomization.

[23] Have had symptomatic herpes zoster infection within 12 weeks prior to randomization.

[24] Have a history of disseminated/complicated herpes zoster (for example, ophthalmic zoster or CNS involvement).

[25] Have a positive test for hepatitis B virus (HBV) defined as:

a. positive for hepatitis B surface antigen (HBsAg), or

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

Note: Patients who are HBcAb-positive and HBV DNA-negative may be enrolled in the study but will require additional HBV DNA monitoring during the study.

[26] Have hepatitis C virus (HCV) infection (hepatitis C antibody-positive and HCV ribonucleic acid [RNA]-positive).

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

[27] Have evidence of HIV infection and/or positive HIV antibodies.

[28] Have had household contact with a person with active TB and did not receive appropriate and documented prophylaxis for TB.

[29] Have evidence of active TB or latent TB

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

- Positive purified protein derivative (PPD) test (≥5 mm induration between approximately 2 and 3 days after application, regardless of vaccination history), medical history, clinical features, and abnormal chest x-ray at screening.
- 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, have no clinical features of active TB, and have a screening chest x-ray with no evidence of active TB may be enrolled if other entry criteria 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 (i.e., chest imaging performed within the past 6 months will not be accepted).

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

- Positive PPD test, no clinical features consistent with active TB, and a chest x-ray with no evidence of active TB at screening; or
- If the PPD test is positive and 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 prior to 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, have no clinical features of active TB, and have a screening chest x-ray with no evidence of active TB may be enrolled if other entry criteria 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 (i.e., chest imaging performed within the past 6 months will not be accepted).

[30] Have received parenteral [i.e., intravenous, intramuscular, intra-articular] corticosteroids within 6 weeks of screening (Visit 1), or are expected to require parenteral corticosteroids during the study.

[31] Have received any of the following medications:

a. Biologic treatments for immunologic disease such as etanercept, infliximab, certolizumab, adalimumab, golimumab, tocilizumab, abatacept, ustekinumab, ixekizumab, secukinumab, or anakinra within 4 weeks of screening.

b. Cyclophosphamide (or any other cytotoxic agent), belimumab, or anifrolumab (or another anti-IFN therapy) within 12 weeks of screening.

c. Rituximab, any other B cell depleting therapies, or intravenous immunoglobulin (IVIg) within 24 weeks of screening.

[32] Have received a JAK inhibitor.

[33] Have been treated with probenecid that cannot be discontinued for the duration of the study.

[34] Have received plasmapheresis within 12 weeks of screening.

[35] Have been exposed to a live vaccine within 12 weeks of 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 have not previously received the herpes zoster vaccine by screening will be encouraged (per local guidelines) to do so prior to randomization; vaccination with live herpes zoster vaccine must occur >4 weeks prior to randomization and start of investigational product. Patients will not be randomized if they were exposed to a live herpes zoster vaccination within 4 weeks of planned randomization. Investigators should review the vaccination status of their patients and follow the local guidelines for vaccination of patients ≥18 years of age with nonlive vaccines intended to prevent infectious disease prior to entering patients into the study.

[36] Are currently enrolled in or have discontinued within 4 weeks of screening from, any other clinical trial involving an investigational product or nonapproved use of a drug or device or any other type of medical research judged not to be scientifically or medically compatible with this study.

[37] Have previously completed or been randomized and withdrawn from this study or have received baricitinib in any other study.

Diagnostics Assessments

[38] Have screening laboratory test values, including thyroid-stimulating hormone (TSH), outside the reference range for the population that, in the opinion of the investigator, pose an unacceptable risk for the patient's participation in the study. 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 have TSH marginally outside the laboratory's normal reference range and are receiving stable thyroxine replacement therapy may participate if the treating physician has documented that the thyroxine replacement therapy is adequate for the patient.

[39] Have any of the following specific abnormalities on screening laboratory tests from the central laboratory:

a. ALT or AST >2 x ULN

b. alkaline phosphatase (ALP) ≥2 x ULN

c. total bilirubin ≥1.5 x ULN

d. hemoglobin <9 g/dL (90.0 g/L)

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

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

g. lymphopenia (lymphocyte count <500 cells/µL) (<0.50 x 10^3/µL or <0.50 GI/L)

h. thrombocytopenia (platelets <50,000 cells/µL) (<50 x 10^3/µL or <50 GI/L)

In the case of any of the aforementioned laboratory abnormalities, the tests may be repeated once by the central laboratory during screening, and values resulting from repeat testing may be accepted for enrollment eligibility if they meet the eligibility criterion.

Other Exclusions

[40] Are largely or wholly incapacitated permitting little or no self-care, such as being bedridden or confined to wheelchair.

[41] In the opinion of the investigator, are at an unacceptable risk for participating in the study.

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

[43] Have a history of intravenous drug abuse, other illicit drug abuse, or chronic alcohol abuse within the 2 years prior to screening or are concurrently using, or expected to use during the study, illicit drugs (including marijuana).

[44] Are unable or unwilling to make themselves available for the duration of the study and/or are unwilling to follow study restrictions/procedures.

[45] 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.

[46] Are Lilly or Incyte employees or their designee.

6.3. Lifestyle Restrictions

Study participants should be instructed not to donate blood or blood products during the study or for 30 days following the last dose of investigational product.

6.4. Screen Failures

Individuals who do not meet the criteria for participation in this study (screen failure) may be rescreened one time. The interval between initial screen failure and rescreening Visit 1 should be at least 4 weeks. When rescreening is performed the individual must sign a new ICF, repeat all screening procedures as described in the Schedule of Activities, and will be assigned a new identification number.

7. Treatments

7.1. Treatments Administered

This study involves a comparison of baricitinib tablets 4-mg or 2-mg administered orally QD with placebo tablets administered orally QD. Table JAHZ.3 shows the treatment regimens.

Table JAHZ.3. Treatment Regimens

Treatment Group	Treatments Administered Day 1 through Day 365
Baricitinib 4-mg QD[a]	1 baricitinib 4-mg tablet and 1 placebo tablet matching baricitinib 2-mg
Baricitinib 2-mg QD	1 baricitinib 2-mg tablet and 1 placebo tablet matching baricitinib 4-mg
Placebo QD	2 placebo tablets: 1 matching baricitinib 4-mg and 1 matching baricitinib 2-mg

Abbreviations: QD = once daily.

a Patients with estimated glomerular filtration rate <60 mL/min/1.73 m^2 at screening who are randomized to the 4-mg dose will receive a dose of 2-mg QD

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
- at the end of the study returning all unused medication to Lilly, or its designee, 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 Labelling

Lilly (or its designee) will provide the following primary study materials:

- tablets containing 4-mg of baricitinib
- tablets containing 2-mg of baricitinib
- tablets containing placebo to match 4-mg baricitinib
- tablets containing placebo to match 2-mg baricitinib

Investigational product will be dispensed to the patient at the principal investigator's study site. Investigational product packaging will contain enough tablets for the longest possible interval between visits.

Investigational product will be labeled according to the country's regulatory requirements. **Patients will be instructed to take 2 tablets, one tablet from each bottle, each day**.

All investigational products will be stored, inventoried, reconciled, and destroyed according to applicable regulations. Investigational products will be supplied by Lilly or its representative in accordance with current Good Manufacturing Practices and will be supplied with lot numbers, expiry dates, and certificates of analysis (as applicable).

7.2. Method of Treatment Assignment

Patients who meet all criteria for enrollment will be randomized in a 1:1:1 ratio (baricitinib 4-mg; baricitinib 2-mg; placebo) to double-blind treatment at Visit 2. Randomization will be stratified by disease activity (total SLEDAI-2K at screening <10; ≥10), corticosteroid dose (<10 mg/day; ≥10 mg/day prednisone or equivalent), and region (defined in the SAP).

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 packages containing double-blind investigational product to each patient, starting at Visit 2 (Week 0), and at each visit subsequent through Visit 15 (Week 48). Each package of clinical trial material will supply sufficient study medication at least for the number of weeks between visits. Site personnel will confirm that they have located the correct packages by entering a confirmation number found on the packages of investigational product into the IWRS before dispensing the packages to the patient.

7.2.1. Selection and Timing of Doses

Investigational product will be provided to patients following randomization at Visit 2 (Week 0).

Two tablets (1 tablet from each bottle) should be taken orally each day, without regard to food and, if possible, at approximately the same time every day. Special timing considerations of administration of investigational product relative to PK sample collection at study visits are described in Section 9.5.

7.2.2. Dose Adjustment for Renal Impairment

The dose adjustment for renal impairment will be managed by IWRS to ensure maintenance of the treatment blind. The eGFR value from screening will be entered into IWRS at Visit 2, and IWRS will assign the treatment doses accordingly.

- Patients with documented renal impairment (defined as screening eGFR ≥40 to <60 mL/min/1.73 m^2), who are randomized to active treatment (either to the baricitinib 4-mg arm or the baricitinib 2-mg arm) will receive a baricitinib 2-mg QD dose by the IWRS.
- No dose adjustment will be made for patients with screening eGFR ≥60 mL/min/1.73 m^2.

The rationale of dose adjustment for patients with documented renal impairment is detailed in Section 5.5.

Interruption criteria due to eGFR during the treatment period is discussed in Section 8.1.1.

7.3. Blinding

This is a double-blind study. To preserve the blinding of the study, a minimum number of Lilly personnel who are not directly involved with investigational sites will see the randomization table and treatment assignments before the study is complete.

In case of an emergency, the investigator has the sole responsibility for determining if unblinding of a patient's treatment assignment is warranted. Patient safety must always be the first consideration in making such a determination. If the investigator decides that unblinding is warranted, the investigator should make every effort to contact the sponsor-designated medical monitor prior to unblinding a patient's treatment assignment unless this could delay emergency treatment of the patient. If a patient's treatment assignment is unblinded, Lilly must be notified immediately.

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 calls resulting in an unblinding event 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-designated medical monitor for the patient to continue in the study.

7.4. Dosage Modification

During this study, treatment assignment and dose adjustment of investigational product will be followed strictly as indicated in Sections 7.1 and 7.2.

See Section 7.7 for permitted concomitant therapy.

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.

Storage and handling instructions are provided on the investigational product packaging and should be followed.

7.6. Treatment Compliance

Patient compliance with investigational product will be assessed at Visit 3 through Visit 16 and at Early Termination during the treatment period by counting returned tablets. Deviations from the prescribed dosage regimen should be recorded in the electronic case report form (eCRF).

A patient treated with baricitinib or placebo will be considered significantly noncompliant if he or she misses >20% of the prescribed doses during the treatment period, unless the patient's investigational product was 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 more than the prescribed amount of investigational product. Patients found to be noncompliant with the investigational product should be assessed to determine the reason for noncompliance and educated and/or managed as deemed appropriate by the investigator to improve compliance.

7.7. Concomitant Therapy

Patients will maintain their usual medication regimen for SLE and for any other concomitant diseases throughout the study unless specifically excluded in the protocol (see Section 6.2, Exclusion Criteria). Patients taking these medications should be on chronic stable doses at the time of randomization, as specified by the Inclusion Criteria (Section 6.1, Inclusion Criterion [8]).

Permitted concomitant background standard of care medications for SLE include:

- A single antimalarial (such as hydroxychloroquine, chloroquine, quinacrine) at a stable therapeutic dose for at least 8 weeks prior to screening (Visit 1). Doses should remain stable throughout the treatment period.
- A single immunosuppressant (such as MTX, azathioprine, mycophenolate, tacrolimus, leflunomide, cyclosporine) at a stable therapeutic dose for at least 8 weeks prior to screening (Visit 1). Doses should remain stable throughout the treatment period.
- An oral corticosteroid, initiated at least 4 weeks prior to screening (Visit 1), at a stable dose ≤40 mg/day prednisone (or equivalent) for at least 2 weeks prior to screening (Visit 1) and through baseline (Visit 2). If the patient is not receiving an antimalarial or immunosuppressant, the dose of corticosteroid must be ≥7.5 mg/day prednisone (or equivalent). See Section 7.7.1 for permitted changes to corticosteroid dosing.

Any changes to the patient's medication must be discussed with the investigator. Patients should be instructed to consult the investigator or other appropriate study personnel at the site before taking any new medications or supplements.

Additional medications are to be avoided during the study unless required to treat an AE or for the treatment of an ongoing medical problem. Between Weeks 0 and 12 and between Weeks 24 and 36, steroids for non-SLE reasons may be permitted on a short-term (no more than 14 days) basis. Patients must be back to prior dose within 14 days (or by Week 14/Week 38) from the initiation of the high-dose steroids for non-SLE reasons (whichever is sooner). If the need for concomitant medication arises, inclusion or continuation of the patient may be at the discretion of the investigator after consultation with Lilly or its designee. Any additional medication whether prescription or over-the-counter, used at baseline and/or during the course of the study must be documented in the eCRF with start and stop dates.

7.7.1. Permitted Changes and Additions to Concomitant Therapy

Changes to corticosteroid dosing are permitted during the treatment period as follows.

Corticosteroid Tapering

Investigators are encouraged to taper corticosteroids according to clinical practice with a goal of ≤7.5 mg/day prednisone (or equivalent) by Week 40 (Visit 13).

Tapering may occur from Week 0 through Week 16 and from Week 24 (after completion of Visit 9) through Week 40 (see Figure JAHZ.1). In most cases, tapering should not be considered for at least 4 weeks after randomization, consistent with clinical practice. If a patient experiences an increase in disease activity during an attempted taper, the patient's corticosteroid dose may be increased once by up to 10 mg/day or up to the baseline dose (whichever is lower). The corticosteroid dose may not exceed the baseline dose.

Corticosteroid dose must be stable for 8 weeks prior to Week 24 (i.e., no changes between Weeks 16 and 24) and for 12 weeks prior to Week 52 (i.e., no changes between Weeks 40 and 52). This design minimizes confounding of assessments at Week 24 and Week 52 (see Figure JAHZ.1).

Corticosteroid Burst

A single corticosteroid burst is permitted for patients with increased disease activity early the treatment period.

During Weeks 0 through 12, patients with intolerable disease activity may receive a single corticosteroid burst of up to 40 mg/day oral prednisone for up to 7 days (in addition to the regular dose, up to a maximum of 60 mg/day prednisone or equivalent total daily dose). The patient must be at or below their preburst corticosteroid dose by the end of the 7-day burst period, which may not extend beyond Week 12. As an alternative to the oral corticosteroid burst, a single 80-mg intramuscular dose of methylprednisone (or equivalent) may be given during Weeks 0 through 11.

7.7.2. Prohibited Changes in Concomitant Therapy

Initiation of or increase in dose of an antimalarial or immunosuppressant is not permitted during the treatment period. Patients requiring an initiation or increase above baseline dose of antimalarials or immunosuppressants during the treatment period will be discontinued from the study drug (see Section 8.1.1).

Patients requiring an increase above the baseline dose of corticosteroid during the treatment period, other than the permitted burst, will be discontinued from study drug (see Section 8.1.1).

7.8. Treatment after the End of the Study

7.8.1. Study Extensions

Patients who complete this study through Week 52 (Visit 16) may be eligible to participate in Study JAIM, if enrollment criteria for Study JAIM are met.

7.8.2. Treatment after Study Completion

Baricitinib will not be made available after conclusion of the study to patients except as described in Section 7.8.1.

8. Discontinuation Criteria

8.1. Discontinuation from Study Treatment

8.1.1. Temporary Interruption of 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. It is recommended that the investigator consult with Lilly (or its designee) before temporarily interrupting therapy for reasons other than those defined in Table JAHZ.4. Retest timing and frequency is at the investigator's discretion.

Investigational product that was temporarily interrupted because of an AE or abnormal laboratory value not specifically covered in Table JAHZ.4 may be restarted at the discretion of the investigator. Investigational product must be held in the following situations and may only be resumed as noted in the table.

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

Hold Investigational Product if the Following Abnormalities Occur:	Investigational Product May be Resumed When:
WBC count <2000 cells/µL (<2.00 x 10^3/µL or <2.00 GI/L)	WBC count ≥2500 cells/µL (≥2.50 x 10^3/µL or ≥2.50 GI/L)
ANC <1000 cells/µL (<1.00 x 10^3/µL or <1.00 GI/L)	ANC ≥1200 cells/µL (≥1.2 x 10^3/µL or ≥1.2 GI/L)
Lymphocyte count <300 cells/µL (<0.30 x 10^3/µL or <0.30 GI/L)	Lymphocyte count ≥500 cells/µL (≥0.50 x 10^3/µL or ≥0.50 GI/L)
Platelet count <25,000/µL (<25 x 10^3/µL or <25 GI/L)	Platelet count ≥50,000/µL (≥50 x 10^3/µL or ≥50 GI/L)
eGFR <40 mL/min/1.73 m^2 for patients with screening eGFR ≥60 mL/min/1.73 m^2	eGFR ≥50 mL/min/1.73 m^2
eGFR <30 mL/min/1.73 m^2 for patients with screening eGFR ≥40 to <60 mL/min/1.73 m^2	eGFR ≥40 mL/min/1.73 m^2
ALT or AST >5x ULN	ALT and AST return to <2 x ULN, and IP is not considered to be the cause of enzyme elevation
Hemoglobin <8 g/dL (<80.0 g/L)	Hemoglobin ≥9 g/dL (≥90.0 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 withheld[a]	Resolution of infection that, in the opinion of the investigator, merits the IP being restarted

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.

a Permanent discontinuation of IP should be considered for patients who develop a serious infection that, in the opinion of the investigator, would pose an unacceptable risk if IP were resumed.

8.1.2. Permanent Discontinuation from Study Treatment

Possible reasons leading to permanent discontinuation of investigational product:

- **Subject Decision**
  - the patient requests to discontinue investigational product.
- **Investigator Decision**
  - the investigator decides that the patient should be discontinued from investigational product.
- **Discontinuation due to a hepatic event or liver test abnormality.** Patients who are discontinued from investigational product due to a hepatic event or liver test abnormality should have additional hepatic safety data collected via eCRF.

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

- ALT or AST >8 x ULN
- ALT or AST >5 x ULN for more than 2 weeks after temporary interruption of investigational product
- ALT or AST >3 x ULN and either total bilirubin level (TBL) >2 x ULN or international normalized ratio (INR) >1.5
- ALT or AST >3 x ULN with the appearance of fatigue, nausea, vomiting, right upper-quadrant pain, or tenderness, fever, rash, and/or eosinophilia (>5%)
- alkaline phosphatase (ALP) >3 x ULN
- ALP >2.5 x ULN and TBL >2 x ULN
- ALP >2.5 x ULN with the appearance of fatigue, nausea, vomiting, right quadrant pain or tenderness, fever, rash, and/or eosinophilia (>5%)

- **Discontinuation due to other laboratory abnormalities:**
  - white blood cell count <1000 cells/µL ($1.00x10^3$/µL or 1.00 GI/L)
  - absolute neutrophil count <500 cells/µL ($0.50x10^3$/µL or 0.50 GI/L)
  - lymphocyte count <200 cells/µL ($0.20x10^3$/µL or 0.20 GI/L)
  - hemoglobin <6.5 g/dL (<65.0 g/L).

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 or another identified factor, laboratory tests may be repeated. The investigator may be able to restart investigational product after consultation with the Lilly-designated medical monitor, only when the laboratory value meets resumption thresholds (Table JAHZ.4) following the resolution of the intercurrent illness or other identified factor.

Discontinuation due to other circumstances:

- pregnancy
- malignancy (except for successfully treated basal cell or squamous epithelial skin cancers)
- hepatitis B virus DNA is detected with a value above the lower limit of quantitation (see Section 9.4.8).
- occurrence of a VTE (DVT/PE) during the study
- serious infection that in the opinion of the investigator, merits the investigational product being discontinued

- if the patient, for any reason, requires treatment with another therapeutic agent that may be effective for treatment of SLE that is listed in the exclusion criteria (see Section 6.2) or is noncompliant to the concomitant therapy requirements (see Section 7.7.2) during the study. Discontinuation from the investigational product must occur prior to introduction of the new agent.

Throughout the trial, investigators should continue to assess benefit/risk for patients to remain in the trial and should consider discontinuing patients if sufficient clinical benefit is not observed with protocol allowed concomitant treatments.

If a patient prematurely discontinues investigational product for any reason, the patient is encouraged to remain in the study through Week 52 (Visit 16) and follow the regular visit schedule per Section 2 (Schedule of Activities).

8.1.3. Discontinuation of Inadvertently Enrolled Patients

If the sponsor or investigator identify 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-designated medical monitor agree it is medically appropriate to continue, the investigator must obtain documented approval from the sponsor-designated medical monitor 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 the protocol.

8.2. Discontinuation from the Study

Patients will be discontinued in the following circumstances:

- 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 (GCP)
- Subject 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 as indicated in the Early Term Visit per Section 2 (Schedule of Activities), Section 9.2 (Adverse Events), and Section 9.4 (Safety) of this protocol and proceed to the post-treatment follow up period.

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 lists 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 Assessment

The SRI-4 is a composite index used to assess disease activity in SLE. The SLEDAI-2K component is used to capture clinically meaningful improvement in disease activity, while the BILAG and Physician's Global Assessment of Disease Activity components ensure that the improvement in overall disease is not accompanied by disease worsening in other organ systems. The SRI-4 response is defined as follows:

- Reduction of ≥4 points from baseline in SLEDAI-2K score (Section 9.1.1.1); and
- No new BILAG A or no more than 1 new BILAG B disease activity scores (Section 9.1.1.2); and
- No worsening (defined as an increase of ≥0.3 points [10 mm] from baseline) in Physician's Global Assessment of Disease Activity (Section 9.1.1.3).

9.1.1.1. Systemic Lupus Erythematosus Disease Activity Index-2000

The SLEDAI-2K is a validated global disease activity instrument that focuses on high-impact disease manifestations across 9 organ systems. It includes 24 clinical and laboratory variables with manifestations graded by the affected organ system. CNS: Seizure, Psychosis, Organic Brain Syndrome, Visual Disturbance, Cranial Nerve Disorder, Lupus Headache, cerebrovascular attack; Vascular: Vasculitis; Musculoskeletal: Arthritis, Myositis; Renal: Urinary Casts, Hematuria, Proteinuria, Pyuria; Mucocutaneous: Rash, Alopecia, Mucosal Ulcers; Cardiovascular and Respiratory: Pleurisy, Pericarditis; Immunologic: Low complement, Increased DNA Binding; Constitutional: Fever; Hematologic: Thrombocytopenia, Leukopenia.

9.1.1.2. British Isles Lupus Assessment Group 2004 Index

The BILAG-2004 Index is a validated global disease activity index designed on the basis of the physician's intention to treat (ITT), focusing on changes in disease manifestations (not present, improving, same, worse, or new) occurring in the last 4 weeks compared with the previous 4 weeks. The instrument assesses 97 clinical signs, symptoms, and laboratory parameters across 9 organ system domains: constitutional, mucocutaneous, neuropsychiatric, musculoskeletal, cardiorespiratory, gastrointestinal, ophthalmic, renal, and hematology. A BILAG A disease activity score is severe disease activity requiring high-dose oral or intravenous corticosteroids, immunomodulators, or high-dose anticoagulation along with high-dose corticosteroids or

immunomodulators. A BILAG B disease activity score is moderate disease activity requiring low-dose oral corticosteroids, intramuscular or intra-articular corticosteroid injections, topical corticosteroids or immunomodulators, antimalarials, or symptomatic therapy. BILAG C corresponds to stable mild disease, BILAG D is inactive disease that was active previously, and BILAG E indicates the system was never involved.

9.1.1.3. Physician's Global Assessment of Disease Activity

The Physician's Global Assessment of Disease Activity is the physician's assessment of the patient's overall disease activity due to SLE, as compared with all possible patients with SLE. The Physician's Global Assessment of Disease Activity is scored using a 100-mm Visual Analog Scale (VAS), where 0 mm (measured from the left starting point of the line) indicates no disease activity and 100 mm (measured from the left starting point of the line) indicates the most severe disease activity possible for all patients with SLE (or death). The Physician's Global Assessment of Disease Activity score is indicated by making a vertical tick mark on the line between 0 and 100 mm. There are benchmarks of 0 (0 mm), 1 (33 mm), 2 (67 mm), and 3 (100 mm) on the line corresponding to no, mild, moderate, and severe SLE disease activity, respectively.

9.1.2. Secondary and Exploratory Efficacy Assessments

Secondary efficacy assessments will include the following:

9.1.2.1. Systemic Lupus Erythematosus Responder Index

The SRI-4 will be assessed at various time points up to Week 52. For assessment description, see Section 9.1.1. The SRI-5, -6, -7, and -8 will also be assessed at various time points up to Week 52, and are similar to the SRI-4 except they require a reduction of ≥5, 6, 7, or 8 points (respectively) from baseline in SLEDAI-2K score (Section 9.1.1.1).

9.1.2.2. Systemic Lupus Erythematosus Disease Activity Index-2000

The SLEDAI-2K will be assessed at various time points up to Week 52. For assessment description, see Section 9.1.1.

9.1.2.3. British Isles Lupus Assessment Group 2004 Index

The BILAG-2004 Index will be assessed at various time points up to Week 52. For assessment description, see Section 9.1.1.

9.1.2.4. Physician's Global Assessment of Disease Activity

The Physician's Global Assessment of Disease Activity will be assessed at various time points up to Week 52. For assessment description, see Section 9.1.1.

9.1.2.5. Safety of Estrogens in Lupus Erythematosus National Assessment (SELENA)-Systemic Lupus Erythematosus Disease Activity Index (SLEDAI) Flare Index

The SELENA-SLEDAI Flare Index (SFI) uses the SLEDAI score, disease activity scenarios, treatment changes, and Physician's Global Assessment of Disease Activity to define mild/moderate and severe flares. The index takes into account the absolute change in total

scores, new or worsening symptoms, and increases in medication use or hospitalization due to the disease activity.

9.1.2.6. Cutaneous Lupus Erythematosus Disease Area and Severity Index

The Cutaneous Lupus Erythematosus Disease Area and Severity Index (CLASI) is a validated scale used to assess cutaneous manifestations of SLE consisting of 2 scores. The first summarizes the activity of the disease while the second is a measure of the damage done by the disease. Activity is scored on the basis of erythema, scale/hyperkeratosis, mucous membrane involvement, acute hair loss, and nonscarring alopecia. Damage is scored in terms of dyspigmentation and scarring, including scarring alopecia.

9.1.2.7. Lupus Low Disease Activity State

The LLDAS is defined as a low level of disease activity attained with or without use of low-dose steroids and/or tolerated standard maintenance doses of SoC medications.

9.1.2.8. Tender/Swollen Joint Count (28 Joints)

The 28 joints to be examined and assessed as tender or not tender for tender joint count and as swollen or not swollen for swollen joint count include 14 joints on each side of the patient's body: the 2 shoulders, the 2 elbows, the 2 wrists, the 10 metacarpophalangeal joints, the 2 interphalangeal joints of the thumb, the 8 proximal interphalangeal joints, and the 2 knees.

9.1.2.9. Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index

The SLICC/ACR Damage Index is scored on 41 items representing damage to 12 organ systems. The index records damage occurring in patients with SLE regardless of its cause and includes specific comorbidities associated with SLE that may be due to treatment-related toxicity.

9.1.2.10. Health Outcomes and Quality of Life Measures

The self-reported questionnaires will be administered via either an electronic patient diary or via an electronic tablet as specified in the Schedule of Activities (Section 2) in countries where the questionnaires have been translated into the native language of the region and linguistically validated.

9.1.2.10.1. Worst Pain Numeric Rating Scale

The Worst Pain Numerical Rating Scale (NRS) is a single-item, patient-administered, 11-point horizontal scale anchored at 0 and 10, with 0 representing "no pain" and 10 representing "pain as bad as you can imagine." Overall severity of a patient's pain is indicated by selecting the number that best describes the worst level of pain in the last 24 hours. Data will be captured on a daily electronic patient diary.

9.1.2.10.2. Worst Joint Pain Numeric Rating Scale

The Worst Joint Pain NRS is a single-item, patient-administered, 11-point horizontal scale anchored at 0 and 10, with 0 representing "no joint pain" and 10 representing "joint pain as bad as you can imagine." Overall severity of a patient's joint pain is indicated by selecting the number that best describes the worst level of joint pain in the last 24 hours. Data will be captured on a daily electronic patient diary.

9.1.2.10.3. Worst Fatigue Numeric Rating Scale

The Worst Fatigue NRS is a single-item, patient-administered, 11-point horizontal scale anchored at 0 and 10, with 0 representing "no fatigue" and 10 representing "as bad as you can imagine." Overall severity of a patient's fatigue is indicated by selecting the number that best describes the worst level of fatigue in the last 24 hours. Data will be captured on a daily electronic patient diary.

9.1.2.10.4. Joint Stiffness Duration

The Duration of Joint Stiffness patient-reported outcomes (PRO) is a single-item, patient-administered scale designed to capture information on the self-reported length of time, in minutes, that a patient's joint stiffness lasted each day. Patients report the duration of time by entering the number of hours and minutes their joint stiffness lasted today. Data will be captured on a daily electronic patient diary.

9.1.2.10.5. Joint Stiffness Severity Numeric Rating Scale

The Joint Stiffness Severity NRS is a single-item, patient-administered, 11-point horizontal scale anchored at 0 and 10, with 0 representing "no joint stiffness" and 10 representing "joint stiffness as bad as you can imagine." Overall severity of a patient's joint stiffness is indicated by selecting the number that best describes the worst level of joint stiffness in the last 24 hours. Data will be captured on a daily electronic patient diary.

9.1.2.10.6. Patient's Global Impression of Severity

The Patient's Global Impression of Severity is a single-item question asking the patient how they would rate their overall lupus symptoms over the last 24 hours. The 5 categories of response range from "no symptoms" to "severe". Data will be captured on a daily electronic patient diary.

9.1.2.10.7. Patient's Global Impression of Change

The Patient's Global Impression of Change is a single-item question asking the patient how they would rate their change in overall lupus symptoms since they started the study. The 7 categories of response range from "very much better" to "very much worse." Data will be captured on an electronic tablet collected at site visits.

9.1.2.10.8. Medical Outcomes Short-Form 36-Item Health Survey Version 2 (SF-36v2)

The SF-36v2 Acute measure is a subjective, generic, health-related quality of life instrument that is patient-reported and consists of 36 questions covering 8 health domains: physical functioning, bodily pain, role limitations due to physical problems, role limitations due to emotional problems, general health perceptions, mental health, social function, and vitality. Each domain is scored by summing the individual items and transforming the scores into a 0 to 100 scale with higher scores indicating better health related quality of life. In addition, 2 summary scores, the PCS and the MCS, will be evaluated based on the 8 SF-36v2 Acute domains. The acute version of this instrument has a 1 week recall period (Brazier et al. 1992; Ware and Sherbourne 1992). Data will be captured on an electronic tablet collected at site visits.

9.1.2.10.9. Functional Assessment of Chronic Illness Therapy–Fatigue Scale

The FACIT-Fatigue scale (Cella and Webster 1997) is a brief, 13-item, symptom-specific questionnaire that specifically assesses the self-reported severity of fatigue and its impact upon daily activities and functioning. The FACIT-Fatigue uses 0 ("not at all") to 4 ("very much") NRS to assess fatigue and its impact in the past 7 days. Scores range from 0 to 52 with higher scores indicating less fatigue. Data will be captured on an electronic tablet collected at site visits.

9.1.2.10.10. Work Productivity and Activity Impairment Questionnaire–Lupus

The Work Productivity and Activity Impairment Questionnaire-Lupus (WPAI-Lupus) records impairment due to Lupus during the past 7 days. The WPAI-Lupus consists of 6 items grouped into four domains, absenteeism (work time missed), presenteesism (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. Data will be captured on an electronic tablet collected at site visits.

9.1.2.10.11. 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 in order to provide 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/her current health state using a 0- to 100-mm VAS. The descriptive system comprises the following 5 dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each dimension has 5 levels: no problems, slight problems, moderate problems, severe problems, and extreme problems. The respondent is asked to indicate his or her health state by ticking (or placing a cross) in the box associated with the most appropriate statement in each of the 5 dimensions. It should be noted that the numerals 1 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 (EuroQol Group 2015 [WWW]). Data will be captured on an electronic tablet collected at site visits.

9.1.3. Appropriateness of Assessments

All assessments and/or concepts included in this study are standard, widely used, and generally recognized as reliable, accurate, and relevant.

9.2. Adverse Events

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

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

Investigators must document their review of each laboratory safety report.

The investigator remains responsible for following, through an appropriate health care option, AEs that are serious or otherwise medically important, considered related to the investigational product or the study, or that caused the patient to discontinue the investigational product before completing the study. The patient should be followed until the event resolves or stabilizes with appropriate diagnostic evaluation. 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, except SLE, as it is 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 the following via eCRF for each AE: start date, stop date (if applicable), severity, and their assessment of the potential relatedness of each AE to protocol procedure or 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 one of the following outcomes:

- death
- initial or prolonged inpatient hospitalization
- a life-threatening experience (that is, immediate risk of dying)
- persistent or significant disability/incapacity
- congenital anomaly/birth defect

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

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 according to SAE-reporting requirements and timelines (see Section 9.2) 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 paper SAE form. If alerts are issued via telephone, they are to be immediately followed with official notification on study-specific SAE forms. The SAE form should be completed by the investigator, and submitted via fax to the Sponsor's global patient safety department. This form includes a fax cover page that is pre-populated with the appropriate fax number. The 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 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 disposition 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 guidance or national regulatory requirements in participating countries require the reporting of SUSARs. Lilly has procedures that will be followed for the identification, recording and expedited reporting of SUSARs that are consistent with global regulations and the associated detailed guidance.

9.2.2. Adverse Events of Special Interest

Adverse events of special interest (AESIs) will include the following:

- infections (including TB, herpes zoster, or opportunistic infections)
- malignancies
- hepatic events (see Section 9.4.9)

- major adverse cardiac events (MACE) (see Section 9.4.10)
- venous thromboembolic events (DVT/PE) (see Section 9.4.10.1)

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 in order to ensure the safety of study participants, monitor quality, and to facilitate process and product improvements.

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

9.3. Treatment of Overdose

Refer to the IB.

9.4. Safety

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

9.4.1. Electrocardiograms

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

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

9.4.2. Vital Signs

For each patient, vital signs measurements (sitting blood pressure, heart rate, and temperature) should be conducted according to the Schedule of Activities (Section 2). Subjects should be seated and relaxed with both feet on the floor for at least 5 minutes prior to taking measurements. Three replicate blood pressure readings should be made at each time point at approximately 30- to 60-second intervals. A single-pulse measurement should be taken simultaneously with at least one of the blood pressure readings. Blood pressure and pulse measurements should be made using either automated or manual equipment. If measurements are machine averaged, the average blood pressure reading should be recorded on the CRF. If measurements are manual or the machine does not provide an average reading, then each individual reading should be recorded on the CRF. Measurements should be made before any scheduled blood draws. Additional measurements of vital signs may be performed at the discretion of the investigator.

9.4.3. Physical Examination

For each patient, physical examinations will be conducted according to the Schedule of Activities (Section 2). One complete physical examination (excluding pelvic and rectal examinations) will be performed at Visit 1 (Screening). All remaining physical examinations throughout the study should include a symptom-directed physical examination. 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 Columbia-Suicide Severity Rating Scale (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 prior to the collection of the C-SSRS. If a suicide-related event is discovered during the C-SSRS but was not captured during the nonleading AE collection, sites should not change the AE form. If an event is serious or leads to discontinuation, this is an exception where the SAE and/or AE leading to discontinuation should be included on the AE form and the process for reporting SAEs should be followed.

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.

9.4.6. 16-Item Quick Inventory of Depressive Symptomatology: Self-Report

The QIDS-SR16 is a self-administered, 16-item instrument intended to assess the existence and severity of symptoms of depression as listed in the American Psychiatric Association's Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (DSM-V) (APA 2013). A patient is asked to consider each statement as it relates to the way they have felt for the past 7 days. There is a 4-point scale for each item ranging from 0 to 3. The 16 items corresponding to 9 depression domains are summed to give a single score ranging from 0 to 27, with higher scores denoting greater symptom severity. The domains assessed by the instrument include (1) sad mood, (2) concentration, (3) self-criticism, (4) suicidal ideation, (5) interest, (6) energy/fatigue, (7) sleep disturbance (initial, middle, and late insomnia or hypersomnia), (8) decrease/increase in appetite/weight, and (9) psychomotor agitation/retardation.

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 chest imaging (other than ultrasound) obtained within 6 months prior to the study are available. The chest x-ray will be reviewed by the investigator or his or her designee to exclude patients with active TB infection. In addition, patients will be tested at screening (Visit 1) for evidence of active or latent TB as described in the exclusion criteria, Section 6.2.

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

9.4.8. Hepatitis B Virus DNA Monitoring

Hepatitis B virus DNA testing will be performed in enrolled patients who tested positive for HBcAb at screening.

Patients who are HBcAb-positive and HBV DNA-negative (undetectable) at Visit 1 will require HBV DNA monitoring every 3 months and at the patient's last visit, regardless of their hepatitis B surface antibody (HBsAb) status.

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

- If a single result is obtained with a value "below limit of quantitation," the test should be repeated within approximately 2 weeks.
- If the repeat test result is "target not detected," monitoring may resume according to the study schedule.
- If the patient has 2 or more test results with a value "below limit of quantitation" during the study, HBV DNA testing should be performed approximately once per month for the remainder of the study and referral to a hepatologist is recommended.

- If a result is obtained with a value above limit of quantitation at any time during the study, the patient will be permanently discontinued from investigational product (see to Section 8.1.2) and should be referred to a hepatology specialist immediately.
    - In selected cases, investigators may temporarily continue investigational product in accordance with current immunomodulator management in the setting of HBV DNA positivity. This option may be considered in consultation with Lilly (or its designee) and after evaluation of individual patient risks and benefits.

9.4.9. Hepatic Safety Monitoring

If a study patient experiences elevated ALT ≥3 x ULN, ALP ≥2 x ULN, or elevated TBL ≥2 x 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 and 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.

Discontinuation criteria of investigational products, either temporary interruption or permanent discontinuation, due to abnormal ALT, AST, TBL, or ALP, are detailed in Section 8.1.2.

Hepatic Safety Data Collection

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

- Elevation of serum ALT to ≥5 x ULN on 2 or more consecutive blood tests
- Elevated serum TBL to ≥ 2 x ULN (except for cases of known Gilbert's syndrome)
- Elevation of serum ALP to ≥2 x 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 a SAE

9.4.10. Safety Monitoring

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

The Lilly-designated medical monitor will monitor safety data throughout the course of the study. Lilly will review SAEs within time frames mandated by company procedures. The Lilly-designated medical monitor will, as is appropriate, consult with the functionally independent Global Patient Safety (GPS) 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 its 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 clinical AE is deemed serious, unexpected, and possibly related to investigational product, Lilly GPS 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 events. Adverse event 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, myocardial infarction, stroke), other cardiovascular events (such as hospitalization for unstable angina, hospitalization for heart failure, serious arrhythmia, resuscitated sudden death, cardiogenic shock, coronary interventions), VTE (DVT/PE), arterial thromboembolic events 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.

9.4.10.1. Venous Thromboembolic Event Assessment

If a patient develops clinical features of a DVT or PE, appropriate local laboratory tests and imaging should be performed, as necessary, for the diagnosis of the event. For confirmed cases, additional laboratory testing is recommended as outlined in Appendix 5. All suspected VTE events will be independently adjudicated by a blinded Clinical Event Committee.

9.5. Pharmacokinetics

A single venous blood sample will be drawn at the times indicated in the Schedule of Activities (Section 2). These blood samples will be used to determine the plasma concentrations of baricitinib using a validated liquid chromatography tandem mass spectrometry method. Blood samples that will be used for other laboratory assessments (for example, chemistry, hematology) may be drawn at approximately the same time as the samples drawn to determine plasma concentrations of baricitinib, with the exception of Visit 2 (Week 0), which requires separate sampling from other labs. The timing will be as follows:

- At Visit 2 (Week 0), patients will take their investigational product in the clinic, and PK samples will be drawn approximately 15 minutes and 1 hour postdose. Samples for the other clinical laboratory assessments must be drawn prior to receiving the first dose.
- At Visit 4 (Week 4), patients will be asked to take their investigational product at home prior to visiting the clinic. The clinic visit should be scheduled so that the blood sample collected during this visit is drawn 2 to 4 hours after the dose is taken at home.

- At Visit 5 (Week 8), patients will be asked to take their investigational product at home prior to visiting the clinic. The clinic visit should be scheduled so that the blood sample collected during this visit is drawn 4 to 6 hours after the dose is taken at home.
- For Visit 6 (Week 12), and Visit 7 (Week 16), patients will take their investigational product in the clinic, and a blood sample will be collected at any time predose on the day of the clinic visits. If the patient has taken the oral dose prior to the visit, the sample may be drawn anytime postdose, and the inability to collect a predose sample will not be considered a protocol violation.
- For an early termination visit prior to Visit 7 (Week 16), a sample may be drawn anytime if the last dose of investigational product was taken within the last 48 hours. If the early termination is due to an AE, then a sample should be drawn if the visit occurred prior to Visit 16 (Week 52).
- In the event of an SAE, up to 2 additional blood samples may be taken at the investigator's discretion. If collected, the PK samples should be collected after the reported event, approximately 6 hours apart and within 24 hours of the patient's last dose.

For visits where PK samples will be collected, the actual date and 24-hour clock time of sample collection, and the date and 24-hour clock time of the 2 doses prior to the sample being drawn, should be recorded. For Visits 4 and 5, these 2 doses should be the dose taken the morning of the day of sample collection and the dose taken the day prior to sample collection. For Visits 6 and 7, these 2 doses should be the dose taken the day prior to sample collection and the dose taken 2 days prior to sample collection.

If the patient fails to follow the directions for a particular visit, the sample should still be collected at that visit, and the date and 24-hour clock time of sample collection and the date and 24-hour clock time of the 2 doses prior to the sample being drawn should be recorded.

Pharmacokinetic samples will be kept in storage at a laboratory facility designated by the sponsor. Pharmacokinetic samples may also be assayed for additional exploratory analyses. Pharmacokinetic results will not be provided to investigative sites until the completion of the study or to the blinded study team until the study has been unblinded.

Bioanalytical samples collected to measure investigational product concentration will be retained for a maximum of 1 year following last patient visit for the study.

9.6. Pharmacodynamics

Samples collected to measure pharmacodynamics markers including, but not limited to anti-dsDNA, complement, and IFN signature, will be identified by the patient number (coded) and retained at a facility selected by Lilly or its designee for a maximum of 1 year following last patient visit for the study.

9.7. Pharmacogenomics

9.7.1. Whole Blood Samples for Pharmacogenetic Research

A whole blood sample will be collected for pharmacogenetic analysis as specified in the Schedule of Activities (Section 2) where local regulations allow and patients provides consent. Patients choosing not to provide consent for whole blood samples to be collected for pharmacogenetic research will still be eligible to participate in the study.

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 SLE. 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 at a facility selected by Lilly or its designee 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 (ERBs)/ institutional review boards (IRBs) impose shorter time limits. 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 become(s) 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, PD, 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 deoxyribonucleic acid (DNA), ribonucleic acid (RNA), proteins, lipids, and other cellular elements.

Serum, plasma, whole blood RNA, and urine samples for 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 SLE, mechanism of action of baricitinib, and/or research method or in validating diagnostic tools or assay(s) related to SLE.

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 at a facility selected by Lilly or its designee for a maximum 15 years after the last patient visit for the study, or for a shorter period if local regulations and ERBs/IRBs impose shorter time limits. 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 become(s) commercially available. Any samples remaining after 15 years will be destroyed.

9.9. Medical Resource Utilization and Health Economics

Health Economics and Medical Resource Utilization parameters will not be evaluated in this study.

10. Statistical Considerations

10.1. Sample Size Determination

Study JAHZ will aim to enroll approximately 750 patients ≥18 years of age in a 1:1:1 ratio between baricitinib 4-mg, 2-mg, and placebo. The proposed samples size will ensure >90% power to detect an absolute difference of 17% between a baricitinib treatment group and the placebo treatment group assuming a 40% placebo response rate for the primary endpoint using a 2-sided α of 0.05. The anticipated effect size represents a clinically relevant difference.

Sample size and power estimates were obtained from nQuery® Advisor 7.0 using the χ^2 test of equal proportions.

10.2. Populations for Analyses

For purposes of analysis, the following populations are defined as outlined in Table JAHZ.5.

Table JAHZ.5. Analysis Populations

Population	Description
Modified Intent-to-Treat (mITT) Population	All patients who are randomized and receive at least 1 dose of investigational product will be included in the mITT population. The intention-to-treat principle remains preserved since the decision of whether or not to begin treatment could not be influenced by knowledge of the assigned treatment in this double-blind study. Patients will be analyzed according to the investigational product to which they were randomized. Unless otherwise specified, the efficacy and health outcomes analyses will be conducted on the mITT population.
Safety Population	The safety population is defined as 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. Patients will be analyzed according to the investigational product to which they were randomized.

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.

Efficacy and safety analyses will be conducted on the mITT and safety populations, respectively, as described in in Table JAHZ.5.

All tests of treatment effects will be conducted at a 2-sided α level of 0.05, unless otherwise stated.

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. Complete details of the planned analyses will be documented in the SAP.

10.3.1.1. Analysis Methods

Treatment comparisons of categorical efficacy and health outcomes variables will be made using a logistic regression analysis with Firth correction (Firth 1993; Heinze and Schemper 2001). Baseline disease activity (total SLEDAI-2K <10; ≥10), baseline corticosteroid dose (<10 mg/day; ≥10 mg/day prednisone or equivalent), region, and treatment group will be explanatory variables in the model. The p-value and 95% (unless otherwise specified) confidence interval (CI) for the odds ratio from the logistic regression model will be used for primary statistical inference. The difference in percentages and 95% CI of the difference in percentages using the Newcombe-Wilson method without continuity correction will be used for

descriptive purposes, unless otherwise specified. Missing data will be imputed using nonresponder imputation (NRI), as described in Section 10.3.1.2.

The primary analyses for all continuous efficacy and health outcome variables will use a restricted maximum likelihood-based mixed model for repeated measures (MMRM) analysis. The model will include treatment, baseline disease activity (total SLEDAI-2K <10; ≥10), baseline corticosteroid dose (<10 mg/day; ≥10 mg/day prednisone or equivalent), region, visit, treatment-by-visit interaction as fixed categorical effects, and baseline value and baseline value-by-visit interaction as fixed continuous effects to estimate change from baseline across postbaseline visits. The following covariance structures will be used to model the between- and within-patient errors: unstructured, heterogeneous Toeplitz, heterogeneous autoregressive, autoregressive (1), and heterogeneous compound symmetry. The variance-covariance structure that results in the smallest Akaike information criterion (AIC) will be used. The Kenward-Roger method will be used to estimate the denominator degrees of freedom. Type III tests for the least-squares (LS) means will be used for the statistical comparisons. The LS mean difference, standard error, p-value, and CIs will also be reported. Treatment group comparisons at specific study visits will be tested using the t-test obtained from the MMRM results. Additional details of the MMRM method are described in Section 10.3.1.2.

Treatment comparisons of continuous efficacy and health outcome variables may also be made using analysis of covariance (ANCOVA). This method may be used when no repeated measures are collected (e.g., SLICC/ACR), or as a sensitivity method to MMRM. Treatment, baseline disease activity (total SLEDAI-2K at screening <10; ≥10), baseline corticosteroid dose (<10 mg/day; ≥10 mg/day prednisone or equivalent), region, and baseline value will be included in the model. Type III tests for LS means will be used for statistical comparison between treatment groups. The LS mean difference, standard error, p-value, and 95% CI may also be reported. Any methods used to handle missing data will be specified in the SAP.

Time to first flare will be analyzed using a Cox proportional hazards model with treatment group, baseline disease activity (SLEDAI-2K <10; SLEDAI-2K ≥10), baseline corticosteroid dose (<10 mg/day; ≥10 mg/day prednisone or equivalent), and region fitted as explanatory variables. Hazard ratios, 95% CI and p-values will be presented for treatment comparisons to placebo. Patients who discontinue treatment early will be censored at the time of last dose of the study drug.

The Fisher exact test will be used for the AEs, discontinuations, 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 ANCOVA with treatment and baseline value in the model. Shift tables for select categorical safety analyses (for example, “high” or “low” laboratory results) will also be produced.

10.3.1.2. Missing Data Imputation

As with any clinical study, patient dropouts and consequently missing data are expected. Patients who discontinue treatment, unless they withdraw consent, will be encouraged to remain in the study and follow the scheduled visits until the primary analysis time point at Week 52 to

provide the efficacy and safety data needed. While every effort will be made to reduce missing data, the missing data imputation methods described below will be used to provide a conservative approach for assessing efficacy endpoints when patients have missing data.

The following imputation rules will be used:

- Nonresponder imputation (NRI): For analysis of categorical efficacy and health outcomes variables such as SRI-4, missing data will be imputed using an NRI method. Patients who discontinue treatment early, increase use of corticosteroids above the baseline dose other than the permitted burst, or have initiation of or an increase above baseline dose in immunosuppressant or antimalarial treatment any time after baseline will be defined as nonresponders.
- Mixed-model repeated measure (MMRM): For continuous variables, the primary analysis will be a MMRM analysis with a missing at random assumption for handling missing data. This analysis takes into account both missingness of data and the correlation of the repeated measurements. For patients who increase use of corticosteroids above the baseline dose other than the permitted burst, or have initiation of or an increase above baseline dose in immunosuppressant or antimalarial treatment any time after baseline, any observed data after that visit will be excluded from the MMRM analysis. Observed data after permanent discontinuation of investigational product will be analyzed as observed, with the exception of any data that occur after increase in use of corticosteroids above the baseline dose other than the permitted burst, or initiation of or an increase above baseline dose in immunosuppressant or antimalarial treatment.
- Tipping point analysis: Various "tipping point" analyses will be performed to evaluate the robustness of statistical analyses for different objectives/estimands to the assumptions about missing data mechanism. This will be implemented by multiple imputation with sensitivity parameters (deltas) quantifying the degree of departure from the missingness at random (MAR) assumption by gradually worsening imputed outcomes for patients treated with baricitinib and/or improving the imputed outcomes for patients treated with placebo. The tipping point is identified as the combination of delta values for baricitinib and placebo that leads to a loss of statistical significance when evaluating baricitinib relative to placebo. Additional details of the tipping point analysis will be provided in the SAP.

Sensitivity analyses that include all randomized patients, including patients who did not receive any dose of investigational product, will be performed on the primary and select secondary outcome measures, if appropriate. Additional sensitivity analyses, including additional methods of handling missing data or analyzing the data, that may be required to satisfy regulatory needs will be specified in the SAP.

10.3.1.3. Multiple Comparisons/Multiplicity

Multiplicity-controlled analyses will be performed on the primary and major secondary endpoints in order 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) will be used. The

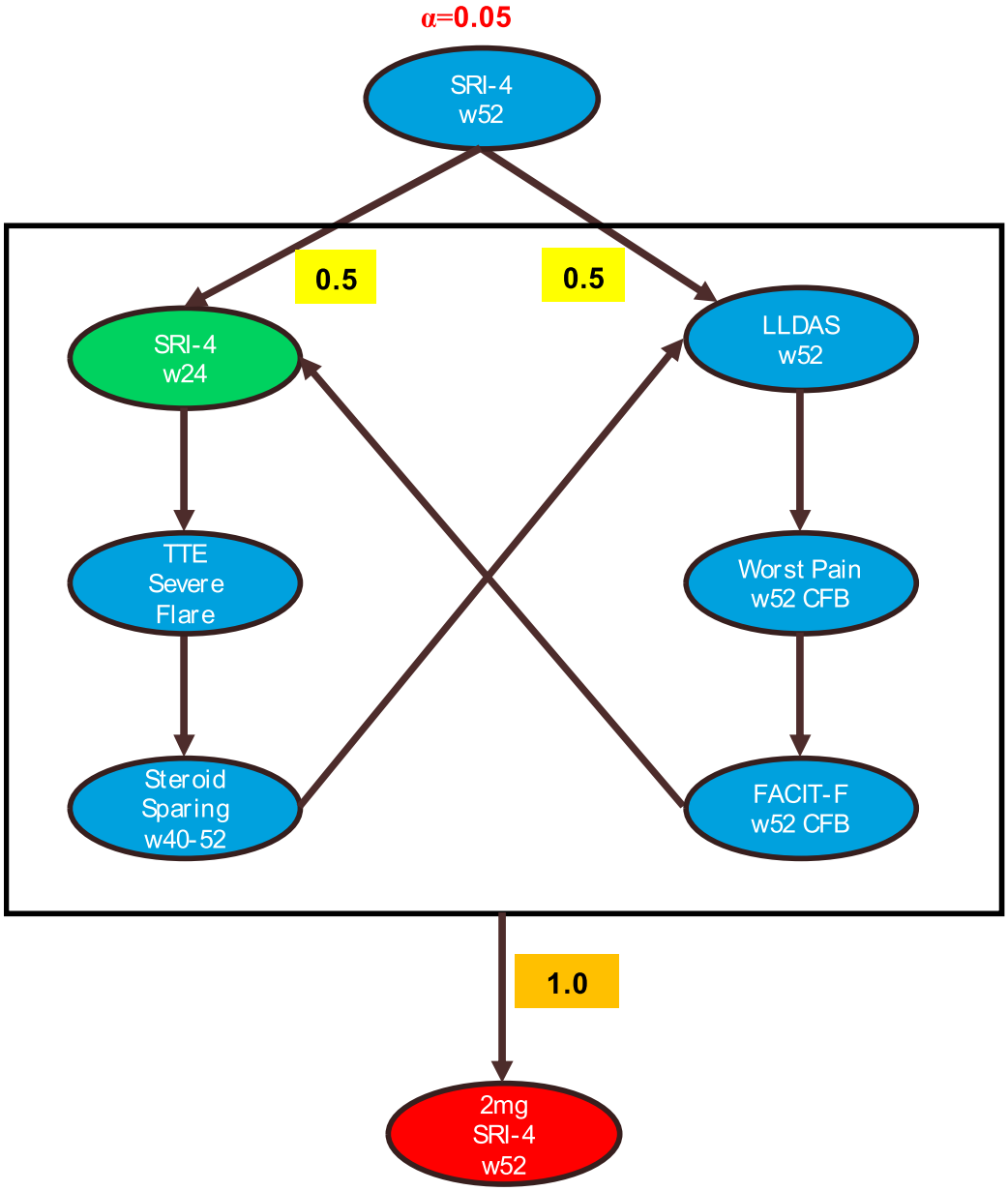
graphical approach is a closed testing procedure; hence, it strongly controls the familywise error rate across all endpoints (Alosh et al. 2014). Details of the graphical testing scheme (including testing order, interrelationships, Type I error allocation, and the associated propagation) are included below (see Figure JAHZ.2). The final graphical testing scheme will be prespecified in the SAP.

The following is a list of primary and major secondary objectives to be tested:

- Primary: Proportion of patients treated with baricitinib 4-mg QD compared to placebo achieving SRI-4 response at Week 52.
- Major Secondary:
    - Proportion of patients treated with baricitinib 4-mg QD compared to placebo achieving an SRI-4 response at Week 24.
    - Proportion of patients treated with baricitinib 4-mg QD compared to placebo receiving >7.5 mg prednisone (or equivalent) at baseline able to decrease dose by ≥25% to a prednisone equivalent dose of ≤7.5 mg/day maintained between Week 40 and Week 52.
    - Time to first severe flare (through Week 52) for patients treated with baricitinib 4-mg QD compared to placebo.
    - Proportion of patients treated with baricitinib 4-mg QD compared to placebo achieving an LLDAS response at Week 52.
    - Change from baseline in Worst Pain NRS at Week 52 for patients treated with baricitinib 4-mg QD compared to placebo
    - Change from baseline in FACIT-Fatigue at Week 52 for patients treated with baricitinib 4-mg QD compared to placebo
    - Proportion of patients treated with baricitinib 2-mg QD compared to placebo achieving an SRI-4 response at Week 52.
    - Proportion of patients treated with baricitinib 2-mg QD compared to placebo achieving an SRI-4 response at Week 24.
    - Proportion of patients treated with baricitinib 2-mg QD compared to placebo receiving >7.5 mg prednisone (or equivalent) at baseline able to decrease dose by ≥25% to a prednisone equivalent dose of ≤7.5 mg/day maintained between Week 40 and Week 52.
    - Time to first severe flare (through Week 52) for patients treated with baricitinib 2-mg QD compared to placebo.
    - Proportion of patients treated with baricitinib 2-mg QD compared to placebo achieving an LLDAS response at Week 52.

- Change from baseline in Worst Pain NRS at Week 52 for patients treated with baricitinib 2-mg QD compared to placebo
- Change from baseline in FACIT-Fatigue at Week 52 for patients treated with baricitinib 2-mg QD compared to placebo

The sequentially rejective, weighted Bonferroni multiple testing procedure is completely specified by the following graph (Figure JAHZ.2).



Abbreviations: CFB = change from baseline; LLDAS = Lupus Low Disease Activity State; SRI-4 = System Lupus Erythematosus Responder Index-4; TTE = time to event; W = week

Note: All tests are 4-mg compared to placebo unless otherwise specified. Testing for baricitinib 2-mg (compared to placebo) will occur in the same order as those prespecified for the 4-mg dose if all previous objectives were met.

Figure JAHZ.2. Illustration of graphical multiple testing procedure with initial α allocation and weights.

The primary objective will first be tested at α = 0.05. If the null hypothesis is not rejected, no further testing is conducted, as the α for that test is considered "spent" and cannot be passed to other endpoints. If the null hypothesis is rejected, all α will be propagated to the next objective(s). The testing process continues, with α propagated according to the weights on the corresponding edges displayed in Figure JAHZ.2, as long as there is at least one hypothesis that can be rejected at its allocated α level. Each time a hypothesis is rejected, the graph is 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 is repeated until all hypotheses have been tested or when no remaining hypotheses can be rejected at their corresponding α levels.

10.3.2. Treatment Group Comparability

10.3.2.1. Patient Disposition

A detailed description of patient disposition by treatment will be summarized with reasons for discontinuation. Frequency counts and percentages will be presented for each treatment group. All patients who discontinue from the study will be identified, and the extent of their participation in the study will be reported along with their reason for discontinuation.

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. A complete list of patient characteristics and baseline clinical measures will be provided in the SAP.

10.3.2.3. Concomitant Therapy

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

10.3.2.4. Treatment Compliance

Compliance with investigational product treatment for baricitinib and placebo will be assessed through 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, 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.

10.3.3. Efficacy Analyses

10.3.3.1. Primary Analyses

The primary efficacy measure is SRI-4 response (defined in Section 9.1.1). A logistic regression model with Firth correction (as described in Section 10.3.1.1) will be used to test the treatment difference between baricitinib 4-mg and placebo in the proportion of patients achieving SRI-4

response at Week 52. Missing data will be imputed using the NRI method described in Section 10.3.1.2.

10.3.3.2. Major Secondary Analyses

The major secondary comparisons will be based on the multiple testing procedure detailed in Section 10.3.1.3.

Treatment comparisons in the proportion of patients achieving a binary response will be analyzed using logistic regression with Firth correction. Treatment comparisons in the continuous measures will be analyzed using MMRM. Treatment comparisons for time to first flare will be analyzed using a Cox proportional hazards model. These methods are described in Section 10.3.1.1. Missing data for the binary outcome measures will be imputed using the NRI method described in Section 10.3.1.2.

10.3.3.3. Other Secondary and Exploratory Analyses

There will be no adjustment for multiple comparisons. Analyses will be conducted for the other secondary and exploratory objectives, which include health outcomes, defined in Section 4. The analysis methods will generally follow those specified in Section 10.3.1.1. Specific details of analyses will be specified in the SAP.

10.3.4. Safety Analyses

All safety data will be summarized by treatment group and analyzed using the safety population. Safety assessments will include AEs, laboratory analytes, vital signs, and questionnaires to assess the existence and severity of depression. Analyses will focus on comparisons of baricitinib to placebo, using the safety analysis methods described in Section 10.3.1.1.

10.3.4.1. Adverse Events

Adverse events are classified based upon the Medical Dictionary for Regulatory Activities (MedDRA). Treatment-emergent adverse events are defined as AEs that first occurred or worsened in severity on or after the date of the first dose of investigational product. The number of TEAEs as well as the number and percentage of patients who experienced at least 1 TEAE will be summarized using MedDRA for each system organ class (or a body system) and each preferred term by treatment group. For events that are gender-specific, the denominator and computation of the percentage will only include patients from the given gender.

Serious adverse events (including deaths), treatment-emergent AESIs, and AEs that lead to investigational product discontinuation or interruption will also be summarized using MedDRA for each system organ class and each preferred term by treatment group. Potential AESIs will be identified by a standardized MedDRA query or a Lilly-defined MedDRA query. Details of the AESIs (listed in Section 9.2.2) and other events of interest and analysis methods will be documented in the SAP or program safety analysis plan.

10.3.4.2. Clinical Laboratory Tests

All clinical laboratory results will be descriptively summarized by treatment group. Individual results that are outside the 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. Categorical variables, including the incidence of abnormal values and incidence of AESIs, will be summarized by frequency and percentage of patients in corresponding categories. Shift tables will be presented for selected measures.

10.3.4.3. Vital Signs, Physical Findings, and Other Safety Evaluation

Observed values and changes from baseline (predose or screening if missing) for vital signs and physical characteristics, and other continuous safety measures will be descriptively summarized by treatment group and time point. Shift tables will be presented where appropriate.

Summary tables or listings for the C-SSRS and the Self-Harm Supplement Form will be produced as needed.

The incidence, duration, and reasons of investigational product interruptions will be summarized and compared descriptively among treatment groups.

Further analyses may be performed and will be planned in the SAP.

10.3.5. Pharmacokinetic/Pharmacodynamic Analyses

All plasma baricitinib concentration-time data will be pooled and evaluated using population PK methods. A covariate screen of patient and study-specific factors will be included in the analyses based on factors investigated in previous and (if any) ongoing PK analyses, and on their relevance to the target population. Exploratory and/or model-based analyses examining the relationships between baricitinib exposure and efficacy and response endpoints will be conducted. Other analyses of efficacy and safety outcome measures may also be assessed as scientifically appropriate and warranted by available data. Details about the analyses to be conducted will be contained in the PK/PD analysis plan.

.

10.3.6. Interim Analyses

A data monitoring committee (DMC) will oversee the conduct of all the Phase 3 clinical trials evaluating baricitinib in patients with SLE. The DMC will consist of members external to Lilly. This DMC will follow the rules defined in the DMC charter, focusing on potential and identified risks for this molecule and for this class of compounds. Data monitoring committee membership will include, at a minimum, specialists with expertise in rheumatology, statistics, and other appropriate specialties. This DMC for studies of patients with SLE may be coordinated with the DMC(s) for other ongoing studies of baricitinib in other indications. Details of the DMC will be documented in a DMC charter.

Access to the unblinded data will be limited to the DMC and statisticians providing the data. These statisticians will be independent from the study team. The study team will not have access to the unblinded data. Only the DMC is authorized to evaluate unblinded interim efficacy and safety analyses. The study sites will receive information about interim results ONLY if they need to know for the safety of their patients. 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. However, the study will not be stopped for positive efficacy results, and there is no planned futility assessment. Hence, there will be no α adjustment for these interim analyses. Adjustments to type I error will be made if an unplanned futility analysis for efficacy is performed.

In addition to the DMC members, a limited number of pre-identified individuals may gain access to the limited unblinded data, as specified in the unblinding plan, prior to the final database lock, in order to initiate the population PK/PD model development processes. These analyses will not be considered interim analyses. 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.

11. References

Alarcón GS, McGwin G Jr, Bartolucci AA, Roseman J, Lisse J, Fessler BJ, Bastian HM, Friedman AW, Reveille JD. Systemic lupus erythematosus in three ethnic groups. IX. Differences in damage accrual. *Arthritis Rheum*. 2001;44(12):2797-2806.

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

Aviña-Zubieta JA, Alarcón GS, Bischoff-Ferrari HA, Fischer-Betz R, Schneider M, Gall V, Liang MH, Schur P, Illei G, Mikdashi J, Petri M, Phillips C, Pouchot J, St.Clair W. Measurement of fatigue in systemic lupus erythematosus: a systematic review. *Arthritis Rheum*. 2007;57(8):1348-1357.

Brazier JE, Harper R, Jones NM, O'Cathain A, Thomas KJ, Usherwood T, Westlake L. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *BMJ*. 1992;305(6846):160-164.

Bretz F, Maurer W, Brannath W, Posch M. A graphical approach to sequentially rejective multiple test procedures. *Stat Med*. 2009;28(4):586-604.

Bruce IN, O'Keeffe AG, Farewell V, Hanly JG, Manzi S, Su L, Gladman DD, Bae SC, Sanchez-Guerrero J, Romero-Diaz J, Gordon C, Wallace DJ, Clarke AE, Bernatsky S, Ginzler EM, Isenberg DA, Rahman A, Merrill JT, Alarcón GS, Fessler BJ, Fortin PR, Petri M, Steinsson K, Dooley MA, Khamashta MA, Ramsey-Goldman R, Zoma AA, Sturfelt GK, Nived O, Aranow C, Mackay M, Ramos-Casals M, van Vollenhoven RF, Kalunian KC, Ruiz-Irastorza G, Lim S, Kamen DL, Peschken CA, Inanc M, Urowitz MB. Factors associated with damage accrual in patients with systemic lupus erythematosus: results from the Systemic Lupus International Collaborating Clinics (SLICC) Inception Cohort. *Ann Rheum Dis.* 2015;74(9):1706-1713.

Cella D, Webster K. Linking outcomes management to quality-of-life measurement. *Oncology (Williston Park)*. 1997;11(11A):232-235.

Cooper GS, Treadwell EL, St.Clair EW, Gilkeson GS, Dooley MA. Sociodemographic associations with early disease damage in patients with systemic lupus erythematosus. *Arthritis Rheum*. 2007;57(6):993-999.

Danza and Ruiz-Irastorza. Infection risk in systemic lupus erythematosus patients: susceptibility factors and preventative strategies. *Lupus*. 2013;22(12):1286-1294.

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.

[EMA] European Medicines Agency. ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials. Step 2b; 2017. EMA/CHMP/ICH/436221/2017.

[EMA] European Medicines Agency. Guidance on clinical investigation of medicinal products for the treatment of systemic lupus erythematosus and lupus nephritis; 2015.

The EuroQol Group. EQ-5D-5L User Guide. Version 2.1. April 2015. Available at: https://euroqol.org/wp-content/uploads/2016/09/EQ-5D-5L_UserGuide_2015.pdf. Accessed April 10 2018.

[FDA] Food and Drug Administration. Guidance for Industry Systemic Lupus Erythematosus – Developing Medical Products for Treatment; 2010.

[FDA] Food and Drug Administration. Benlysta Prescribing Information, 2012. https://www.accessdata.fda.gov/drugsatfda_docs/label/2012/125370s016lbl.pdf. Accessed April 10, 2018.

Firth D. Bias reduction of maximum likelihood estimates. *Biometrika*. 1993;80(1):27-38.

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*. 2016. Doi: 10.1002/art.39953.

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.

Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. *Clin Chem*. 1972 Jun;18(6):499-502.

Furie R, Khamashta M, Merrill JT, Werth VP, Kalunian K, Brohawn P, Illei GG, Drappa J, Wang L, Yoo S, CD1013 Study Investigators. Anifrolumab, an anti-interferon-α receptor monoclonal antibody, in moderate-to-severe systemic lupus erythematosus. *Arthritis Rheumatol*. 2017;69(2):376-86.

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.

Guttman-Yassky E, Silverberg JI, Nemoto O, Forman SB, Wilke A, Prescilla R, de la Pena A, Nunes FP, Janes J, Gamalo M, Donley D, Paik J, DeLozier AM, Nickoloff BJ, Simpson EL. Baricitinib in adult patients with moderate-to-severe atopic dermatitis: A phase 2 parallel, double-blind, randomized, placebo-controlled, multiple-dose study. *J Am Acad Dermatol*. 2018; [epub ahead of print]. Available at: http://dx.doi.org/10.1016/j.jaad.2018.01.018

Heinze G, Schemper M. A solution to the problem of monotone likelihood in Cox regression. *Biometrics*. 2001;57(1):114-119.

Hochberg MC, for the Diagnostic and Therapeutic Criteria Committee of the American College of Rheumatology. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus [letter]. *Arthritis Rheum*. 1997;40:1725.

Hoffman RW, Merrill JT, Alarcón-Riquelme MM, Petri M, Dow ER, Nantz E, Nisenbaum LK, Schroeder KM, Komocsar WJ, Perumal NB, Linnik MD, Airey DC, Liu Y, Rocha GV, Higgs

RE. Gene expression and pharmacodynamic changes in 1,760 systemic lupus erythematosus patients from two phase III trials of BAFF blockade with tabalumab. *Arthritis Rheumatol*. 2017;69(3):643-654.

Illei GG, Shirota Y, Yarboro CH, Daruwalla J, Tackey E, Takada K, Fleisher T, Balow JE, Lipsky PE. Data on safety, preliminary efficacy, and impact on circulating plasma cells from an open-label phase I dosage-escalation study. *Arthritis Rheum*. 2010;62(2):542-552.

Kaslow RA. High rate of death caused by systemic lupus erythematosus among U.S. residents of Asian descent. *Arthritis Rheum*. 1982;25(4):414-418.

Lim SS, Bayakly AR, Helmick CG, Gordon C, Easley KA, Drenkard C. The incidence and prevalence of systemic lupus erythematosus, 2002-2004: The Georgia Lupus Registry. *Arthritis Rheum*. 2014;66(2):357-368.

Linker-Israeli M, Wallace DJ, Prehn J, Michael D, Honda M, Taylor KD, Paul-Labrador M, Fischel-Ghodsian N, Fraser PA, Klinenberg JR. Association of IL-6 gene alleles with systemic lupus erythematosus (SLE) and with elevated IL-6 expression. *Genes and Immunity*. 1999;1:45-52.

Ozel F, Argon G. The effects of fatigue and pain on daily activities in systemic lupus erythematosus. *Agri.* 2015;27(4):181-189.

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.

GA, Reinhardt A, Ramsey S, Wittkowski H, Hashkes PJ, Berkun Y, Schalm S, Murias Montealegre Sanchez S, Dare JA, Brown D, Stone DL, Gao L, Klausmeier T, Foell D, de Jesus AA, Chapelle DC, Kim H, Dill S, Colbert R, Failla L, Kost B, O'Brien M, Reynolds JC, Folio LR, Calvo KR, Paul KM, Weir N, Brofferio A, Soldatos A, Biancotto A, Cowen EW, Digiovanna JG, Gadina M, Lipton AJ, Hadigan C, Holland SM, Fontana J, Alawad AS, Brown RJ, Rother KI, Heller T, Brooks KM, Kumar P, Brooks SR, Waldman M, Singh HK, Nickeleit V, Silk M, Prakash A, Janes JM, Ozen S, Wakim PG, Brogan PA, Macias WL, and Goldbach-Mansky R. JAK1/2 inhibition with baricitinib in the treatment of autoinflammatory interferonopathies. J Clin Invest 2018; Epub ahead of print.

Somers EC, Marder W, Cagnoli P, Lewis EE, DeGuire P, Gordon C, Helmick CG, Wang L, Wing JJ, Dhar P, Leisen J, Shaltis D, McCune WJ. Population-based incidence and prevalence of systemic lupus erythematosus. *Arthritis Rheum*. 2014;66(2):369-378.

Tan EM, Cohen AS, Fries JF, Masi AT, McShane DJ, Rothfield NF, Schaller JG, Talal N, Winchester RJ. The 1982 revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum.* 1982;25(11):1271-1277.

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 C, Rooney T, Schlichting D, Macias W, de Bono S, Tanaka Y. Baricitinib versus placebo or adalimumab in rheumatoid arthritis.

Tuttle KR, Brosius FC III, 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 [abstract]. In: American Diabetes Association 75th Scientific Sessions; 2015 June 5-9; Boston (MA). *Diabetes.* 2015;64(Suppl 1A). Abstract 114-LB.

UpToDate resource page. Epidemiology and pathogenesis of systemic lupus erythematosus. 2018. Available at: https://www.uptodate.com/contents/epidemiology-and-pathogenesis-of-systemic-lupus-erythematosus?source=search_result&search=epidemiology+and+pathogensis+of+systemic+lupus&selectedTitle=1%7E150#H1. Accessed April 10, 2018.

Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36), I: Conceptual framework and item selection. *Med Care*. 1992;30(6):473-483.

Yu T, Enioutina EY, Brunner HI, Vinks AA, Sherwin CM. Clinical pharmacokinetics and pharmacodynamics of biologic therapeutics for treatment of systemic lupus erythematosus. *Clin Pharmacokinet*. 2017;56(2):107-125.

Yurkovich M, Vostretsova K, Chen W, Avina-Zubieta JA. Overall and cause-specific mortality in patients with systemic lupus erythematosus: a meta-analysis of observational studies. *Arthritis Care Res*. 2014;66(4):608-616.

12. Appendices

Appendix 1. Abbreviations and Definitions

Term	Definition
ACR	American College of Rheumatology
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.
AESI	Adverse events of special interest
AGS	Aicardi-Goutières Syndrome
ALP	alkaline phosphatase
ALT	alanine aminotransferase
ANA	antinuclear antibody
ANC	absolute neutrophil count
ANCOVA	analysis of covariance
AST	aspartate aminotransferase
BILAG	British Isles Lupus Assessment Group
blinding/masking	A double-blind study is one in which neither the [patient/subject] 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.
CANDLE	Chronic atypical neutrophilic dermatosis with lipodystrophy and elevated temperature
BPI	Brief Pain Inventory
CFR	Code of Federal Regulations
CI	confidence interval
CLASI	Cutaneous Lupus Erythematosus Disease Area and Severity Index
CNS	central nervous system
CIOMS	Council for International Organizations of Medical Sciences
Companion diagnostic	An in vitro diagnostic device (assay or test) that provides information that is essential for the safe and effective use of a corresponding therapeutic product

complaint	A complaint is any written, electronic, or oral communication that alleges deficiencies related to the identity, quality, purity, durability, reliability, safety or effectiveness, or performance of a drug or drug delivery system.
CRP	clinical research physician: Individual responsible for the medical conduct of the study. Responsibilities of the CRP may be performed by a physician, clinical research scientist, global safety physician or other medical officer.
CSR	clinical study report
C-SSRS	Columbia-Suicide Severity Rating Scale
DMC	data monitoring committee
DNA	deoxyribonucleic acid
DSM-V	Diagnostic and Statistical Manual of Mental Disorders
DVT	deep vein thrombosis
ECG	electrocardiogram
eCRF	electronic case report form
eGFR	estimated glomerular filtration rate
eCOA	electronic clinical outcome assessment
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
ERB	ethics review board
FACIT	Functional Assessment of Chronic Illness Therapy-Fatigue
GCP	good clinical practice
GPS	Global Patient Safety
HBcAb	hepatitis B core antibody
HBsAb	hepatitis B surface antibody
HBsAg	hepatitis B surface antigen
HBV	hepatitis B virus
HCV	hepatitis C virus

IB	Investigator's Brochure
ICF	informed consent form
ICH	International Council for Harmonisation
IFN	interferon
Ig	immunoglobulin
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.
IL	interleukin
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.
IRB	institutional review board
ITT	intention to treat: The principle that asserts that the effect of a treatment policy can be best assessed by evaluating on the basis of the intention to treat a patient (that is, the planned treatment regimen) rather than the actual treatment given. It has the consequence that patients allocated to a treatment group should be followed up, assessed, and analyzed as members of that group irrespective of their compliance to the planned course of treatment.
IUD	intrauterine device
IVRS/IWRS	interactive voice-response system/interactive web-response system
JAK	Janus kinase
JDM	juvenile dermatomyositis
LLDAS	Lupus low disease activity state
LS	least squares
LTE	long-term extension
MACE	major adverse cardiovascular events
MCS	mental component score
MDRD	Modification of Diet in Renal Disease
MedDRA	Medical Dictionary for Regulatory Activities

MI	myocardial infarction
mITT	modified intent-to-treat population
MMRM	mixed model for repeated measures
MTX	methotrexate
NRI	nonresponder imputation
NRS	Numeric Rating Scale
NSAID	nonsteroidal anti-inflammatory drug
PCS	physical component score
PE	pulmonary embolism
PK/PD	pharmacokinetics/pharmacodynamics
PPD	purified protein derivative
PPS	per-protocol set: The set of data generated by the subset of patients who sufficiently complied with the protocol to ensure that these data would be likely to exhibit the effects of treatment, according to the underlying scientific model.
PRO	patient-reported outcomes
PSAP	program safety analysis plan
QD	once daily
QIDS-SR16	Quick Inventory of Depressive Symptomatology
QTc	corrected QT interval
RA	rheumatoid arthritis
RNA	ribonucleic acid
SAE	serious adverse event
SAP	statistical analysis plan
SAVI	Stimulator of Interferon Genes (STING)-Associated Vasculopathy With Onset During Infancy
screen	The act of determining if an individual meets minimum requirements to become part of a pool of potential candidates for participation in a clinical study.
SF-36v2	Short-Form 36-item Health Survey version 2
SFI	SELENA-SLEDAI Flare Index

SLE	systemic lupus erythematosus
SLEDAI-2K	Systemic Lupus Erythematosus Disease Activity Index 2000
SLICC	Systemic Lupus Erythematosus International Collaborating Clinics
SoC	Standard of care
SRI-4	Systemic Lupus Erythematosus Responder Index-4
STAT	signal transducers and activators of transcription
SUSARs	suspected unexpected serious adverse reactions
TB	tuberculosis
TBL	total bilirubin level
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 and does not necessarily have to have a causal relationship with this treatment.
TSH	thyroid-stimulating hormone
TYK	tyrosine kinase
ULN	upper limit of normal
VAS	Visual Analog Scale
VTE	venous thromboembolic event
WPAI-Lupus	Work Productivity and Activity Impairment-Lupus

Appendix 2. Clinical Laboratory Tests

Clinical Laboratory Tests[a]

Hematology[b]
Hemoglobin
Hematocrit
Erythrocyte count (RBC)
Absolute reticulocyte count
Mean cell volume (MCV)
Mean platelet volume
Mean cell hemoglobin (MCH)
Mean cell hemoglobin concentration (MCHC)
Leukocytes (WBC)
Absolute count of
- Neutrophils, segmented
- Neutrophils, juvenile (bands)
- Lymphocytes
- Monocytes
- Eosinophils
- Basophils
- Platelets

Cell morphology
Differential and blood smear[e]

Lymphocyte subsets (T, B, NK, and T cell subsets)

Urinalysis[b,n,]
Color
Specific gravity
pH
Urine Protein
Glucose
Ketones
Blood
Bilirubin
Urobilinogen
Leukocyte esterase
Nitrite
Microscopic examination of sediment[j]
Urine creatinine
Urine protein
Urine protein-creatinine ratio

Clinical Chemistry[b]
Serum Concentrations of:
Sodium
Potassium
Total bilirubin
Direct bilirubin
Alkaline phosphatase (ALP)
Alanine aminotransaminase (ALT)/serum glutamic pyruvic transaminase
Aspartate aminotransferase (AST)/serum glutamic oxaloacetic transaminase
Blood urea nitrogen (BUN)
Creatinine
Estimated glomerular filtration rate (eGFR)[c]
Uric acid
Calcium
Glucose[d]
Albumin
Creatine phosphokinase (CPK)
Total protein
Iron studies (iron, TIBC, and ferritin)

Lipid Profile[d,f]
Total cholesterol
High-density lipoprotein cholesterol (HDL-C)
Low-density lipoprotein cholesterol (LDL-C)
Triglycerides

Other Tests
Serum pregnancy test (females only)[g]
Urine pregnancy test (females only)[g]
Follicle-stimulating hormone (FSH) (females only)[h]
Hepatitis B surface antigen (HBsAg)
Hepatitis B surface antibody (HBsAb)
Hepatitis B core antibody (HBcAb), total
Hepatitis C antibody[i]
HBV DNA[k]
HCV RNA[i]
Human immunodeficiency virus (HIV) serology
PPD or QuantiFERON®-TB Gold, or T-SPOT®.TB[l]
Thyroid-stimulating hormone (TSH)

Autoantibodies	Serum immunoglobulins (IgG, IgA, IgM)
Anti-nuclear antibody (ANA)	Stored serum, plasma, urine, mRNA[m], and DNA samples for exploratory biomarker analyses
Anti-double-stranded DNA (anti-dsDNA) antibody	Baricitinib plasma concentration (PK sample)
Anti-Smith	Lupus anticoagulant (dRVVT)
Autoantibodies (anti-RNP, anti-Sm, anti-SSA/Ro, anti-SSB/La)	Complement C3 and C4
Anti-cardiolipin antibodies IgG, IgA, IgM	Anti-beta2 glycoprotein-I (IgG, IgM)

Abbreviations: dRVVT = Dilute Russell's Viper Venom Time; dsDNA = double-stranded deoxyribonucleic acid; HBV DNA = hepatitis B virus deoxyribonucleic acid; HCV RNA = hepatitis C virus ribonucleic acid; Ig = immunoglobulin; mRNA = messenger ribonucleic acid; PK = pharmacokinetics; PPD = purified protein derivative; RBC = red blood cells; TIBC = total iron binding capacity; WBC = white blood cells.

[a] All labs will be assayed/calculated by a Lilly-designated laboratory unless otherwise noted.

[b] Unscheduled blood chemistry hematology, and urinalysis panels may be performed at the discretion of the investigator.

[c] eGFR for serum creatinine will be calculated by the central laboratory using the Modification of Diet in Renal Disease (MDRD) isotope dilution mass spectrometry traceable method.

[d] Fasting laboratory values for glucose and lipids will be required at baseline, Weeks 12, 24, 36, and 52. 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.

[e] Differential and blood smear may be performed if necessary.

[f] Lipid panel consists of direct HDL-C, triglycerides, cholesterol, and LDL-C (calculation from Friedewald et al. 1972).

[g] For all women of child-bearing potential, a serum pregnancy test (central laboratory) will be performed at Visit 1. Urine pregnancy tests (local laboratory) will also be performed at each subsequent study visit per the Schedule of Activities.

[h] For female patients ≥50 and <55 years of age who have had a cessation of menses for at least 6 and less than 12 months, an FSH test will be performed to confirm non-child-bearing potential (FSH ≥40 mIU/mL).

[i] A confirmatory test of HCV RNA will be performed if the patient is positive for HCV antibody.

[j] Microscopic examination of sediment will be performed if abnormalities are noted on the routine urinalysis.

[k] HBV DNA testing will be performed in patients who tested positive for HBcAb at screening.

[l] 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 must be performed centrally; the T-SPOT® must be performed locally. See Exclusion Criterion [29] for details.

[m] Interferon signature may be assayed from stored mRNA sample.

[n] Urine should be collected as a clean catch sample.

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 that the patient understands the nature of the study, the potential risks and benefits of participating in the study, and that their participation is voluntary.
- ensuring that informed consent is given by each patient or legal representative. 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 study.
- ensuring that a copy of the ICF is provided to the participant or the participant's legal representative and is kept on file.
- ensuring that the medical record includes a statement that written informed consent was obtained before the participant was enrolled in the study and the date the written consent was obtained. The authorized person obtaining the informed consent must also sign the ICF.

Appendix 3.1.2. Recruitment

Lilly or its designee is responsible for the central recruitment strategy for patients. Individual investigators may have additional local requirements or processes.

Appendix 3.1.3. Ethical Review

The investigator or an appropriate local representative 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 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 protocol and related amendments and addenda, current Investigator Brochure (IB) and updates during the course of the study
- informed consent form
- Other relevant documents (for example, curricula vitae, advertisements)

Appendix 3.1.4. Regulatory Considerations

This study will be conducted in accordance with the protocol and with the following:

- 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 3.1.5. Investigator Information

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

Physicians with other specialties and experience in treatment of patients with systemic lupus erythematosus (SLE) may also participate as investigators.

Appendix 3.1.6. 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 3.1.7. Final Report Signature

The clinical study report (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
- 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 (e.g., 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 data record will serve as source data will documented by the site in their study file.

Case report form data collected by a third-party will be encoded by the third-party and stored electronically in the third-party's database system. Validated data will subsequently be transferred to the sponsor's data warehouse, using standard Lilly file transfer processes.

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 or its designee, 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 or its designee judges it necessary for medical, safety, regulatory, or other reasons consistent with applicable laws, regulations, and GCP.

The publication policy for Study I4V-MC-JAHZ is described in the Clinical Trial Agreement.

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, or its designee, 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 cells; WBC = white blood cells.

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

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

Appendix 5. 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 Lilly designated medical monitor. 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 6. Protocol Amendment I4V-MC-JAHZ(a) Summary - A Randomized, Double-Blind, Placebo-Controlled, Parallel-Group, Phase 3 Study of Baricitinib in Patients with Systemic Lupus Erythematosus

Overview

Protocol I4V-MC-JAHZ, A Randomized, Double-Blind, Placebo-Controlled, Parallel-Group, Phase 3 Study of Baricitinib in Patients with Systemic Lupus Erythematosus, has been amended. The new protocol is indicated by amendment (a) and will be used to conduct the study in place of any preceding version of the protocol.

The overall changes and rationale for the changes made to this protocol are described in the following table: Editorial revisions with no impact on protocol design or implementation were also made. These revisions are not noted in this protocol amendment summary except where contained in a section with substantive changes.

Amendment Summary for Protocol I4V-MC-JAHZ Amendment (a)

Section # and Name	Description of Change	Brief Rationale
1. Synopsis 10.3.1.1 Analysis Methods 10.3.3.1 Primary Analyses 10.3.3.2 Major Secondary Analyses 11. References	Modified logistic regression analyses to be logistic regression with the Firth correction. Removed text that referred to changing to Fisher exact test when logistic regression sample size requirements are not met.	Logistic regression with Firth correction is a penalized likelihood method that can handle both few and many events. Changing to this method allowed deleting the condition of checking the number of responders in categories before deciding which methodology to use.
2. Schedule of Activities 6.1 Inclusion Criteria Appendix 2. Clinical Laboratory Tests	Changed the text in Inclusion Criterion [9]d, footnote “v” of the Schedule of Activities, and footnote “h” of Appendix 2.	Clarified the definition of post-menopausal.
2. Schedule of Activities 6.2 Exclusion Criteria [29] 9.4.7 Chest x-Ray and Tuberculosis Testing	Chest images from the prior 6 months are not limited solely to x-ray images, but other imaging modalities (other than ultrasound) are allowed.	Data from types of chest imaging other than x-ray can be accepted for tuberculosis screening (e.g. CT, MRI).
9.4.10 Safety Monitoring	Added arterial thromboembolic event as an event that will be adjudicated.	ATEs will be adjudicated by a blinded Clinical Event Committee.
10.3.1.1 Analysis Methods	Added text stating that the variance-covariance structure that results in the smallest Akaike information criterion (AIC) will be used.	Due to the large number of visits in the study, using unstructured covariance will result in fitting *many* parameters. This revised strategy will potentially allow a more parsimonious model to be fit.

10.3.1.2 Missing Data Imputation	Removed text stating that no additional imputation methods will be applied to the mixed-model repeated measure (MMRM) analysis. Language regarding tipping point analysis was edited.	Deleted text to allow flexibility for other kinds of imputation, if needed. The language regarding tipping point analysis was not consistent with ICH E9 guidance regarding sensitivity vs. supplemental analysis (EMA/CHMP/ICH/436221/2017). The language was changed to allow correct application of ICH-E9 principles in the SAP.
10.3.6 Other Analyses 10.3.6.1 Subgroup Analyses	The section on subgroup analysis has been removed from the protocol.	Subgroups and their analyses will be specified within the integrated efficacy analysis plan. Subgroups are more appropriately assessed as part of an integrated plan than in the individual studies, because of increased sample size in the integrated plan compared to individual studies.

Revised Protocol Sections

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

1. Synopsis

Statistical Analysis:

Treatment comparisons of categorical efficacy and health outcomes variables will be made using a logistic regression analysis with Firth correction (Firth 1993; Heinze and Schemper 2001). ~~b~~Baseline disease activity, baseline corticosteroid dose, region, and treatment group will be explanatory variables in the model. The p-value and 95% confidence interval (CI) for the odds ratio from the logistic regression model will be used for primary statistical inference. Patients who discontinue treatment early, increased use of corticosteroids above the baseline dose other than the permitted burst, or have initiation of or an increase above the baseline dose in immunosuppressant or antimalarial treatment any time after baseline will be defined as nonresponders and the data will be imputed using the nonresponder imputation (NRI) method.

Treatment comparisons of continuous efficacy and health outcomes variables will be made using a restricted maximum likelihood-based mixed model for repeated measures (MMRM) analysis. The model will include treatment, baseline disease activity, baseline corticosteroid dose, region, visit, treatment-by-visit interaction as fixed categorical effects, and baseline value and baseline value-by-visit interaction as fixed continuous effects to estimate change from baseline across postbaseline visits. ~~An unstructured covariance structure will be used to model the between- and within-patient errors.~~ Type III tests for the least squares (LS) means will be used for the statistical comparisons.

2. Schedule of Activities

Table JAHZ.1. Schedule of Activities

	Screening	Double-Blinded Treatment																Post-Treatment Follow-up
Visit Number	**V1**	**V2[a]**	**V3**	**V4**	**V5**	**V6**	**V7**	**V8**	**V9**	**V10**	**V11**	**V12**	**V13**	**V14**	**V15**	**V16[b]**	**ET[c]**	**V801[d]**
Study Week		**0**	**1**	**4**	**8**	**12**	**16**	**20**	**24**	**28**	**32**	**36**	**40**	**44**	**48**	**52**	**any**	**56 or last dose + 4 weeks**
Study Day	**-42 to -3**	**1**	**8 ± 3**	**29 ± 4**	**57 ± 4**	**85 ± 4**	**113 ± 4**	**141 ± 4**	**169 ± 4**	**197 ± 4**	**225 ± 4**	**253 ± 4**	**281 ± 4**	**309 ± 4**	**337 ± 4**	**365 ± 4**	**any**	**last dose + 28 ± 5 days**
Procedures and Assessments																		
Administer TB test[h]	X																	
Chest x-ray[j]	X																	
Laboratory Tests																		
FSH[v]	X																	
Urinalysis[bb]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

Abbreviations: ET = early termination; FSH = follicle-stimulating hormone; V = visit.

a At Visit 2, all baseline assessments must be performed and baseline laboratory samples must be drawn PRIOR to administration of the first dose of investigational product for randomized patients.

b Patients who complete the treatment period through Visit 16 may be eligible to participate in the long-term extension study (I4V-MC-JAIM [JAIM]). For patients enrolled in JAIM, Visit 1 of Study JAIM should occur on the same day as Visit 16 of Study I4V-MC-JAIA (JAIA).

c Those patients who discontinue from the study prior to Visit 16 should complete the ET visit and proceed to post-treatment follow-up.

d Patients should return for Visit 801, a post-treatment follow-up visit, 28±5 days after the last dose of investigational product. Patients who complete the treatment period through Visit 16 and enter the long-term extension (LTE; JAIM) will not complete Visit 801 in Study JAIA, as they will complete post-treatment follow-up at the conclusion of JAIM. Patients who discontinue investigational product (IP) but remain in the study for at least 28±5 days without IP can combine their Visit 16/ET with their Visit 801 (post-treatment follow-up visit).

h TB test(s) include PPD, QuantiFERON®-TB Gold, and T SPOT®. See Exclusion Criterion [29] for details 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 must be performed 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, have no clinical features of active TB, and have a screening chest x-ray with no evidence of active TB are not required to undergo the protocol-specific TB testing (i.e., PPD, QuantiFERON®-TB Gold, or T SPOT®) but must have a chest x-ray at screening (i.e.: ~~a~~ chest ~~x-ray~~imaging performed within the past 6 months will not be accepted).

j A chest x-ray (posterior-anterior view) will be performed at screening unless chest imaging (excluding ultrasound)~~one~~ has been performed within the past 6 months and the ~~x-ray~~imaging and reports are available for review by the investigator.

v For female patients ≥50~~40~~ and <55~~60~~ years of age who have had a cessation of menses for at least 6 and less than 12 months, an FSH test will be performed to confirm non-child-bearing potential (FSH ≥40 mIU/mL).

bb Urine should be collected as a clean catch sample.

6.1. Inclusion Criteria

Type of Patient and Disease Characteristics

[9] Male or nonpregnant, nonbreastfeeding female patient

a. …

d. The following contraception methods are considered acceptable (the patient should choose 2, and 1 must be highly effective [defined as less than 1% failure rate per year when used consistently and correctly]):

- Highly effective birth control methods:
  - ...
- Effective birth control methods:
  - …

Patients of non–child-bearing potential are not required to use birth control and they are defined as:

- Women who are infertile due to surgical sterilization (hysterectomy, bilateral oophorectomy, or tubal ligation)~~ who has had either~~
- <u>Post-menopausal – defined either as</u>
  - <u>A woman at least 50 years of age with an intact uterus, not on hormone therapy, who has either</u>
    - Cessation of menses for at least 1 year
    - At least 6 months of spontaneous amenorrhea with follicle-stimulating hormone >40 mIU/mL
- …

6.2. Exclusion Criteria

Medical Conditions

[29] Have evidence of active TB or latent TB

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

- Positive purified protein derivative (PPD) test (≥5 mm induration between approximately 2 and 3 days after application, regardless of vaccination history), medical history, clinical features, and abnormal chest x-ray at screening.
- 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, have no clinical features of active TB, and have a screening chest x-ray with no evidence of active TB may be enrolled if other entry criteria 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 (i.e., ~~a~~ chest ~~x-ray~~imaging performed within the past 6 months will not be accepted).

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

- Positive PPD test, no clinical features consistent with active TB, and a chest x-ray with no evidence of active TB at screening; or
- If the PPD test is positive and 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 prior to 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, have no clinical features of active TB, and have a screening chest x-ray with no evidence of active TB may be enrolled if other entry criteria 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 (i.e., ~~a~~ chest ~~x-ray~~imaging performed within the past 6 months will not be accepted).

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~~imaging (other than ultrasound) obtained within 6 months prior to the study are available. The chest x-ray will be reviewed by the investigator or his or her designee to exclude patients with active TB infection. In addition, patients will be tested at screening (Visit 1) for evidence of active or latent TB as described in the exclusion criteria, Section 6.2.

9.4.10. Safety Monitoring

Investigators will monitor vital signs and carefully review findings that may be associated with cardiovascular events. Adverse event 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, myocardial infarction, stroke), other cardiovascular events (such as hospitalization for unstable angina, hospitalization for heart failure, serious arrhythmia, resuscitated sudden death, cardiogenic shock, coronary interventions), VTE (DVT/PE), arterial thromboembolic events 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.

10.3.1.1. Analysis Methods

Treatment comparisons of categorical efficacy and health outcomes variables will be made using a logistic regression analysis with Firth correction (Firth 1993; Heinze and Schemper 2001). ~~The primary analysis of categorical efficacy and health outcomes variables will use a logistic regression analysis with b~~Baseline disease activity (total SLEDAI-2K ~~at screening~~ <10; ≥10), baseline corticosteroid dose (<10 mg/day; ≥10 mg/day prednisone or equivalent), region, and treatment group will be explanatory variables in the model. The p-value and 95% (unless otherwise specified) confidence interval (CI) for the odds ratio from the logistic regression model ~~are~~ will be used for primary statistical inference. The difference in percentages and 95% CI of the difference in percentages using the Newcombe-Wilson method without continuity correction ~~are~~ will be used for descriptive purposes, unless otherwise specified. ~~When logistic regression sample size requirements are not met (<5 responders in any category for any factor), the p-value from the Fisher exact test is produced instead of the odds ratio and CI.~~ Missing data will be imputed using nonresponder imputation (NRI), as described in Section 10.3.1.2.

The primary analyses for all continuous efficacy and health outcome variables will use a restricted maximum likelihood-based mixed model for repeated measures (MMRM) analysis. The model will include treatment, baseline disease activity (total SLEDAI-2K ~~at screening~~ <10; ≥10), baseline corticosteroid dose (<10 mg/day; ≥10 mg/day prednisone or equivalent), region, visit, treatment-by-visit interaction as fixed categorical effects, and baseline value and baseline value-by-visit interaction as fixed continuous effects to estimate change from baseline across

postbaseline visits. ~~An unstructured covariance structure will be used to model the between- and within-patient errors. If this analysis fails to converge, other structures will be tested using the following prespecified order: heterogeneous Toeplitz, heterogeneous autoregressive, or heterogeneous compound symmetry.~~ The following covariance structures will be used to model the between- and within-patient errors: unstructured, heterogeneous Toeplitz, heterogeneous autoregressive, autoregressive (1), and heterogeneous compound symmetry. The variance-covariance structure that results in the smallest Akaike information criterion (AIC) will be used. The Kenward-Roger method will be used to estimate the denominator degrees of freedom. Type III tests for the least-squares (LS) means will be used for the statistical comparisons. The LS mean difference, standard error, p-value, and CIs will also be reported. Treatment group comparisons at specific study visits will be tested using the t-test obtained from the MMRM results. Additional details of the MMRM method are described in Section 10.3.1.2.

10.3.1.2. Missing Data Imputation

The following imputation rules will be used:

- …
- Mixed-model repeated measure (MMRM): For continuous variables, the primary analysis will be a MMRM analysis with a missing at random assumption for handling missing data. This analysis takes into account both missingness of data and the correlation of the repeated measurements. ~~No additional imputation methods will be applied to the MMRM analysis.~~ For patients who increase use of corticosteroids above the baseline dose other than the permitted burst, or have initiation of or an increase above baseline dose in immunosuppressant or antimalarial treatment any time after baseline, any observed data after that visit will be excluded from the MMRM analysis. Observed data after permanent discontinuation of investigational product will be analyzed as observed, with the exception of any data that occur after increase in use of corticosteroids above the baseline dose other than the permitted burst, or initiation of or an increase above baseline dose in immunosuppressant or antimalarial treatment.
- Tipping point analysis: Various "tipping point" analyses will be performed to evaluate the robustness of statistical analyses for different objectives/estimands to the assumptions about missing data mechanism. This will be implemented by multiple imputation with sensitivity parameters (deltas) quantifying the degree of departure from the missingness at random (MAR) assumption by gradually worsening imputed outcomes for patients treated with baricitinib and/or improving the imputed outcomes for patients treated with placebo. The tipping point is identified as the combination of delta values for baricitinib and placebo that leads to a loss of statistical significance when evaluating baricitinib relative to placebo. Additional details of the tipping point analysis will be provided in the SAP. ~~Sensitivity analysis using multiple imputation under the missing not at random~~

~~assumption will be provided for primary and major secondary endpoints to investigate the missing data mechanism assumption. Specifically, tipping point analyses will be performed to evaluate the robustness of the statistical analyses of both categorical and continuous efficacy data by gradually worsening imputed outcomes for patients treated with baricitinib and improving the imputed outcomes for patients treated with placebo. The tipping point is identified as the delta value which will lead to a loss of statistical significance when evaluating baricitinib relative to placebo. All observed data will be analyzed as observed without censoring. Additional details of the tipping point analysis will be provided in the SAP~~.

10.3.3.1. Primary Analyses

The primary efficacy measure is SRI-4 response (defined in Section 9.1.1). A logistic regression model with Firth correction (as described in Section 10.3.1.1) will be used to test the treatment difference between baricitinib 4-mg and placebo in the proportion of patients achieving SRI-4 response at Week 52. Missing data will be imputed using the NRI method described in Section 10.3.1.2.

10.3.3.2. Major Secondary Analyses

The major secondary comparisons will be based on the multiple testing procedure detailed in Section 10.3.1.3.

Treatment comparisons in the proportion of patients achieving a binary response will be analyzed using logistic regression with Firth correction. Treatment comparisons in the continuous measures will be analyzed using MMRM. Treatment comparisons for time to first flare will be analyzed using a Cox proportional hazards model. These methods are described in Section 10.3.1.1. Missing data for the binary outcome measures will be imputed using the NRI method described in Section 10.3.1.2.

~~*10.3.6. Other Analyses*~~

~~10.3.6.1. Subgroup Analyses~~

~~To assess whether the treatment effect is similar across subgroups for the primary efficacy outcome, a logistic regression model will be used and will include treatment, stratification variables, the subgroup variable (e.g., sex) and the subgroup-by-treatment interaction. If the interaction is statistically significant at $\alpha = 0.10$, the nature of the interaction will be explored, that is, within each subgroup the treatment effect will be estimated.~~

~~Subgroups to be evaluated will include demographic characteristics (such as region or country, gender, age, and race), renal function, and baseline disease characteristics. 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.~~

11. References

[EMA] European Medicines Agency. ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials. Step 2b; 2017. EMA/CHMP/ICH/436221/2017.

Firth D. Bias reduction of maximum likelihood estimates. *Biometrika*. 1993;80(1):27-38.

Heinze G, Schemper M. A solution to the problem of monotone likelihood in Cox regression. *Biometrics*. 2001;57(1):114-119.

Appendix 2. Clinical Laboratory Tests

Clinical Laboratory Tests[a]

Hematology[b]	**Other Tests**
Mean platelet volume	Follicle-stimulating hormone (FSH) (females only)[h]
	Hepatitis B core antibody (HBcAb), total
Urinalysis[b,n]	

[a] All labs will be assayed/calculated by a Lilly-designated laboratory unless otherwise noted.

[b] Unscheduled blood chemistry hematology, and urinalysis panels may be performed at the discretion of the investigator.

[h] For female patients ≥50 and <55 years of age who have had a cessation of menses for at least 6 and less than 12 months, an FSH test will be performed to confirm non-child-bearing potential (FSH ≥40 mIU/mL)~~To confirm postmenopausal status for women ≥40 and <60 years of age, an FSH test will be performed. Non-child-bearing potential is defined as an FSH ≥40 mIU/mL and a cessation of menses for at least 12 months~~.

[n] Urine should be collected as a clean catch sample.

Leo Document ID = d45a04aa-a6e4-4fb6-9de1-fad7b2239ba3

Approver: Brenda J Crowe (AM\RM89719)
Approval Date & Time: 07-Dec-2018 20:49:45 GMT
Signature meaning: Approved

Approver: Stephanie de Bono (AM\YE91114)
Approval Date & Time: 07-Dec-2018 20:54:15 GMT
Signature meaning: Approved