

Protocol (c) I4V-CR-JAGS

Randomized, Double-Blind, Placebo-Controlled, Phase 3 Study Evaluating the Efficacy and Safety of Baricitinib in Patients with Moderately to Severely Active Rheumatoid Arthritis Who Have Had an Inadequate Response to Methotrexate Therapy

NCT02265705

Approval Date: 07-Dec-2016

**1. Protocol I4V-CR-JAGS(c)
A Randomized, Double-Blind, Placebo-Controlled, Phase 3
Study Evaluating the Efficacy and Safety of Baricitinib in
Patients with Moderately to Severely Active Rheumatoid
Arthritis Who Have Had an Inadequate Response to
Methotrexate Therapy**

Confidential Information

The information contained in this protocol is confidential and is intended for the use of clinical investigators. It is the property of Eli Lilly and Company or its subsidiaries and should not be copied by or distributed to persons not involved in the clinical investigation of baricitinib, unless such persons are bound by a confidentiality agreement with Eli Lilly and Company or its subsidiaries. This document and its associated attachments are subject to United States Freedom of Information Act (FOIA) Exemption 4.

Baricitinib (LY3009104)

A Phase 3, randomized, double-blind, placebo-controlled, parallel-group study of baricitinib for the treatment of moderately to severely active rheumatoid arthritis during a 52-week treatment period in patients with inadequate response to methotrexate therapy.

Eli Lilly and Company
Indianapolis, Indiana USA 46285

Protocol Electronically Signed and Approved by Lilly: 05-Nov-2013.
Amendment (a) Electronically Signed and Approved by Lilly
on 03-Apr-2014.
Amendment (b) Electronically Signed and Approved by Lilly
on 27-Feb-2015.
Amendment (c) Electronically Signed and Approved by Lilly
on approval date provided below

Approval Date: 07-Dec-2016 GMT

2. Synopsis

Study Rationale

Baricitinib (LY3009104) is an oral Janus kinase 1 (JAK1)/Janus kinase 2 (JAK2) selective inhibitor representing a potentially effective therapy for treatment of patients with moderately to severely active rheumatoid arthritis (RA). The rationale for the current study is to confirm the efficacy and to continue to define the safety profile of 4 mg baricitinib when administered once daily (QD) to patients with RA who have had an inadequate response to methotrexate (MTX) therapy. The safety and tolerability data from this study are intended to inform the current understanding of the benefit-risk relationship for baricitinib in patients with RA.

Clinical Protocol Synopsis: Study I4V-CR-JAGS(b)

Name of Investigational Product: Baricitinib (LY3009104)	
Title of Study: A Randomized, Double-Blind, Placebo- Controlled, Phase 3 Study Evaluating the Efficacy and Safety of Baricitinib in Patients with Moderately to Severely Active Rheumatoid Arthritis Who Have Had an Inadequate Response to Methotrexate Therapy	
Number of Planned Patients: Entered: 411 Enrolled/Randomized: 288 (144 on baricitinib, 144 on placebo) Completed: Approximately 245	Phase of Development: 3
Length of Study: 2 years Planned first patient visit: Jun 2014 Planned last patient visit: Jun 2016	
<p>Objectives: The primary objective of the study is to determine whether baricitinib is superior to placebo in the treatment of patients with moderately to severely active rheumatoid arthritis (RA) despite methotrexate treatment (that is, inadequate response to methotrexate [MTX-IR]), as assessed by the proportion of patients achieving a 20% improvement in American College of Rheumatology criteria (ACR20) at Week 12.</p> <p>The secondary objectives of the study are to evaluate the efficacy of baricitinib versus placebo as assessed by:</p> <ul style="list-style-type: none"> • change from baseline to Week 12 in Health Assessment Questionnaire-Disability Index (HAQ-DI) score compared to placebo • change from baseline to Week 12 in Disease Activity Score modified to include the 28 diarthroidal joint count (DAS28)-high-sensitivity C-reactive protein (hsCRP) compared to placebo • proportion of patients achieving a Simplified Disease Activity Index (SDAI) score ≤ 3.3 at Week 12 compared to placebo • mean duration of morning joint stiffness in the 7 days prior to Week 12 compared to placebo as collected in diaries • mean severity of morning joint stiffness in the 7 days prior to Week 12 compared to placebo as collected in diaries • mean Worst Tiredness numeric rating scale (NRS) in the 7 days prior to Week 12 compared to placebo as collected in diaries • mean Worst Joint Pain NRS in the 7 days prior to Week 12-compared to placebo as collected in diaries <p>The exploratory objectives in the study include:</p> <ul style="list-style-type: none"> • change from baseline at Weeks 16, 24 and 52 in structural joint damage as measured by modified Total Sharp Score (mTSS; van der Heijde 2000) • proportion of patients with mTSS change ≤ 0 from baseline at Weeks 16, 24, and 52 • change from baseline to Weeks 16, 24, and 52 in joint space narrowing and bone erosion scores • proportion of patients achieving DAS28-hsCRP ≤ 3.2 and DAS28-hsCRP < 2.6 at Weeks 12, 24, and 52 • proportion of patients achieving HAQ-DI improvement ≥ 0.22 and ≥ 0.3 at Weeks 12, 24, and 52 • proportion of patients achieving ACR20 at Week 24 and Week 52 • proportion of patients achieving 50% improvement in American College of Rheumatology criteria (ACR50) at Weeks 12, 24, and 52 • proportion of patients achieving 70% improvement in American College of Rheumatology criteria (ACR70) at Weeks 12, 24, and 52 • percentage change from baseline to Weeks 12, 24, and 52 in individual components of the ACR Core Set • change from baseline through Week 52 in DAS28-hsCRP • change from baseline through Week 52 in DAS28-erythrocyte sedimentation rate (ESR) • proportion of patients achieving DAS28-ESR ≤ 3.2 and DAS28-ESR < 2.6 at Weeks 12, 24, and 52 • change from baseline to Week 24 and Week 52 in HAQ-DI score • change from baseline to Weeks 12, 24, and 52 in Clinical Disease Activity Index (CDAI) score 	

- change from baseline to Weeks 12, 24, and 52 in SDAI score
- proportion of patients achieving a CDAI score ≤ 2.8 at Weeks 12, 24, and 52
- proportion of patients achieving an SDAI score ≤ 3.3 , Week 24, and Week 52
- proportion of patients achieving ACR/European League Against Rheumatism (EULAR) remission according to the Boolean-based definition at Weeks 12, 24, and 52
- proportion of patients achieving moderate and good EULAR responses based on the DAS28 at Weeks 12, 24, and 52
- evaluation of hybrid ACR (bounded) response at Weeks 12, 24 and 52
- change from baseline to Weeks 12, 24, and 52 in Functional Assessment of Chronic Illness Therapy Fatigue (FACIT-F) scale scores
- change from baseline to Weeks 12, 24 and 52 in European Quality of Life-5 Dimensions-5 Level (EQ-5D-5L) scores
- change from baseline to Weeks 12, 24 and 52 in Work Productivity and Activity Impairment-Rheumatoid Arthritis (WPAI-RA) scores
- evaluation of healthcare resource utilization through Week 52
- change from baseline to Weeks 12, 24 and 52 in duration of morning joint stiffness assessed at the visit using an electronic patient-reported outcomes (ePRO) tablet
- change from baseline to Weeks 12, 24 and 52 in Worst Tiredness NRS assessed at the visit using an ePRO tablet
- change from baseline to Weeks 12, 24 and 52 in Worst Joint Pain NRS assessed at the visit using an ePRO tablet.

Study Design: Study I4V-CR-JAGS(b) (JAGS) will be a 52-week, Phase 3, multicenter, randomized, double-blind, placebo- controlled, parallel-group, study comparing the efficacy of baricitinib versus placebo in patients with moderately to severely active RA who have had an insufficient response to MTX and who have never been treated with a biologic disease-modifying antirheumatic drug (DMARD) (that is, MTX-IR patients). Baricitinib will be administered as a 4-mg tablet QD. The dose for patients with renal impairment, defined as estimated glomerular filtration rate (eGFR) < 60 mL/min/1.73 m², will be 2 mg baricitinib QD. Patients not assigned to the baricitinib treatment arm will receive either a 4-mg or 2-mg matching placebo tablet.

A total of 288 patients are planned for enrollment in this study (144 to receive baricitinib, and 144 to receive placebo). After the Week 52 study visit, eligible patients may proceed to a separate extension study (Study I4V-MC-JADY [JADY]) lasting for up to 2 years or to the posttreatment follow-up period of this study (Part C).

Study JAGS will consist of 4 parts:

- Screening: Screening period lasting from 3 to 42 days prior to Visit 2 (Week 0)
- Part A: double-blind, placebo-controlled period from Week 0 through Week 24
- Part B: open-label period from Week 24 through Week 52
- Part C: posttreatment follow-up period.

Diagnosis and Main Criteria for Inclusion and Exclusions: The proposed patient population for Study I4V-CR-JAGS(b) (JAGS) will be ambulatory adult patients with moderately to severely active RA who have had an insufficient response to MTX and who have never been treated with a biologic DMARD. Patients will be eligible for participation only if they have an inadequate response to MTX as evidenced by the presence of at least 6/68 tender joints and 6/66 swollen joints and an hsCRP measurement ≥ 6 mg/L.

Investigational Product, Dosage, and Mode of Administration or Intervention:

Baricitinib 4 mg administered orally QD

Planned Duration of Treatment: 52 weeks.

Screening period: 3 to 42 days

Treatment period: 52 weeks

Follow-up period: 28 days

Reference Therapy, Dose, and Mode of Administration or Comparative Intervention:

Placebo tablets matching baricitinib will be administered orally QD.

Patients will continue to take their background MTX therapy during the course of the study.

Criteria for Evaluation:

Efficacy

The following efficacy measures will be assessed in this study:

- ACR20, ACR50, and ACR70 indices
- HAQ-DI
- DAS28-hsCRP and DAS28-ESR
- Hybrid ACR (bounded) response measure
- SDAI
- CDAI
- EULAR responses based on DAS28
- mTSS (includes joint space narrowing score and bone erosion score)

Health Outcomes

The following health outcome measures will be administered in this study:

- duration and severity of morning joint stiffness
- recurrence of joint stiffness during the day
- tiredness severity numeric rating scale (Worst Tiredness NRS)
- pain severity numeric rating scale (Worst Joint Pain NRS)
- FACIT-F
- WPAI-RA
- EQ-5D-5L
- Quick Inventory of Depressive Symptomatology Self-Rated-16 (QIDS-SR₁₆)
- healthcare resource utilization

Safety

The following safety measures will be assessed in this study:

- adverse events (AEs)
- adverse events of special interest (AESIs)
- serious adverse events (SAEs)
- suspected unexpected serious adverse reactions (SUSARs)

- physical examinations
- electrocardiograms (ECGs)
- chest x-ray and tuberculosis (TB) testing
- vital signs (blood pressure and heart rate) and physical characteristics
- standard laboratory tests (including hematology, clinical chemistry, urinalysis, lipid profile, eGFR, iron studies, hsCRP, and ESR)
- concomitant medications

AESIs will include:

- severe or opportunistic infections
- myelosuppressive events of anemia, leukopenia, neutropenia, lymphopenia, and thrombocytopenia
- thrombocytosis
- elevations in alanine aminotransferase (ALT) or aspartate aminotransferase (AST) (>3 times ULN) with total bilirubin (>2 times ULN)

Patients with these laboratory value specified events will be identified using the same criteria for the interruption of investigational product with the exception of anemia, which will be identified using the same criteria for the discontinuation of investigational product, and thrombocytosis, which will be defined as a platelet count >600,000/ μ L.

Pharmacogenetics

Samples will be stored, and analysis may be performed on genetic variants thought to play a role in active RA and the JAK/signal transducers and activators of transcription (STAT) signaling pathways.

Statistical Methods:

Stratification:

Randomization will be stratified at the country level and joint erosion status (1-2 joint erosions plus seropositivity versus at least 3 joint erosions)

Sample size:

288 patients (144 in baricitinib, 144 in placebo) will provide:

- 99% power to detect a difference between the baricitinib and placebo treatment groups in ACR20 response rate (60% versus 35%) at Week 12 based on a 2-sided chi-square test at a significance level of 0.05.

The above sample size and power estimates are based on nQuery advisor 7.0.

General considerations:

All statistical tests of treatment effects will be performed at 2-sided significance levels of 0.05, unless otherwise stated. Treatment comparisons of discrete efficacy variables will be made using a logistic regression analysis with country, treatment groups and joint erosion status in the model. The proportions and 95% confidence interval (CI) will be reported. Treatment comparisons of continuous efficacy variables will be made using analysis of covariance (ANCOVA) with country, treatment groups, joint erosion status and baseline value in the model. Type III sums of squares for the least-squares (LS) means will be used for the statistical comparison; the 95% CI will also be reported.

When a restricted maximum likelihood (REML)-based mixed model for repeated measures (MMRM) is used, the model will include the fixed, categorical effects of treatment, country, joint erosion status, visit, treatment-by-visit-interaction, as well as the continuous, fixed covariates of baseline score and baseline score-by-visit-interaction. An unstructured (co)variance structure will be used to model the within-patient errors. If this analysis fails to converge, other structures will be tested. Type III sums of squares for the LS means will be used for the statistical comparison; the 95% CI will also be reported. Treatment group comparisons with placebo at Week 12 and all other visits will be tested. Further details on the MMRM will be described in the statistical analysis plan (SAP).

Fisher's exact test will be used for all AEs, discontinuation, and other categorical safety variables. Continuous vital

signs and other continuous safety variables including laboratory parameters will be analyzed by using ANCOVA with treatment and baseline value in the model.

Analysis population:

Unless otherwise specified, the efficacy and safety analyses will be conducted on a modified intent-to-treat basis (Gillings and Koch 1991). This analysis set includes all data from all randomized patients who received at least 1 dose of study drug. Patients will be analyzed according to the treatment to which they were assigned.

Sensitivity analyses of the primary endpoints may be conducted and will be described in the statistical analysis plan.

Missing data imputation:

In accordance with precedent set with other Phase 3 RA trials (Cohen et al. 2006; Keystone et al. 2004, 2008, 2009; Smolen et al. 2008, 2009), the following methods for imputation of missing data will be used:

- Nonresponder imputation (NRI): All patients who discontinue the study or the study treatment at any time for any reason will be defined as nonresponders for the NRI analysis for categorical variables, such as ACR20/50/70, from the time of discontinuation and onward. Patients who receive rescue therapy at Week 16 will be analyzed as nonresponders after Week 16 and onward. Randomized patients without available data at a postbaseline visit will be defined as nonresponders for the NRI analysis at that visit.
- Linear extrapolation method: The linear extrapolation method will be used for analysis of the structural progression endpoint (van der Heijde mTSS) to impute missing data at time points when analyses are conducted (including the analyses at Week 24 and Week 52). For patients who discontinue the study or the study treatment or for patients who miss a radiograph for any reason, baseline data and the most recent radiographic data prior to discontinuation or the missed radiograph, adjusted for time, will be used for linear extrapolation to impute missing data at subsequent time points. For patients who receive rescue therapy starting at Week 16 or at any time point thereafter, baseline data and the most recent radiographic data prior to initiation of rescue therapy, adjusted for time, will be used for linear extrapolation. All patients originally randomized to placebo will be switched to active treatment at Week 24; thus, there will be no observed data for the placebo arm for the structural comparison at subsequent time points. Therefore, baseline data and the most recent radiographic data prior to initiation of the new therapy, adjusted for time, will be used for linear extrapolation at Week 52. The linear extrapolation method has been established as an appropriate missing data imputation method in other Phase 3 RA trials (Keystone et al. 2004, 2008, 2009; Cohen et al. 2006; Smolen et al. 2008, 2009). Missing postbaseline values will be imputed only if both a baseline value and a postbaseline value from another time point are available, as long as the patient was on the same treatment at each applicable time point.
- Multiple imputation: Other methods than linear extrapolation that include multiple imputation will be employed for analysis of the structural progression endpoint (van der Heijde mTSS) to impute missing data at Week 24. Details will be provided in the SAP.
- Modified baseline observation carried forward (mBOCF): The mBOCF method will be used for the analysis of secondary continuous endpoints (unless otherwise stated). For patients who discontinue the study or the study treatment because of an AE, including death, the baseline observation will be carried forward to the corresponding endpoint for evaluation. For patients who discontinue the study for reason(s) other than an AE, the last nonmissing postbaseline observation prior to discontinuation will be carried forward to the corresponding endpoint for evaluation. For patients who receive rescue therapy starting at Week 16, the last nonmissing observation at or before rescue will be carried forward to subsequent time points for evaluation.
- Modified last observation carried forward (mLOCF): For all continuous measures including safety analyses, the mLOCF will be a general approach to impute missing data unless otherwise specified. For patients who receive rescue therapy starting at Week 16, the last nonmissing observation at or before rescue will be carried

forward to subsequent time points for evaluation. For all other patients discontinuing from the study or the study treatment for any reason, the last nonmissing postbaseline observation before discontinuation will be carried forward to subsequent time points for evaluation. After mLOCF imputation, data from patients with nonmissing baseline and postbaseline observations will be included in the analyses. The mLOCF will be used as a sensitivity method to the mBOCF method with respect to the secondary endpoints.

- Other methods for data imputation for analysis of endpoints other than mTSS may be used and will be described in the SAP.

Analyses methods:

Primary analysis:

A logistic regression model with treatment, country and joint erosion status in the model will be used to test the treatment difference between baricitinib and placebo in the proportion of patients achieving ACR20 response at Week 12. The NRI method as described above will be used to impute missing data.

Secondary analyses:

- Comparison between baricitinib and placebo in change from baseline to Week 12 in HAQ-DI score: An ANCOVA analysis with treatment, country, joint erosion status and baseline score in the model will be used to test the treatment difference between baricitinib and placebo in change from baseline to Week 12 in HAQ-DI. The mBOCF and mLOCF approaches as described above will be used to impute missing data.
- Comparison between baricitinib and placebo in change from baseline to Week 12 in DAS28-hsCRP: An ANCOVA model with treatment, country, joint erosion status, and baseline score in the model will be used to test the treatment difference between baricitinib and placebo in change from baseline to Week 12 in DAS28-hsCRP. The mBOCF and mLOCF approaches as described above will be used to impute missing data.
- Comparison between baricitinib and placebo in proportion of patients achieving an SDAI score ≤ 3.3 at Week 12: A logistic regression model with treatment, country, and joint erosion status in the model will be used to test the treatment difference between baricitinib and placebo in the proportion of patients achieving an SDAI score ≤ 3.3 response at Week 12. The NRI method as described above will be used to impute missing data.
- Comparison between baricitinib and placebo in mean duration of morning stiffness in the 7 days prior to Week 12: An ANCOVA analysis with treatment, country, joint erosion status and baseline score (Day 1 on the diaries) in the model will be used to test the treatment difference between baricitinib and placebo in mean duration of morning stiffness.
- Comparison between baricitinib and placebo in mean severity of morning joint stiffness in the 7 days prior to Week 12: An ANCOVA analysis with treatment, country, joint erosion status and baseline score (Day 1 on the diaries) in the model will be used to test the treatment difference between baricitinib and placebo in mean severity of morning joint stiffness.
- Comparison between baricitinib and placebo in mean “Worst Tiredness” NRS in the 7 days prior to Week 12: An ANCOVA model with treatment, country, joint erosion status, and baseline score (Day 1 on the diaries) in the model will be used to test the treatment difference between baricitinib and placebo in mean “Worst Tiredness” NRS.
- Comparison between baricitinib and placebo in mean “Worst Joint Pain” NRS in the 7 days prior to Week 12: An ANCOVA model with treatment, country, joint erosion status, and baseline score (Day 1 on the diaries) in the model will be used to test the treatment difference between baricitinib and placebo in mean “Worst Joint Pain” NRS.

Besides the analyses listed in the objective section, the following analyses will be performed as well.

- change from baseline in mTSS
- proportion of patients with mTSS change ≤ 0
- change from baseline in joint space narrowing and bone erosion scores
- proportion of patients achieving DAS28-hsCRP ≤ 3.2 and < 2.6

- proportion of patients achieving HAQ-DI improvement ≥ 0.22 and ≥ 0.3
- proportions of patients achieving ACR20 (other than primary), ACR50, and ACR70
- percentage change from baseline in individual components of the ACR Core Set
- change from baseline in DAS28-hsCRP
- change from baseline in DAS28-ESR
- proportion of patients achieving DAS28-ESR ≤ 3.2 and < 2.6
- change from baseline in HAQ-DI
- change from baseline in CDAI score
- change from baseline in SDAI score
- proportion of patients achieving CDAI score ≤ 2.8
- proportion of patients achieving SDAI score ≤ 3.3 (ACR/EULAR remission according to the Index-based definition)
- proportion of patients achieving ACR/EULAR remission according to the Boolean-based definition
- proportion of patients achieving moderate response and good response based on EULAR response criteria
- hybrid ACR (bounded) response measure
- change from baseline in fatigue score in the FACIT-F
- change from baseline in health states, VAS score, and index score of EQ-5D-5L
- change from baseline in absenteeism, presenteeism, work productivity loss, and activity impairment scores of WPAI-RA
- individual items and total score in healthcare resource utilization
- change from baseline in duration of morning joint stiffness assessed at the visit using an ePRO tablet
- change from baseline in Worst Tiredness NRS assessed at the visit using an ePRO tablet
- change from baseline in Worst Joint Pain NRS assessed at the visit using an ePRO tablet.

Approaches similar to the primary and above secondary analyses will be used to analyze other exploratory and health outcome measures. Analysis methods will consist of ANCOVA and logistic regression, as described above, using mBOCF, linear extrapolation, NRI, and mLOCF missing data imputation methods, as applicable.

In addition, continuous parameters that are collected at repeated visits throughout the double-blind period may be analyzed using a MMRM approach as described in the general consideration section.

Safety:

All safety data will be descriptively summarized by treatment groups and analyzed using the mITT population. For the purposes of calculating changes from baseline during Part A and Part B treatment, baseline may be considered as the original baseline for Part A – prior to any study drug received – or at the visits that mark the transition between treatments received, depending on the purpose of the summary or comparison being made. Summaries within Parts A and B will be conducted by treatment group used within each respective part. Additional summaries by treatment group that continue unchanged across multiple parts of the study will be conducted, as appropriate.

Treatment-emergent adverse events are defined as AEs that first occurred or worsened in severity after the first dose of study treatment. AEs will be considered treatment emergent for nonresponders assigned to baricitinib if the AEs first occurred or worsened in severity after the visit at which rescue therapy is assigned. The number of TEAEs as well as the number and percentage of patients who experienced at least 1 TEAE will be summarized using Medical Dictionary for Regulatory Activities (MedDRA) for each system organ class (SOC) (or a body system) and each preferred term by treatment group. TEAEs will also be summarized by relationship to treatment and by severity within each treatment group. Serious adverse events (SAEs) and AEs that lead to study drug discontinuation will also be summarized by treatment group. Fisher's exact test will be used to perform the

statistical comparisons between baricitinib and placebo.

All clinical laboratory results will be descriptively summarized by treatment group. Individual results that are outside of normal range will be flagged in data listings. Quantitative clinical hematology, chemistry, and urinalysis variables obtained at the baseline to postbaseline visits will be summarized as changes from baseline by treatment group and analyzed using ANCOVA with treatment and baseline value in the model. Categorical variables, including the incidence of abnormal values and incidence of AESIs, will be summarized by frequency and percentage of patients in corresponding categories. Shift tables will be presented for selected measures. Observed values and changes from baseline (predose or screening if missing) for vital signs and physical characteristics will be descriptively summarized by treatment group and time point. Change from baseline to postbaseline in vital signs, QIDS-SR16, and body weight will be analyzed using ANCOVA with treatment and baseline value in the model.

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

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

3. Table of Contents

A Randomized, Double-Blind, Placebo- Controlled, Phase 3 Study Evaluating the Efficacy and Safety of Baricitinib in Patients with Moderately to Severely Active Rheumatoid Arthritis Who Have Had an Inadequate Response to Methotrexate Therapy

Section	Page
1. Protocol I4V-CR-JAGS(c) A Randomized, Double-Blind, Placebo-Controlled, Phase 3 Study Evaluating the Efficacy and Safety of Baricitinib in Patients with Moderately to Severely Active Rheumatoid Arthritis Who Have Had an Inadequate Response to Methotrexate Therapy	1
2. Synopsis	2
3. Table of Contents.....	11
4. Abbreviations and Definitions.....	18
5. Introduction	24
6. Objectives.....	27
6.1. Primary Objective	27
6.2. Secondary Objectives.....	27
6.3. Exploratory Objective	27
7. Investigational Plan.....	29
7.1. Summary of Study Design.....	29
7.1.1. Screening Period	30
7.1.2. Part A and Part B: 24-Week Double-Blind Placebo-Controlled, (Part A) Period and 28-Week Open-Label (Part B) Period.....	30
7.1.3. Part C: Posttreatment Follow-Up Period.....	31
7.1.4. Study Extensions.....	31
7.2. Discussion of Design and Control	31
8. Study Population.....	33
8.1. Entry Criteria	33
8.1.1. Inclusion Criteria	33
8.1.2. Exclusion Criteria	34
8.2. Enrollment Criteria.....	38
8.2.1. Inclusion Criteria	38
8.2.2. Exclusion Criteria	38

8.2.3.	Rationale for Exclusion of Certain Study Candidates	41
8.3.	Discontinuations	41
8.3.1.	Discontinuation of Patients.....	41
8.3.1.1.	Interruption of Study Drug.....	41
8.3.1.2.	Discontinuation of Study Drug	42
8.3.1.3.	Discontinuation from the Study	43
8.3.2.	Discontinuation of Study Sites	44
8.3.3.	Discontinuation of the Study	44
9.	Treatment.....	45
9.1.	Treatments Administered	45
9.2.	Materials and Supplies	45
9.3.	Method of Assignment to Treatment	46
9.4.	Rationale for Selection of Doses in the Study.....	46
9.5.	Selection and Timing of Doses.....	47
9.5.1.	Special Treatment Considerations	47
9.6.	Continued Access to Investigational Product.....	48
9.7.	Blinding.....	48
9.8.	Concomitant Therapy.....	49
9.9.	Treatment Compliance	50
10.	Efficacy, Health Outcome Measures, Safety Evaluations, Sample Collection and Testing, and Appropriateness of Measurements	52
10.1.	Efficacy Measures.....	52
10.1.1.	Primary Efficacy Measure.....	52
10.1.2.	Secondary Efficacy Measures.....	52
10.1.2.1.	Disease Activity Score-Erythrocyte Sedimentation Rate and Disease Activity Score–High Sensitivity C-Reactive Protein.....	52
10.1.2.2.	ACR50 and ACR70 Indices.....	52
10.1.2.3.	Hybrid American College of Rheumatology Response Measure.....	52
10.1.2.4.	ACR/EULAR Rheumatoid Arthritis Remission	53
10.1.2.5.	Simplified Disease Activity Index (SDAI).....	53
10.1.2.6.	Clinical Disease Activity Index (CDAI).....	54
10.1.2.7.	European League Against Rheumatism Responder Index	54
10.1.2.8.	Van der Heijde Modified Total Sharp Score (mTSS)	54
10.2.	Health Outcome Measures.....	55
10.3.	Safety Evaluations.....	57
10.3.1.	Adverse Events (AEs).....	57
10.3.1.1.	Adverse Events of Special Interest.....	58

10.3.1.2.	Serious Adverse Events (SAEs)	58
10.3.1.2.1.	Suspected Unexpected Serious Adverse Reactions (SUSARs)	59
10.3.2.	Other Safety Measures	60
10.3.2.1.	Measures Evaluated During Screening Period	60
10.3.2.2.	Measures Evaluated During Screening and Treatment Periods	60
10.3.3.	Safety Monitoring	61
10.3.4.	Complaint Handling	62
10.4.	Sample Collection and Testing	62
10.4.1.	Samples for Standard Laboratory Testing	62
10.4.2.	Exploratory Stored Samples	63
10.4.2.1.	Pharmacogenetics	63
10.4.2.2.	Nonpharmacogenetic/Biomarker Stored Samples	64
10.5.	Appropriateness of Measurements	64
11.	Data Quality Assurance	65
11.1.	Data Capture System	65
12.	Sample Size and Statistical Methods	66
12.1.	Determination of Sample Size	66
12.2.	Statistical and Analytical Plans	66
12.2.1.	General Considerations	66
12.2.2.	Patient Disposition	68
12.2.3.	Patient Characteristics	68
12.2.4.	Concomitant Therapy	68
12.2.5.	Treatment Compliance	69
12.2.6.	Primary Outcome and Methodology	69
12.2.7.	Efficacy Analyses	69
12.2.8.	Health Outcome Analyses	71
12.2.9.	Safety Analyses	71
12.2.10.	Subgroup Analyses	72
12.2.11.	Planned Analyses	72
13.	Informed Consent, Ethical Review, and Regulatory Considerations	73
13.1.	Informed Consent	73
13.2.	Ethical Review	73
13.3.	Regulatory Considerations	73
13.3.1.	Investigator Information	74
13.3.2.	Protocol Signatures	74
13.3.3.	Final Report Signature	74

14. References 75

List of Tables

Table	Page
Table JAGS.8.1. Criteria for Temporary Discontinuation of Investigational Product	42
Table JAGS.8.2. Criteria for Permanent Discontinuation of Investigational Product	43
Table JAGS.9.1. Treatment Regimens	45
Table JAGS.10.1. Scoring Method for the Hybrid American College of Rheumatology Response Measure	53
Table JAGS.10.2. Categorization of Patients as Nonresponders, Moderate Responders, or Good Responders	54

List of Figures

Figure

Page

Figure JAGS.7.1. Illustration of study design for Clinical Protocol I4V-CR-JAGS. 30

List of Attachments

Attachment		Page
Attachment 1.	Protocol JAGS Study Schedule	79
Attachment 2.	Protocol JAGS Clinical Laboratory Tests.....	86
Attachment 3.	Protocol JAGS Hepatic Monitoring Tests for Treatment-Emergent Abnormality	88
Attachment 4.	Protocol Amendment I4V-CR-JAGS (c) Summary A Randomized, Double-Blind, Placebo-Controlled, Phase 3 Study Evaluating the Efficacy and Safety of Baricitinib in Patients with Moderately to Severely Active Rheumatoid Arthritis Who Have Had an Inadequate Response to Methotrexate Therapy	89

4. Abbreviations and Definitions

Term	Definition
ACR	American College of Rheumatology
ACR20	20% improvement in American College of Rheumatology criteria
ACR50	50% improvement in American College of Rheumatology criteria
ACR70	70% improvement in American College of Rheumatology criteria
adverse event (AE)	Any untoward medical occurrence in a patient or clinical investigation subject administered a pharmaceutical product and 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 event of special interest
ALT	alanine aminotransferase
ANCOVA	analysis of covariance
anti-CCP	anticyclic citrullinated peptide
AST	aspartate aminotransferase
audit	a systematic and independent examination of the trial-related activities and documents to determine whether the evaluated trial-related activities were conducted, and the data were recorded, analyzed, and accurately reported according to the protocol, applicable standard operating procedures (SOPs), good clinical practice (GCP), and the applicable regulatory requirement(s)
blinding	A procedure in which one or more parties to the trial are kept unaware of the treatment assignment(s). Single-blinding usually refers to the patient(s) being unaware, and double-blinding usually refers to the patient(s), investigator(s), monitor(s), and in some cases, select sponsor personnel being unaware of the treatment assignment(s).
case report form (CRF) and electronic case report form (eCRF)	sometimes referred to as clinical report form: a printed or electronic form for recording study participants' data during a clinical study, as required by the protocol. This study uses an electronic case report form.
CDAI	Clinical Disease Activity Index
cDMARD	conventional disease-modifying antirheumatic drugs
CI	confidence interval
CL	central compartment clearance

clinical research physician	Individual responsible for the medical conduct of the study. Responsibilities of the clinical research physician may be performed by a physician, clinical research scientist, global safety physician or other medical officer.
CNS	central nervous system
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.
compliance	adherence to all the trial-related requirements, good clinical practice (GCP) requirements, and the applicable regulatory requirements
CRP	C-reactive protein
CSR	Clinical Study Report
DAS	Disease Activity Score
DAS28	Disease Activity Score modified to include the 28 diarthroidal joint count
DMARD	disease-modifying antirheumatic drug
DNA	deoxyribonucleic acid
ECG	Electrocardiogram
efficacy/ effectiveness	Efficacy is the ability of a treatment to achieve a beneficial intended result. Effectiveness is the measure of the produced effect of an intervention when carried out in a clinical environment.
eGFR	estimated glomerular filtration rate
end of study (trial)	the date of the last visit or last scheduled procedure shown in the study schedule for the last active patient in the study
enroll/randomize	The act of assigning a patient to a treatment. Patients who are enrolled/randomized in the trial are those who have been assigned to a treatment.
enter	The act of obtaining informed consent for participation in a clinical trial from patients deemed eligible or potentially eligible to participate in the clinical trial. Patients entered into a trial are those who sign the informed consent document directly or through their legally acceptable representatives.
ePRO	electronic Patient-Reported Outcome
EQ-5D-5L	European Quality of Life-5 Dimensions-5 Level
ESR	erythrocyte sedimentation rate
ethical review board (ERB)	a board or committee (institutional, regional, or national) composed of medical and nonmedical members whose responsibility is to verify that the safety, welfare, and human rights of the patients participating in a clinical trial are

	protected
EULAR	European League Against Rheumatism
EULAR28	European League Against Rheumatism Responder Index based on the 28-joint count
FACIT-F	Functional Assessment of Chronic Illness Therapy Fatigue
FSH	follicle-stimulating hormone
GCP	good clinical practice
HAQ-DI	Health Assessment Questionnaire–Disability Index
HBsAb	hepatitis B surface antibody
HBV	hepatitis B virus
HCV	hepatitis C virus
HDL	high-density lipoprotein
HIV	human immunodeficiency virus
hsCRP	high-sensitivity C-reactive protein
IB	Investigator’s Brochure
IC₅₀	inhibitory concentration of 50%
ICF	informed consent form
ICH	International Conference on Harmonisation
IL-6	interleukin-6
intention to treat (ITT)	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.
investigator	A person responsible for the conduct of the clinical study at a study site. If a study is conducted by a team of individuals at a study site, the investigator is the responsible leader of the team and may be called the principal investigator.
IWRS	interactive web-response system
JAK	Janus kinase
JAK1	1 of 4 identified members of the family of Janus kinases
JAK2	1 of 4 identified members of the family of Janus kinases

JAK3	1 of 4 identified members of the family of Janus kinases
Ka	absorption rate constant
LC/MS/MS	liquid chromatography, tandem mass spectrometry
LDL	low-density lipoprotein
legal representative	an individual, judicial, or other body authorized under applicable law to consent on behalf of a prospective patient, to the patient's participation in the clinical trial
LS means	least square means
mBOCF	modified baseline observation carried forward
MCS	Mental Component Score
MDRD	Modification of Diet in Renal Disease
MedDRA	<i>Medical Dictionary for Regulatory Activities</i>
mITT	modified intent-to-treat
mLOCF	modified last observation carried forward
MMRM	mixed model for repeated measures
MRI	magnetic resonance imaging
mRNA	messenger ribonucleic acid
mTSS	modified Total Sharp Score
MTX	Methotrexate
MTX-IR	patients who have had an inadequate response to methotrexate
NRI	nonresponder imputation
NRS	numeric rating scale
NSAID	nonsteroidal anti-inflammatory drug
patient	a study participant who has the disease or condition for which the investigational product is targeted
PCS	Physical Component Score
PI	principal investigator
PK	Pharmacokinetic
per-protocol (set)	the set of data generated by the subset of patients who sufficiently complied

(PP)	with the protocol to ensure that these data would be likely to exhibit the effects of treatment, according to the underlying scientific model
PPD	purified protein derivative
QD	once daily
QIDS-SR₁₆	Quick Inventory of Depressive Symptomatology Self-Rated-16
RA	rheumatoid arthritis
REML	restricted maximum likelihood
SAE	serious adverse event
SAP	statistical analysis plan
screening	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
SD	standard deviation
SDAI	Simplified Disease Activity Index
SJC	swollen joint count
SOC	system organ class
STAT	signal transducers and activators of transcription
subject	An individual who is or becomes a participant in clinical research, either as a recipient of the investigational product(s) or as a control. A subject may be either a healthy human or a patient.
SUSARs	suspected unexpected serious adverse reactions
TB	Tuberculosis
TJC	tender joint count
TNF	tumor necrosis factor
TNF-α	tumor necrosis factor-alpha
treatment-emergent adverse event (TEAE)	any untoward medical occurrence that either occurs or worsens at any time after treatment baseline and which does not necessarily have to have a causal relationship with this treatment (also called treatment-emergent signs and symptoms [TESS])
TSH	thyroid-stimulating hormone
TYK2	tyrosine kinase 2
ULN	upper limit of normal
V1	central compartment volume of distribution

V2	peripheral volume of distribution
VAS	visual analog scale
WBC	white blood cell
WPAI-RA	Work Productivity and Activity Impairment-Rheumatoid Arthritis

A Randomized, Double-Blind, Placebo- Controlled, Phase 3 Study Evaluating the Efficacy and Safety of Baricitinib in Patients with Moderately to Severely Active Rheumatoid Arthritis Who Have Had an Inadequate Response to Methotrexate Therapy

5. Introduction

Rheumatoid arthritis (RA) is a systemic inflammatory autoimmune disease. The disease has variable expression and outcome ranging from mild, limited disease to severe disease associated with progressive joint destruction, significantly compromised quality of life, and reduced survival (Smolen and Steiner 2003; Colmegna et al. 2012). Substantial comorbidity can be seen outside of the musculoskeletal system, with excess cardiovascular risk, dyslipidemia, and propensity for infection partially related to the need for treatment with immunomodulatory agents (Curtis et al. 2007; Kitas and Gabriel 2011).

Management of RA has improved substantially in recent years. In addition to reduction of signs and symptoms, improvement of physical function, and inhibition of structural damage, better patient outcomes and clinical remission are now considered achievable goals. Therefore, the current recommended primary target for treatment of RA should be a state of clinical remission (Smolen et al. 2010; Felson et al. 2011). There are several definitions for clinical remission based on composite scores (commonly including: tender joint counts [TJCs]/swollen joint counts [SJC], level of acute phase reactants, and assessment of a patient or physician global response). These scores include the Disease Activity Score (DAS), DAS modified to include the 28 diarthroidal joint count (DAS28), Simplified Disease Activity Index (SDAI), Clinical Disease Activity Index (CDAI), and a Boolean definition as recently proposed jointly by American College of Rheumatology (ACR)/European League Against Rheumatism (EULAR) (Felson et al. 2011). The current focus on arresting disease activity results from an understanding that persistent joint inflammation leads to progressive joint destruction manifested by cartilage loss, erosive damage to juxta-articular bone, and resultant functional impairment; that erosive bone changes occur within months of disease onset; and that early and aggressive treatment increases the likelihood of disease control (Klareskog et al. 2009; Colmegna et al. 2012).

Despite a variety of approved agents for RA, complete or sustained disease remission is unusual. Conventional disease-modifying anti-rheumatic drugs (cDMARDs) have been used with some success. Patients often receive 1 or more of these medications (for example, methotrexate [MTX], sulfasalazine, hydroxychloroquine, leflunomide, azathioprine, gold salts, and cyclosporine), typically in combination with low-dose oral or intra-articular glucocorticoids. In addition to cDMARDs, biological agents that block or antagonize critical inflammatory mediators, T cells, or B cells can reduce pain and swelling and provide joint protection against structural damage. The efficacy of these biologics, particularly in combination with MTX, has been shown to have a clinically important effect on the signs and symptoms of RA (Fleischmann 2005); however, a majority of patients do not go into remission or achieve a 50%

improvement in ACR criteria (ACR50) in a clinical trial setting. During treatment with tumor necrosis factor-alpha (TNF- α) antagonists, approximately 30% of patients fail to achieve a 20% improvement in ACR criteria (ACR20; primary failure), and more patients lose efficacy during therapy (secondary failure; Rubbert-Roth and Finckh 2009). Disease progression can still occur even for patients who achieve apparent adequate control of their signs and symptoms with cDMARDs and/or biologic therapies (Klareskog et al. 2009; Rubbert-Roth and Finckh 2009). Accordingly, a significant unmet need remains for more effective and better tolerated treatments for RA.

Members of the Janus kinase (JAK) family of protein tyrosine kinases (JAK1, JAK2, JAK3, and tyrosine kinase 2 [TYK2]) play an important role in signal transduction following cytokine and growth factor binding to their receptors. Aberrant production of cytokines and growth factors has been associated with a number of chronic inflammatory conditions, including RA. Many of the pro-inflammatory cytokines implicated in the pathogenesis of RA, including interleukin-6 (IL-6) and interferon-gamma, use cell signaling that involves the JAK/signal transducers and activators of transcription (STAT) pathways. Inhibition of JAK-STAT signaling can target multiple RA-associated cytokine pathways and thereby reduce inflammation, cellular activation, and proliferation of key immune cells as demonstrated with other JAK inhibitors (Williams et al. 2008; Kremer et al. 2009).

Baricitinib is an orally available, selective JAK inhibitor with excellent potency and selectivity for JAK1 (inhibitory concentration of 50% [IC₅₀] = 5.9 nM) and JAK2 (IC₅₀ = 5.7 nM) and less potency for JAK3 (IC₅₀ \geq 400 nM) or TYK2 (IC₅₀ = 53 nM) (Fridman et al. 2010). Baricitinib is being developed for treatment of patients with moderately to severely active RA who are either intolerant to MTX treatment or who have had an inadequate response to disease-modifying antirheumatic drugs (DMARDs), either conventional or biologic. One completed and 2 currently ongoing Phase 2 studies have evaluated the clinical utility of baricitinib as treatment for patients with active RA. In the completed study (I4V-MC-JADC [JADC], conducted by Incyte Corporation), administration of baricitinib at doses of 4, 7, or 10 mg daily for up to 24 weeks resulted in improved signs and symptoms in patients with an inadequate response to DMARDs, including biologics. A relatively flat dose response was observed for the efficacy parameters, with the 4-mg dose performing as well as the 7- and 10-mg doses. The proportions of patients who were ACR responders generally decreased in the off-treatment follow-up period (from Week 24 to Week 28).

In an ongoing Phase 2 study (I4V-MC-JADA [JADA]), administration of baricitinib at doses of 1, 2, 4, and 8 mg daily for up to 12 weeks with continued administration of the 2-, 4-, and 8-mg doses for up to 24 weeks resulted in improved signs and symptoms in patients with an inadequate response to MTX. The observed treatment effect in the 4- and 8-mg dose groups was similar, confirming the flat dose response observed in Study JADC. The treatment effect in these dose groups was larger than that observed in the 1- and 2-mg dose groups. The 1- and 2-mg dose groups offered some clinical effectiveness relative to placebo treatment. Patients completing 24 weeks of treatment could enter an extension study that is currently ongoing. A third Phase 2 study (I4V-JE-JADN) is ongoing in Japan. The safety profile for baricitinib has been informed

by results from nonclinical and clinical studies evaluating a wide range of doses (up to 20 mg once daily [QD]). There are a number of potential risks recognized for baricitinib that will be followed carefully in the Phase 3 development program. Important laboratory monitoring guidance for baricitinib include decreased hemoglobin; increased total cholesterol, high-density lipoprotein (HDL), low-density lipoprotein (LDL), and triglycerides; and decreased neutrophils and other phagocytic white cell lines (within the normal reference range). More information about the known and expected benefits, risks, and reasonably anticipated adverse events (AEs) may be found in the investigator's brochure (IB). Information on AEs expected to be related to the study drug may be found in Section 7 (Development Core Safety Information) of the IB. Information on serious AEs (SAEs) expected in the study population independent of drug exposure will be assessed by the sponsor in aggregate periodically during the course of the study and may be found in Section 6 (Effects in Humans) of the IB.

There are 6 Phase 3 Studies including I4V-CR-JAGS. The remaining 5 Phase 3 studies are described below. Study I4V-MC-JADV (JADV) is investigating the efficacy and safety of baricitinib in patients with active RA. Study JADV is a double-blind, placebo- and active-controlled study investigating the efficacy and safety of baricitinib (4 mg QD) in patients with active RA who have had an inadequate response to MTX (MTX-IR). The study is 52 weeks in duration and includes assessment of structural joint damage. Study I4V-MC-JADZ is a double-blind, active-controlled study investigating the efficacy and safety of baricitinib (4 mg QD), administered as monotherapy or in combination with MTX, in patients with early RA who have had limited or no treatment with DMARDs. Methotrexate, administered as monotherapy and titrated to the highest tolerated dose, serves as the active comparator. Study I4V-MC-JADX (JADX) is a double-blind, placebo-controlled study investigating the efficacy and safety of baricitinib (2 mg QD or 4 mg QD) administered with background cDMARD(s) in patients with active RA who have had an inadequate response to conventional DMARDs. Study I4V-MC-JADW (JADW) is a double-blind, placebo-controlled study investigating the efficacy and safety of baricitinib (2 mg QD or 4 mg QD) administered with background cDMARD(s) in patients with active RA who have had an inadequate response to biologic DMARDs. Patients completing any of the above studies are eligible for enrollment in Study I4V-MC-JADY (JADY), a long-term extension study investigating the safety of baricitinib with prolonged administration (up to 2 additional years of treatment).

Given the continuing unmet medical need in patients with moderately to severely active RA, the efficacy of baricitinib demonstrated in Phase 2 studies, and the acceptable safety profile for baricitinib observed through the current stage of development, Phase 3 testing is appropriate.

Study I4V-CR-JAGS (JAGS) is a 52-week, double-blind, placebo-controlled study to assess the efficacy and safety of baricitinib (4 mg QD) in patients with moderately to severely active RA who have had an inadequate response to MTX (MTX-IR).

6. Objectives

6.1. Primary Objective

The primary objective of the study is to determine whether baricitinib is superior to placebo in the treatment of patients with moderately to severely active RA despite methotrexate treatment (that is, MTX-IR), as assessed by the proportion of patients achieving ACR20 at Week 12.

6.2. Secondary Objectives

The secondary objectives of the study are to evaluate the efficacy of baricitinib versus placebo as assessed by:

- change from baseline to Week 12 in Health Assessment Questionnaire-Disability Index (HAQ-DI) score
- change from baseline to Week 12 in DAS28-high-sensitivity C-reactive protein (hsCRP)
- proportion of patients achieving an SDAI score ≤ 3.3 at Week 12
- mean duration of morning joint stiffness in the 7 days prior to Week 12 as collected in diaries
- mean severity of morning joint stiffness in the 7 days prior to Week 12 as collected in diaries
- mean Worst Tiredness numeric rating scale (NRS) in the 7 days prior to Week 12 as collected in diaries
- mean Worst Joint Pain NRS in the 7 days prior to Week 12 as collected in diaries.

6.3. Exploratory Objective

The exploratory objectives in the study include efficacy evaluation of baricitinib as assessed by:

- change from baseline at Weeks 16, 24 and 52 in structural joint damage as measured by modified Total Sharp Score (mTSS; van der Heijde 2000)
- proportion of patients with mTSS change ≤ 0 from baseline at Weeks 16, 24, and 52
- change from baseline to Weeks 16, 24, and 52 in joint space narrowing and bone erosion scores
- proportion of patients achieving DAS28-hsCRP ≤ 3.2 and DAS28-hsCRP < 2.6 at Weeks 12, 24, and 52
- proportion of patients achieving HAQ-DI improvement ≥ 0.22 and ≥ 0.3 at Weeks 12, 24, and 52
- proportion of patients achieving ACR20 at Week 24 and Week 52
- proportion of patients achieving ACR50 at Weeks 12, 24, and 52
- proportion of patients achieving 70% improvement in ACR criteria (ACR70) at Weeks 12, 24, and 52
- percentage change from baseline to Weeks 12, 24, and 52 in individual components of the ACR Core Set
- change from baseline through Week 52 in DAS28-hsCRP
- change from baseline through Week 52 in DAS28-erythrocyte sedimentation rate (ESR)

- proportion of patients achieving DAS28-ESR ≤ 3.2 and DAS28-ESR < 2.6 at Weeks 12, 24, and 52
- change from baseline to Week 24 and Week 52 in HAQ-DI score
- change from baseline to Weeks 12, 24, and 52 in CDAI score
- change from baseline to Weeks 12, 24, and 52 in SDAI score
- proportion of patients achieving a CDAI score ≤ 2.8 at Weeks 12, 24, and 52
- proportion of patients achieving an SDAI score ≤ 3.3 at 24, and Week 52
- proportion of patients achieving ACR/EULAR remission according to the Boolean-based definition at Weeks 12, 24, and 52
- proportion of patients achieving moderate and good EULAR responses based on the DAS28 at Weeks 12, 24, and 52
- evaluation of hybrid ACR (bounded) response at Weeks 12, 24 and 52
- change from baseline to Week 12, Week 24, and Week 52 in Functional Assessment of Chronic Illness Therapy Fatigue (FACIT-F) scale scores
- change from baseline to Weeks 12, 24 and 52 in European Quality of Life-5 Dimensions-5 Level (EQ-5D-5L) scores
- change from baseline to Weeks 12, 24 and 52 in Work Productivity and Activity Impairment-Rheumatoid Arthritis (WPAI-RA) scores
- evaluation of healthcare resource utilization through Week 52
- change from baseline to Weeks 12, 24 and 52 in duration of morning joint stiffness assessed at the visit using an electronic patient-reported outcomes (ePRO) tablet
- change from baseline to Weeks 12, 24 and 52 in Worst Tiredness NRS assessed at the visit using an ePRO tablet
- change from baseline to Weeks 12, 24 and 52 in Worst Joint Pain NRS assessed at the visit using an ePRO tablet.

7. Investigational Plan

7.1. Summary of Study Design

Study JAGS will be a 52-week, Phase 3, multicenter, randomized, double-blind, placebo-controlled, parallel-group study comparing the efficacy of baricitinib versus placebo in patients with moderately to severely active RA who have had an insufficient response to MTX and who have never been treated with a biologic DMARD (that is, MTX-IR patients).

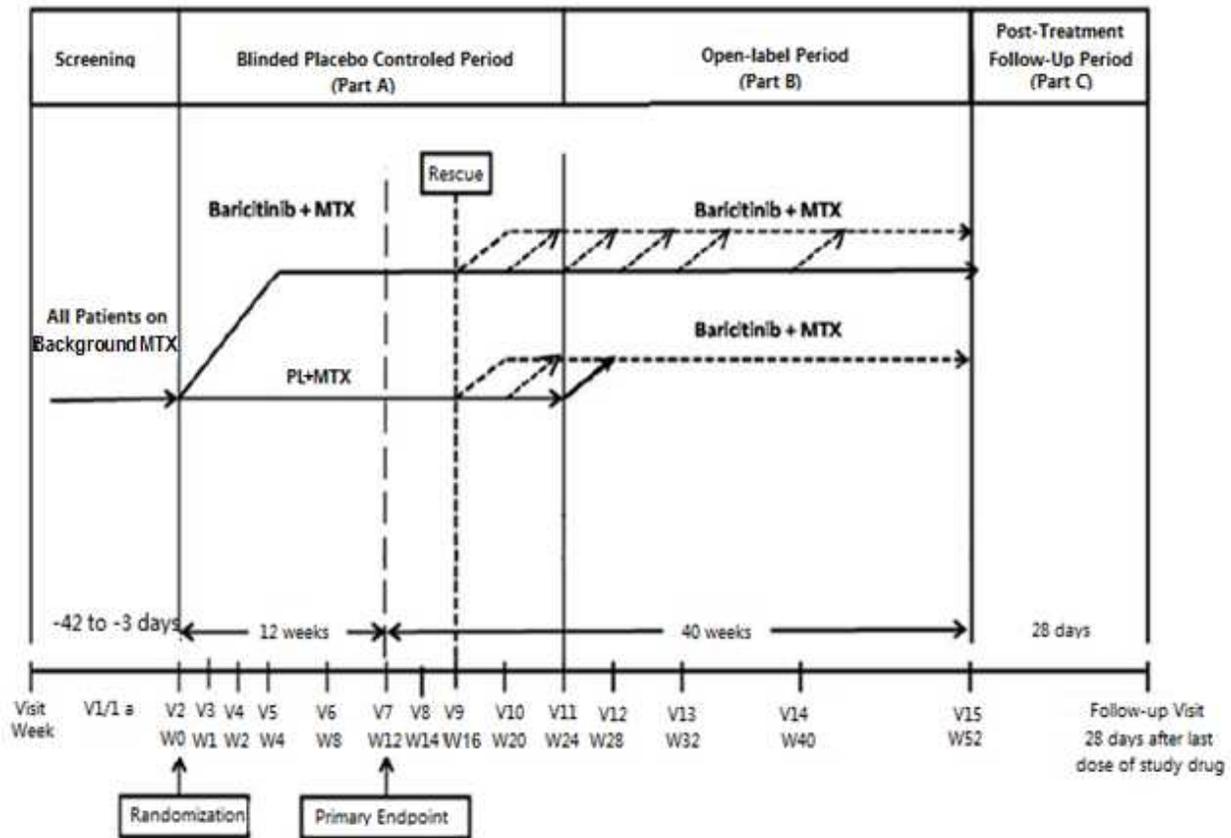
Baricitinib will be administered as a 4-mg tablet QD. The dose for patients with renal impairment, defined as estimated glomerular filtration rate (eGFR) <60 mL/min/1.73 m², will be 2 mg baricitinib QD. Patients not assigned to the baricitinib treatment arm will receive either a 4-mg or 2-mg matching placebo tablet.

Planned enrollment is approximately 288 randomized patients: 144 to receive baricitinib, and 144 to receive placebo. After the Week 52 study visit, eligible patients may proceed to a separate extension study (Study JADY) lasting for up to 49 months or to the posttreatment follow-up period of this study (Part C).

Study JAGS will consist of 4 parts:

- Screening: Screening period lasting from 3 to 42 days prior to Visit 2 (Week 0)
- Part A: double-blind, placebo-controlled period from Week 0 through Week 24
- Part B: open-label period from the end of Week 24 through Week 52
- Part C: posttreatment follow-up period

Figure JAGS.7.1 illustrates the study design.



Abbreviations: MTX = methotrexate, V = visit, W = week.

Note: Diagonal dashed arrows indicate an option for rescue therapy. The diagonal solid arrow indicates a mandatory change to baricitinib treatment (at Week 24 for placebo-treated patients).

Figure JAGS.7.1. Illustration of study design for Clinical Protocol I4V-CR-JAGS.

7.1.1. Screening Period

This is a screening period of not less than 3 days and not more than 42 days. In exceptional circumstances, the screening window can be extended after consultation with the Sponsor. Patients will be screened for study eligibility at Visit 1 and will be randomized at Visit 2 to continue in Part A.

7.1.2. Part A and Part B: 24-Week Double-Blind Placebo- Controlled, (Part A) Period and 28-Week Open-Label (Part B) Period

Patients on a background of MTX will be randomized at a 1:1 ratio to 1 of 2 treatment arms (baricitinib or placebo):

- Patients randomized to baricitinib will be administered a 4-mg tablet QD (or 2 mg baricitinib tablet QD if eGFR <60 mL/min/1.73 m²) starting at Week 0 (Visit 2) through Week 52 (Visit 15).

- Patients randomized to placebo will be administered a placebo tablet QD starting at Week 0 (Visit 2) through Week 24 (the morning of Visit 11). At Week 24, patients will be administered baricitinib 4 mg QD (or 2 mg baricitinib QD if eGFR <60 mL/min/1.73 m²) through Week 52 (Visit 15). Patients will be eligible for rescue treatment beginning at Week 16 (Visit 9). Details of the rescue are listed in Section 9.5.1.

Patients will remain on background MTX, nonsteroidal anti-inflammatory drugs (NSAIDs), analgesics, and/or corticosteroids (see Section 9.8 for guidelines on adjustments to concomitant therapy). Patients will take their assigned doses of study drugs as directed. The primary efficacy endpoint will be at Week 12, and the final visit during Part A will be at Week 24. The final visit during Part B will be at Week 52 (Visit 15). The structural joint damage endpoints will be at Weeks 16, 24, and 52. All patients who complete the Week 52 study visit will be eligible for inclusion in extension Study JADY.

7.1.3. Part C: Posttreatment Follow-Up Period

Patients who are not enrolled in the extension Study JADY will have a Follow-Up Visit (Visit 801) approximately 28 days after the last dose of study drug.

Patients who discontinue from the study prior to Week 52 for any reason will have an Early Termination Visit performed, including x-rays. The Early Termination Visit should be performed as soon as logistically possible because this visit includes assessment of safety. Infrequently, patient and investigator availability may be such that the Early Termination Visit and the Follow-Up Visit may occur on or about the same date. In this instance, the visits may be combined and should occur approximately 28 days after the last dose of study drug. All activities required for the Early Termination Visit should be completed according to the Study Schedule ([Attachment 1](#)).

7.1.4. Study Extensions

Patients who complete this study through Visit 15 will be eligible to participate in Study JADY, if enrollment criteria for Study JADY are met.

7.2. Discussion of Design and Control

Based on efficacy, safety, and pharmacokinetic (PK) data from the Phase 2 studies (JADC and JADA), Eli Lilly and Company (Lilly) has selected a single dose of 4 mg QD baricitinib (with dose adjustment for renal dysfunction) for evaluation in this study (see Section 9.4 for rationale).

The Week 12 primary efficacy endpoint was chosen based on the results of Studies JADC and JADA. In these studies, efficacy as measured by ACR20 response was observed as early as 2 weeks after initial administration of baricitinib followed by near-maximum efficacy response achieved by Week 12. The 52-week treatment period will allow evaluation of the long-term effect of baricitinib on structural joint damage.

All patients will be required to be on background MTX treatment at baseline. Other background therapies, including NSAIDs and low dose oral corticosteroids, are permitted during the study

for patients who are on stable doses of these treatments at baseline. Patients should remain on the same dose of MTX throughout the study; however, the MTX dose may be adjusted for safety reasons.

To prevent potential unblinding due to observed efficacy or laboratory changes, a “dual assessor” approach will be used to evaluate efficacy and safety. The Joint Assessor (or designee) should be a rheumatologist or skilled arthritis assessor. The Joint Assessor will be responsible for completing the joint counts. To ensure consistent joint evaluation throughout the trial, individual patients should be evaluated by the same Joint Assessor for all study visits. The Joint Assessor must not access or discuss with the patient the patient-reported assessments, Physician’s Global Assessment of Disease Activity (visual analog scale [VAS]), and safety assessments.

The Safety Assessor (or designee) should be a rheumatologist (or medically qualified physician) and will have access to both safety and efficacy data. The Safety Assessor may be the Principal Investigator (PI). The Safety Assessor will be responsible for completing the Physician’s Global Assessment of Disease Activity (VAS). To ensure consistent Physician’s Global Assessment of Disease Activity (VAS) throughout the trial, the instrument should be evaluated by the same physician at all study visits. The Safety Assessor will have access to source documents, laboratory results, and case report forms (CRFs) and will be responsible for making treatment decisions based on a patient’s clinical response and laboratory parameters.

All patients in Study JAGS will be offered rescue therapy starting at Week 16 in line with precedent from recent clinical trials. Patients who are determined to be nonresponders to their originally assigned treatment will be reassigned to treatment with baricitinib at Week 16. Patients originally assigned to baricitinib will be rescued to the same dose of baricitinib to maintain the study blind. Patients not rescued at Week 16 may be rescued to treatment with baricitinib at the discretion of the investigator anytime thereafter. Patients initially randomized to placebo and not rescued will be reassigned to receive baricitinib at Week 24 to limit exposure to inactive treatment. Patients not experiencing improvement in signs and symptoms following approximately 4 weeks of rescue treatment should be discontinued from the study. Patients may be rescued only once. See Section [9.5.1](#) for details of rescue treatment.

8. Study Population

The study population will be comprised of patients diagnosed with RA with active disease who have failed to adequately respond to MTX. Study investigator(s) will review patient records and screening test results to determine that the patient meets all inclusion and exclusion criteria to qualify for participation in the study.

Individuals who do not meet the criteria for participation in this study (screen failure) may be re-screened. Individuals may be re-screened up to 1 time. The interval between re-screenings should be approximately 30 days. Each time re-screening is performed, the individual must sign a new informed consent form (ICF) and will be assigned a new identification number.

8.1. Entry Criteria

Patients may enter into the study once they have signed the ICF. Patients may be assessed for study entry on the basis of readily available information (that is, information that does not require informed consent from the subject).

8.1.1. Inclusion Criteria

Patients are eligible for entry into the study (ie, eligible to sign consent) only if they meet **all** of the following criteria:

- [1] are at least 18 years of age
- [2] have a diagnosis of adult-onset RA as defined by the ACR/EULAR 2010 Criteria for the Classification of RA (Aletaha et al. 2010)
- [3] have moderately to severely active RA defined as the presence of at least 6/68 tender joints and at least 6/66 swollen joints
 - a. If significant surgical treatment of a joint has been performed, that joint cannot be counted in the TJC and SJC for entry or enrollment purposes.
- [4] have a C-reactive protein (CRP) (or hsCRP) measurement \geq 6mg/L based on the most recent data (if available)
- [5] have had regular use of MTX for at least the 12 weeks prior to study entry at a dose that, in accordance with local clinical practice, is considered acceptable to adequately assess clinical response. The dose of MTX must have been a stable, unchanging oral dose of 7.5 to 25 mg/week (or the equivalent injectable dose) for at least the 8 weeks prior to study entry. The dose of MTX is expected to remain stable throughout the study and may be adjusted only for safety reasons.
 - a. For patients entering the trial on MTX doses $<$ 15 mg/week, there must be clear documentation in the medical record that higher doses of MTX were not tolerated or that the dose of MTX is the highest acceptable dose based on local clinical practice guidelines.

- b. Local standard of care should be followed for concomitant administration of folic acid.

[6] are able to read, understand, and give written informed consent.

8.1.2. Exclusion Criteria

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

- [7] are currently receiving corticosteroids at doses >10 mg of prednisone per day (or equivalent) **or** have been receiving an unstable dosing regimen of corticosteroids within 2 weeks of study entry or within 6 weeks of planned randomization
- [8] have started treatment with NSAIDs (for which the NSAID use is intended for treatment of signs and symptoms of RA) within 2 weeks of study entry or within 6 weeks of planned randomization **or** have been receiving an unstable dosing regimen of NSAIDs within 2 weeks of study entry or within 6 weeks of planned randomization
- [9] are currently receiving concomitant treatment with MTX, hydroxychloroquine, **and** sulfasalazine or combination of any 3 cDMARDs
- [10] are currently receiving or have received cDMARDs (for example, gold salts, cyclosporine, azathioprine, or any other immunosuppressives) other than MTX, hydroxychloroquine (up to 400 mg/day), or sulfasalazine (up to 3000 mg/day) within 8 weeks prior to study entry
 - a. Doses of hydroxychloroquine **or** sulfasalazine should be stable for at least 8 weeks prior to study entry; if either has been recently discontinued, the patient must not have taken any dose within 4 weeks prior to study entry.
 - b. Immunosuppression related to organ transplantation is not permitted.
- [11] have received leflunomide in the 12 weeks prior to study entry (or within 4 weeks prior to study entry if the standard 11 days of cholestyramine is used to washout leflunomide)
- [12] have started a new physiotherapy treatment for RA in the 2 weeks prior to study entry
- [13] have ever received any biologic DMARD (such as tumor necrosis factor (TNF), interleukin-1, interleukin-6 (IL-6), or T-cell- or B-cell-targeted therapies)
- [14] have received interferon therapy (such as Roferon-A, Intron-A, Rebetrone, Alferon-N, Peg-Intron, Avonex, Betaseron, Infergen, Actimmune, Pegasys) within 4 weeks prior to study entry or are anticipated to require interferon therapy during the study

- [15] have received any parenteral corticosteroid administered by intramuscular or intravenous injection within 2 weeks prior to study entry or within 6 weeks prior to planned randomization or are anticipated to require parenteral injection of corticosteroids during the study
- [16] have had 3 or more joints injected with intraarticular corticosteroids or hyaluronic acid within 2 weeks prior to study entry or within 6 weeks prior to planned randomization
 - a. Joints injected with intraarticular corticosteroids or hyaluronic acid within 2 weeks prior to study entry or within 6 weeks prior to planned randomization cannot be counted in the TJC and SJC for entry or enrollment purposes.
- [17] have active fibromyalgia that, in the investigator's opinion, would make it difficult to appropriately assess RA activity for the purposes of this study
- [18] have a diagnosis of any systemic inflammatory condition other than RA such as, but not limited to, juvenile chronic arthritis, spondyloarthropathy, Crohn's disease, ulcerative colitis, psoriatic arthritis, active vasculitis or gout
 - a. Patients with secondary Sjögren's syndrome are not excluded.
- [19] have a diagnosis of Felty's syndrome
- [20] have had any major surgery within 8 weeks prior to study entry or will require major surgery during the study that, in the opinion of the investigator in consultation with Lilly or its designee, would pose an unacceptable risk to the patient
- [21] have experienced any of the following within 12 weeks of study entry: myocardial infarction, unstable ischemic heart disease, stroke, or New York Heart Association Stage IV heart failure
- [22] 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 a risk when taking investigational product or could interfere with the interpretation of data
- [23] are largely or wholly incapacitated permitting little or no self-care, such as being bedridden or confined to a wheelchair
- [24] have an eGFR based on the most recent available serum creatinine using the Modification of Diet in Renal Disease (MDRD) method of $<40 \text{ mL/min/1.73 m}^2$
- [25] have a history of chronic liver disease with the most recent available aspartate aminotransferase (AST) or alanine aminotransferase (ALT) >1.5 times the ULN or the most recent available total bilirubin ≥ 1.5 times the ULN

- [26] have a history of lymphoproliferative disease; or have signs or symptoms suggestive of possible lymphoproliferative disease, including lymphadenopathy or splenomegaly; or have active primary or recurrent malignant disease; or have been in remission from clinically significant malignancy for <5 years
- a. Patients with cervical carcinoma in situ that has been 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.
- [27] have been exposed to a live vaccine within 12 weeks prior to planned randomization or are expected to need/receive a live vaccine during the course of the study (with the exception of herpes zoster vaccination)
- a. All patients who have not received the herpes zoster vaccine at study entry will be encouraged to do so prior to randomization; vaccination must occur >30 days prior to randomization and start of study drug. Patients will be excluded if they were exposed to herpes zoster vaccination within 30 days of planned randomization.
 - b. Investigators should review the vaccination status of their patients and follow the local guidelines for adult vaccination with nonlive vaccines intended to prevent infectious disease prior to entering patients into the study.
- [28] have a current or recent (<30 days prior to study entry) clinically serious viral, bacterial, fungal, or parasitic infection
- [29] have had symptomatic herpes zoster infection within 12 weeks prior to study entry
- [30] have a history of disseminated/complicated herpes zoster (for example, multidermatomal involvement, ophthalmic zoster, central nervous system (CNS) involvement, or postherpetic neuralgia)
- [31] are immunocompromised and, in the opinion of the investigator, are at an unacceptable risk for participating in the study
- [32] have a history of active hepatitis B virus (HBV), hepatitis C virus (HCV), or human immunodeficiency virus (HIV)
- [33] have had household contact with a person with active tuberculosis (TB) and did not receive appropriate and documented prophylaxis for TB
- [34] have evidence of active TB or have previously had evidence of active TB and did not receive appropriate and documented treatment
- [35] are pregnant or nursing at the time of study entry

- [36] are females of childbearing potential who do not agree to use 2 forms of highly effective birth control when engaging in intercourse while enrolled in the study and for at least 28 days following the last dose of orally administered investigational product
- a. Females of nonchildbearing potential are defined as women ≥ 60 years of age, women ≥ 40 and < 60 years of age who have had a cessation of menses for at least 12 months, or women who are congenitally or surgically sterile (that is, have had a hysterectomy or bilateral oophorectomy or tubal ligation).
 - b. The following birth control methods are considered highly effective (the patient should choose 2 to be used with their partner):
 - oral, injectable, or implanted hormonal contraceptives
 - condom with a spermicidal foam, gel, film, cream, or suppository
 - occlusive cap (diaphragm or cervical/vault caps) with a spermicidal foam, gel, film, cream, or suppository
 - intrauterine device
 - intrauterine system (for example, progestin-releasing coil)
 - vasectomized male (with appropriate postvasectomy documentation of the absence of sperm in the ejaculate)
- [37] are males who do not agree to use 2 forms of highly effective birth control (see above) while engaging in sexual intercourse with female partners of childbearing potential while enrolled in the study and for at least 28 days following the last dose of orally administered investigational product
- [38] have donated > 500 mL of blood within 30 days prior to study entry or intend to donate blood during the course of the study
- [39] have a history of chronic alcohol abuse, intravenous drug abuse, or other illicit drug abuse within the 2 years prior to study entry
- [40] have previously been randomized in this study or any other study investigating baricitinib
- [41] are unable or unwilling to make themselves available for the duration of the study and/or are unwilling to follow study restrictions and procedures
- [42] have received prior treatment with an oral JAK inhibitor
- [43] 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
- [44] are Lilly or Incyte employees or either's designee

- [45] are currently enrolled in or have discontinued within 30 days of study entry from a clinical trial involving an investigational product or nonapproved use of a drug or device or are concurrently enrolled in any other type of medical research judged not to be scientifically or medically compatible with this study.

8.2. Enrollment Criteria

8.2.1. Inclusion Criteria

Entered patients are eligible for enrollment into the study (that is, eligible for randomization) only if they continue to meet **all** entry criteria (see above) at the time of randomization and meet the following enrollment criteria:

- [46] have moderately to severely active RA defined as the presence of at least 6/68 tender joints and at least 6/66 swollen joints assessed at Visit 1 and Visit 2 at the time of randomization
- [47] have an hsCRP measurement $\geq 6\text{mg/L}$ based on Visit 1 laboratory results by central laboratory testing

The hsCRP measurement may be repeated once within approximately 2 weeks of the initial value, and the value resulting from repeat testing may be accepted for enrollment eligibility if it meets the eligibility criterion.

- [48] have at least 1 joint erosion in hand, wrist, or foot joints based on radiographic interpretation by the central reader and be rheumatoid factor or anticyclic citrullinated peptide (anti-CCP) antibody positive; **or** have at least 3 joint erosions in hand, wrist, or foot joints based on radiographic interpretation by the central reader regardless of rheumatoid factor or anti-CCP antibody status (radiographic images acquired within 4 weeks of study entry may be submitted to the central reader to confirm eligibility). If a patient undergoes rescreening, radiographic images acquired as part of initial screening and within 10 weeks of randomization may be used.

8.2.2. Exclusion Criteria

Entered patients are ineligible for enrollment (that is, ineligible for randomization) and should be discontinued from the study if they meet any of the following criteria:

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

The TSH test may be repeated once within approximately 2 weeks of the initial value, and the value resulting from repeat testing may be accepted for enrollment eligibility if it meets the eligibility criterion.

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

- AST or ALT >1.5 times the ULN
- total bilirubin \geq 1.5 times the ULN
- hemoglobin <10.0 g/dL (100.0 g/L)
- total white blood cell count <2500 cells/ μ L
- neutropenia (absolute neutrophil count <1200 cells/ μ L)
- lymphopenia (lymphocyte count <750 cells/ μ L)
- thrombocytopenia (platelets <100,000/ μ L)
- eGFR <40 mL/min/1.73 m²

In the case of any of the aforementioned laboratory abnormalities, the tests may be repeated once within approximately 2 weeks of the initial values, and values resulting from repeat testing may be accepted for enrollment eligibility if they meet the eligibility criterion.

[51] have screening electrocardiogram (ECG) abnormalities that, in the opinion of the investigator or the sponsor, are clinically significant and indicate an unacceptable risk for the patient's participation in the study (for example, Fridericia's corrected QT interval >500 msec)

[52] have symptomatic herpes simplex at the time of study enrollment

[53] have evidence of active TB as documented by a positive purified protein derivative (PPD) test (\geq 5-mm induration between approximately 2 and 3 days after application, regardless of vaccination history), medical history, clinical symptoms, and abnormal chest x-ray at screening

- a. The QuantiFERON[®]-TB Gold or the T-SPOT[®].TB test (if available) 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.

[54] have evidence of latent TB (as documented by a positive PPD, no clinical symptoms consistent with active TB, and a normal chest x-ray at screening) unless patients complete at least 4 weeks of appropriate treatment prior to randomization and agree to complete the remainder of treatment while in the trial

- a. 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 or T-SPOT.TB_test (if available). If the test results are not negative, the patient will be considered to have latent TB (for purposes of this study).
- b. The QuantiFERON-TB Gold or T-SPOT.TB_test (if available) 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).
- c. Exceptions include patients with a history of active or latent TB who have documented evidence of appropriate treatment.

[55] have a positive test for HBV defined as:

- a. positive for hepatitis B surface antigen, **or**
- b. positive for anti-hepatitis B core antibody but negative for hepatitis B surface antibody (HBsAb), **or**
- c. for patients enrolled in other countries, if required, positive for anti-HBsAb and positive for HBV deoxyribonucleic acid (DNA).

If any of the HBV tests have an indeterminate result, confirmatory testing will be performed by an alternate method.

[56] have HCV (positive for anti-hepatitis C antibody with confirmed presence of HCV)

[57] have evidence of human HIV infection and/or positive HIV antibodies

[58] Are women ≥ 40 and < 60 years of age who have had a cessation of menses for at least 12 months, have a follicle-stimulating hormone (FSH) value < 40 mIU/mL, and do not agree to use 2 forms of highly effective methods of birth control when engaging in intercourse while enrolled in the study and for at least 28 days following the last dose of orally administered investigational product, unless they are congenitally or surgically sterile (that is, have had a hysterectomy, bilateral oophorectomy, or tubal ligation)

[59] Have received treatment with traditional medicine that has known immunosuppressant or immunomodulatory effects, such as Pafulin or Leigongteng (Common Threewingnut Root) within 30 days before Visit 2

[60] Have a current or recent (during screening period) clinically serious viral, bacterial, fungal, or parasitic infection

If patients are entered into the study while taking medications that require discontinuation before randomization, those patients should only be asked to discontinue these medications after it has been determined that they will be eligible for randomization.

Patients who are entered into the study but do not meet enrollment criteria should be discontinued from the study. These patients can be re-entered into the trial (ie, give repeat consent) if the investigator believes that the patient might meet enrollment criteria at a future date. The investigator should wait approximately 1 month before re-entering a patient into the study.

8.2.3. Rationale for Exclusion of Certain Study Candidates

The rationale for the **entry** Exclusion Criteria is as follows: Exclusion Criteria [7] to [16] exclude individuals who are taking or who may take RA medications or treatments that interfere with the ability to assess the safety and efficacy of baricitinib. Exclusion Criteria [17] to [26] exclude individuals with previous or concomitant medical conditions that increase the risk for their participation in the study. Exclusion Criteria [27] to [34] exclude individuals who are at an increased risk for infections or infectious complications. Exclusion Criteria [35] to [37] exclude individuals who are pregnant, breastfeeding, at risk for becoming pregnant, or at risk for impregnating their partner during the study. Exclusion Criteria [38] to [45] exclude individuals who may not be compliant with study-related procedures or whose participation in the study may introduce bias.

The rationale for the **enrollment** Exclusion Criteria is as follows: Exclusion Criteria [49] to [52] exclude individuals with concomitant medical conditions that increase the risk for their participation in the study. Exclusion Criteria [53] to [57] and [60] exclude individuals who are at an increased risk for infections or infectious complications. Exclusion Criterion [58] excludes individuals who are at risk for becoming pregnant during the study. Exclusion Criterion [59] excludes individuals who are taking or who may take traditional medicines that may interfere with the ability to assess the safety and efficacy of baricitinib.

8.3. Discontinuations

8.3.1. Discontinuation of Patients

8.3.1.1. Interruption of Study Drug

In some circumstances it may be necessary to temporarily interrupt treatment as a result of AEs or abnormal laboratory values as described in [Table JAGS.8.1](#) that may have an unclear relationship to study drug. Except in cases of emergency, it is recommended that the investigator consult with Lilly (or its designee) before temporarily interrupting therapy.

The investigator must obtain approval from Lilly (or its designee) before restarting study drug that was temporarily discontinued because of an AE or abnormal laboratory value. Study drug must be held in the following situations involving laboratory abnormalities and may be resumed as noted in [Table JAGS.8.1](#).

Table JAGS.8.1. Criteria for Temporary Discontinuation of Investigational Product

Hold Investigational Product if the Following Laboratory Test Results Occur:	Investigational Product May Be Resumed after Approval from Lilly (or its Designee) and When:
WBC count <2000 cells/ μ L	WBC count \geq 2500 cells/ μ L
ANC <1000 cells/ μ L	ANC >2000 cells/ μ L
Lymphocyte count <500 cells/ μ L	Lymphocyte count \geq 750 cells/ μ L
Platelet count <75,000/ μ L	Platelet count \geq 100,000/ μ L
eGFR <40 mL/min/1.73 m ² (from serum creatinine) for patients receiving the baricitinib 4 mg QD dose	eGFR \geq 50 mL/min/1.73 m ²
eGFR <30 mL/min/1.73 m ² (from serum creatinine) for patients receiving the baricitinib 2 mg QD dose	eGFR \geq 40 mL/min/1.73 m ²
ALT or AST >5 times ULN	ALT and AST return to <2 times ULN and investigational product is not considered to be the cause of enzyme elevation
Hemoglobin <8 g/dL	Hemoglobin \geq 10 g/dL

Abbreviations: ALT = alanine aminotransferase; ANC = absolute neutrophil count; AST = aspartate aminotransferase; eGFR = estimated glomerular filtration rate; min = minute; QD = once daily; ULN = upper limit of normal; WBC = white blood cell.

8.3.1.2. Discontinuation of Study Drug

Any patient who is permanently discontinued from study drug for an AE or abnormal laboratory result should have the AE or abnormal laboratory value reported as an SAE (reported as Other Reason Serious). If any of the criteria listed in Section 8.3.1.1 above recur after study drug is restarted, the patient should be permanently discontinued from study drug. In addition, patients will be permanently discontinued from study drug if they experience any of the criteria listed in [Table JAGS.8.2](#).

Table JAGS.8.2. Criteria for Permanent Discontinuation of Investigational Product

Permanently Discontinue Investigational Product if any of the Following are Observed:
ALT or AST >8 times the ULN
ALT or AST >5 times the ULN persisting for more than 2 weeks after temporary interruption of investigational product
ALT or AST >3 times the ULN and total bilirubin level >2 times the ULN
ALT or AST >3 times the ULN with the appearance of fatigue, nausea, vomiting, right upper quadrant pain or tenderness, fever, rash, and/or eosinophilia (>5%)
WBC count <1000 cells/ μ L
Absolute neutrophil count <500 cells/ μ L
Lymphocyte count <200cells/ μ L
Hemoglobin <6.5 g/dL
Symptomatic herpes zoster
Pregnancy
Malignancy
HBV DNA \geq 29 IU/mL ^a
Severe infection that, in the opinion of the investigator, merits the investigational product being discontinued

Abbreviations: ALT = alanine aminotransferase; AST = aspartate aminotransferase; ULN = upper limit of normal; WBC = white blood cell.

^a. If an HBV DNA result of “target detected” 29 IU/mL or greater, then the patient should be referred to a hepatology specialist immediately. In selected cases, investigators may temporarily continue study drug 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 evaluation of individual patient risks and benefits. Refer to Section 10.3.2.2 for additional instruction on HBV DNA monitoring.

8.3.1.3. Discontinuation from the Study

The criteria for enrollment must be followed explicitly. If a patient who does not meet enrollment criteria is inadvertently enrolled, the investigator should contact Lilly or its designee to discuss whether the patient should be discontinued from study drug. If the patient is discontinued from study drug, the patient should remain in the study to provide the follow-up data needed for the primary analysis (Week 12) of the entire intention-to-treat (ITT) population.

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

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

- Patient enrolls in any other clinical trial involving an investigational product or enrollment in any other type of medical research judged not to be scientifically or medically compatible with this study.
- Investigator decides that the patient should be withdrawn. If this decision is made because of an intolerable AE or a clinically significant laboratory value, the study drug is to be discontinued and appropriate measures are to be taken. Lilly (or designee) is to be notified immediately.

- Lilly (or designee) decides that the patient should be withdrawn.
- Patient is noncompliant with protocol procedures.
- Patient withdraws consent.
- Investigator or Lilly (or designee), for any reason, stops the study.

Patients who discontinue the study early will have Early Termination procedures and follow-up performed as shown in the study schedule ([Attachment 1](#)).

8.3.2. Discontinuation of Study Sites

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

8.3.3. Discontinuation of the Study

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

9. Treatment

9.1. Treatments Administered

This study involves a comparison of baricitinib 4 mg administered orally QD, and placebo administered orally QD. Patients will continue to take background MTX therapy during the course of the study. [Table JAGS.9.1](#) shows the treatment regimens.

Table JAGS.9.1. Treatment Regimens

Treatment Group	Treatments Administered during Part A	Treatments Administered during Part B (Patients Not Rescued)	Treatments Administered during Part B (after Rescue ^a)
Baricitinib treatment	<ul style="list-style-type: none"> Baricitinib 4 mg^b oral once-daily tablet 	<ul style="list-style-type: none"> Baricitinib 4 mg^b oral once-daily tablet 	<ul style="list-style-type: none"> Baricitinib 4 mg^b oral once-daily tablet
Placebo comparator	<ul style="list-style-type: none"> Baricitinib placebo oral once-daily tablet 	<ul style="list-style-type: none"> Baricitinib 4 mg^b oral once-daily tablet 	<ul style="list-style-type: none"> Baricitinib 4 mg^b oral once-daily tablet

^a See Section 9.5.1 for details of rescue medication.

^b The dose for patients with renal impairment, defined as estimated glomerular filtration rate <60 mL/min/1.73 m², will be 2 mg once daily.

The investigator or his/her designee will be responsible for explaining the correct use of investigational agent(s) to the patient, verifying that instructions are followed properly, maintaining accurate records of investigational product dispensing and collection, and returning all unused medication to Lilly or its designee at the end of the study.

In some cases, sites may destroy the material if, during the investigator site selection, the evaluator has verified and documented that the site has appropriate facilities and written procedures to dispose of clinical trial materials.

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

9.2. Materials and Supplies

Lilly (or 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 PI'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. 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).

9.3. Method of Assignment to Treatment

Patients who meet all criteria for enrollment will be randomized in a 1:1 ratio (baricitinib:placebo) to double-blind treatment at Week 0 (Visit 2). Randomization will be stratified by country and joint erosion status (1-2 joint erosions plus seropositivity versus at least 3 joint erosions). 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 of baricitinib/placebo containing double-blind investigational product to each patient. Each package of clinical trial material will supply sufficient medication for the number of weeks between visits. Site personnel will confirm that they have located the correct package by entering a confirmation number found on the packages of investigational product into the IWRS before dispensing the package to the patient.

9.4. Rationale for Selection of Doses in the Study

Clinical Efficacy Considerations

A single initial dose of 4 mg QD baricitinib in this study was selected based on the efficacy, safety, and PK data from Phase 2 Studies JADC and JADA. In Study JADC, baricitinib doses of 4, 7, and 10 mg QD were studied and found to produce similar degrees of efficacy across the measured parameters at both the Week 12 and Week 24 time points. Study JADA included lower doses of baricitinib (1 and 2 mg QD) to find the minimum clinically effective dose and to confirm the flat dose-response curve in the dose range of 4 to 8 mg QD. In Study JADA, a relatively flat dose-response curve was observed in the 4- and 8-mg QD dose groups at Week 12 across almost all measured efficacy parameters. The 1- and 2-mg QD dose groups are biologically active but did not produce adequate response at Week 12 as assessed by ACR50, ACR70, and low disease activity and remission rates when compared to placebo and the 4- and 8-mg dose groups. All 4 doses were associated with reductions in hsCRP and ESR values over time. Assessment of the Week 12 magnetic resonance imaging (MRI) data suggests the potential for the 4- and 8-mg QD baricitinib doses to have a favorable impact on reducing structural damage given the observed significant improvement in osteitis.

Week 24 data from Study JADA demonstrated persistence of the flat dose response between the 4- and 8-mg QD doses. Over the second 12-week period, continued administration of the 2-mg QD dose did not result in further improvement in the clinically important endpoints of ACR50, ACR70, DAS28-hsCRP ≤ 3.2 , DAS28-hsCRP < 2.6 , and CDAI and SDAI remission rates. Week 24 MRI data further support the potential for the 4- and 8-mg QD doses to have a beneficial effect on prevention of structural damage.

Clinical Safety Considerations

Baricitinib was generally safe and well tolerated in single doses ranging from 1 to 20 mg and in repeat oral doses ranging from 2 to 20 mg. The most commonly reported treatment-emergent

AEs (TEAEs) were in the infections and infestations system organ class (SOC). The most common alterations in laboratory values involved decreases in hemoglobin, hematocrit, total red blood cells, and white blood cells (neutrophils and other white cell lines), and increases in platelet counts, HDL, LDL, total cholesterol, and triglycerides. Patients with RA experienced a mean increase in ALT/AST compared to placebo, with the mean values remaining within the normal reference range and clinically significant increases in ALT/AST occurring uncommonly. Patients with impaired renal function have increased exposure to baricitinib.

Dose Adjustment for Renal Impairment

The therapeutic dose range for baricitinib was identified from concentration-efficacy relationships for ACR20 and DAS28-hsCRP as well as concentration-safety relationships identified for absolute neutrophil counts and hemoglobin. The starting dose for this Phase 3 study will be 4 mg QD. As demonstrated in a study of subjects with normal renal function versus those with mild to end-stage renal disease (Study I4V-MC-JADL), baricitinib exposure increases with decreased renal function. Based on PK simulations of baricitinib exposures for the mild and moderate categories of renal function (stratified as eGFR 60 to <90 mL/min/1.73 m² and eGFR 30 to <60 mL/min/1.73 m², respectively) dose adjustment is not required for patients with eGFR ≥60 mL/min/1.73 m². Patients with eGFR <60 mL/min/1.73 m² will receive a dose of 2 mg QD, which will ensure that exposures are comparable to, but do not exceed, those of the 4-mg QD dose. Lilly intends to enroll patients with eGFR no lower than 40 mL/min/1.73 m² in the Phase 3 trials. Baricitinib will be discontinued if eGFR falls below 30 mL/min/1.73 m². Baricitinib administration to patients with severe renal impairment (eGFR <30 mL/min/1.73 m²) is not intended.

Summary

The clinical data known to date suggest a dosing strategy using a 4-mg QD dose as having the potential for providing clinically relevant efficacy response while minimizing potential risks of AEs.

9.5. Selection and Timing of Doses

Each patient will receive blinded investigational product (baricitinib 4 mg or baricitinib placebo tablets orally QD) beginning on the day of Visit 2 (Week 0). Baricitinib 4-mg oral dosing will continue daily through Visit 15 (Week 52). Baricitinib placebo oral dosing will continue daily through Visit 11 (Week 24); at Visit 11 (Week 24), all patients randomized to placebo will be reassigned to baricitinib.

Oral investigational product should be taken at home at approximately the same time each day, as much as possible:

- At Visit 2 (Week 0), patients will take their investigational product in the clinic.

9.5.1. Special Treatment Considerations

In consideration of the disease severity, all patients in Study JAGS will be offered rescue therapy starting at Week 16 if they are determined to be nonresponders.

All patients who are non-responders (in both baricitinib and placebo arm) will receive rescue therapy at Week 16. Patients who are nonresponders and who were randomized to placebo will be rescued with baricitinib. To maintain study integrity through investigational product blinding, patients who were originally randomized to baricitinib will continue to receive baricitinib, with appropriate ongoing assessment of response and use of concomitant medication as needed.

Nonresponse at Week 16 is defined as lack of improvement of at least 20% in both TJC and SJC at both Week 14 and Week 16 compared to baseline. At Week 16, the IWRS will assign rescue treatment based on TJC and SJC values.

After Week 16, rescue therapy will be offered to patients at the discretion of the investigator based on TJCs and SJCs. Once a patient has been rescued, new NSAIDs, corticosteroids, and/or analgesics may be added or doses of ongoing concomitant NSAIDs, corticosteroids, and/or analgesics may be increased at the discretion of the investigator.

A patient may only be rescued once. If a patient continues to meet nonresponse criteria for 4 weeks after rescue or at any time point thereafter, that patient should be discontinued from the study.

Nonresponders who are rescued will receive an oral tablet daily for the remainder of the study. Rescue therapy will be administered through Week 52.

9.6. Continued Access to Investigational Product

After the Week 52 study visit, eligible patients may proceed to a separate extension study (Study JADY) lasting for up to 48 months.

9.7. Blinding

This is a double-blind study. To preserve the blinding of the study, a minimum number of Lilly personnel will see the randomization table and treatment assignments before the study is complete. Emergency unblinding for AEs may be performed through an IWRS, which may supplement or take the place of emergency codes generated by a computer drug-labeling system. This option may be used ONLY if the patient's well-being requires knowledge of the patient's treatment assignment. All operations resulting in an unblinding event are recorded and reported by the IWRS.

The investigator should make every effort to contact the Lilly clinical research physician prior to unblinding a patient's treatment assignment. If a patient's treatment assignment is unblinded, Lilly must be notified immediately by telephone.

If an investigator, site personnel performing assessments, or patient is unblinded, the patient must be discontinued from the study. In cases for which there are ethical reasons to have the patient remain in the study, the investigator must obtain specific approval from a Lilly clinical research physician for the patient to continue in the study.

9.8. Concomitant Therapy

Patients will be instructed to consult the investigator or other appropriate study personnel at the site before taking any new medications or supplements during the study.

As listed below, all patients will continue to take the background MTX therapy at a stable dose throughout the study.

Additional drugs are to be avoided during the study unless required to treat an AE or for the treatment of an ongoing medical condition. Investigators should follow local guidelines for the management of lipid disorders. If the need for other concomitant medications arises, discontinuation of the patient from study drug or the study will be at the discretion of the investigator in consultation with Lilly (or designee).

Patients will not be permitted to have received a live vaccination up to 12 weeks prior to baseline or at any time during the study (with the exception of herpes zoster vaccination, which is not permitted within 30 days of planned randomization or at any time during the study).

Treatment with concomitant DMARDs during the study is permitted as outlined below and in the inclusion/exclusion criteria:

- Use of concomitant oral MTX for at least 12 weeks with treatment at a stable dose of 7.5 to 25 mg/week (or the equivalent injectable dose) for at least 8 weeks prior to entry into the study is required for study participation. Patients should remain on the same dose of MTX throughout Part A and Part B of the study; however, the MTX dose may be adjusted for safety reasons.
- Patients may also be on concomitant hydroxychloroquine or sulfasalazine. If receiving hydroxychloroquine (up to 400 mg/day) or sulfasalazine (up to 3000 mg/day), the patient must be receiving a stable dose for at least 8 weeks prior to entry into the study, and the dose must remain stable throughout the study.
- Concomitant use of NSAIDs is permitted during Part A of the study only if the patient was on a stable dose for at least 6 weeks before planned randomization. An increase of NSAID dose and/or introduction of new NSAIDs are not permitted during Part A of the study unless a patient receives rescue therapy. After rescue or during Part B of the study, new NSAIDs or increases in doses of ongoing concomitant NSAIDs are permitted. Dose reductions and/or termination of NSAIDs are permitted at any time.
- During the study, concomitant use of analgesics is permitted. Increase of an analgesic dose and/or introduction of a new analgesic is not permitted during Part A of the study unless a patient receives rescue therapy. After rescue, new analgesics or increases in doses of ongoing concomitant analgesics are permitted. Dose reductions and/or termination of analgesics are permitted at any time.
- Prednisone (or equivalent) at doses up to 10 mg per day is allowed during this study but must be maintained at stable levels from 6 weeks prior to randomization through the treatment phase of the study, unless a patient receives rescue therapy. After rescue, new corticosteroids or increases in doses of ongoing concomitant corticosteroids are permitted. Patients who were not previously on prednisone (or equivalent) prior to

randomization should not initiate corticosteroid therapy during the study. Patients should not receive other systemic corticosteroids during the study including intra-muscular or intra-articular corticosteroids. Topical, intranasal, intra-ocular, and inhaled corticosteroids are permitted.

- If an unforeseen intra-articular glucocorticoid injection is required during the study, it will be noted as a protocol deviation.
- It is strongly encouraged to avoid hyaluronic acid injections during the course of the study. If required, use should be minimized and discussed with the Sponsor.
- Local standard of care should be followed for concomitant administration of folic acid.

The following concomitant medications are prohibited after randomization:

- Live vaccines, including herpes zoster vaccination. Nonlive seasonal vaccinations and/or emergency vaccination, such as rabies or tetanus vaccinations, are allowed.
- Any DMARDs, other than stable doses of background DMARDs being used at the time of study entry
- Any biologic therapy (such as TNF, interleukin-1, IL-6, T-cell-, or B-cell-targeted therapies) for any indication
- Any interferon therapy (such as Roferon-A, Intron-A, Rebetrone, Alferon-N, Peg-Intron, Avonex, Betaseron, Infergen, Actimmune, Pegasys)
- Any parenteral corticosteroid administered by intramuscular or IV injection
- Any traditional medicine that is known to have immunosuppressant and/or immunomodulatory effects, such as Pafulin or Leigongteng (Common Three Throwingnut Root)

In addition to the concomitant medications guidelines outlined above, all patients who have not already received the herpes zoster vaccine at study entry will be encouraged to do so. Lilly will provide this vaccine where available. After receiving the herpes zoster vaccine, all patients must wait at least 30 days before randomization to study treatment.

Any joints that had been injected with intraarticular corticosteroids within 6 weeks prior to randomization will be censored from the TJs and SJs for the duration of the study.

All concomitant medications taken during the study must be recorded in the Concomitant Medication sections of the CRF.

9.9. Treatment Compliance

Patient compliance with study medication will be assessed at Visits 5 through 7 and Visits 9 through 15 (and, if necessary, at Early Termination) during the treatment period by counting returned tablets. Deviations from the prescribed dosage regimen should be recorded in the CRF.

For baricitinib or oral placebo, a patient will be considered significantly noncompliant if he or she misses >20% of the prescribed doses during the study 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 medication. 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.

10. Efficacy, Health Outcome Measures, Safety Evaluations, Sample Collection and Testing, and Appropriateness of Measurements

Study procedures and their timing (including tolerance limits for timing) are summarized in the Study Schedule ([Attachment 1](#)).

10.1. Efficacy Measures

10.1.1. Primary Efficacy Measure

The primary efficacy measure will be the proportion of patients achieving ACR20 at Week 12. ACR20 is defined as at least 20% improvement in the following ACR Core Set values:

- TJC (68 joint count)
- SJC (66 joint count)
- An improvement of $\geq 20\%$ in at least 3 of the following 5 assessments:
 1. Patient's Assessment of Pain (VAS)
 2. Patient's Global Assessment of Disease Activity (VAS)
 3. Physician's Global Assessment of Disease Activity (VAS)
 4. Patient's Assessment of Physical Function as measured by the HAQ-DI
 5. Acute phase reactant as measured by hsCRP.

10.1.2. Secondary Efficacy Measures

The secondary efficacy measures described below will be collected at the times shown in the Study Schedule ([Attachment 1](#)).

10.1.2.1. Disease Activity Score-Erythrocyte Sedimentation Rate and Disease Activity Score-High Sensitivity C-Reactive Protein

The DAS28 is a measure of disease activity in 28 joints that consists of a composite numeric score of the following variables: TJC, SJC, hsCRP *or* ESR, and Patient's Global Assessment of Disease Activity (Vander Cruyssen et al. 2005). The 28 joints to be examined and assessed as tender or not tender for TJC and as swollen or not swollen for SJC 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 (Smolen et al. 1995).

10.1.2.2. ACR50 and ACR70 Indices

The ACR50 and ACR70 responses are secondary efficacy measures that are calculated as improvements of at least 50% and of at least 70%, respectively, in the ACR Core Set values identified in Section [10.1.1](#).

10.1.2.3. Hybrid American College of Rheumatology Response Measure

The hybrid ACR (bounded) response measure will be obtained as described by the ACR Committee to Reevaluate Improvement Criteria (2007); see [Table JAGS.10.1](#).

Table JAGS.10.1. Scoring Method for the Hybrid American College of Rheumatology Response Measure

ACR Status	Mean % Change in Core Set Measures ^a			
	<20	≥20, <50	≥50, <70	≥70
Not ACR20	Mean % change	19.99	19.99	19.99
ACR20 but not ACR50	20	Mean % change	49.99	49.99
ACR50 but not ACR70	50	50	Mean % change	69.99
ACR70	70	70	70	Mean % change

Abbreviations: ACR = American College of Rheumatology; ACR20 = 20% improvement in ACR criteria; ACR50 = 50% improvement in ACR criteria; ACR70 = 70% improvement in ACR criteria.

^a 1) Calculate the average percentage change in core set measures. For each core set measure, subtract score after treatment from baseline score and determine percentage improvement in each measure. Next, if a core set measure worsened by >100%, limit that percentage change to 100% (set equal to -100% bound). Then, average the percentage changes for all core set measures. 2) Determine whether the patient has achieved ACR20, ACR50, or ACR70. 3) Using the table above, obtain the hybrid ACR response measure. To use the table, take the ACR20, ACR50, or ACR70 status of the patient (left column) and the mean percentage improvement in core set items; the hybrid ACR score is where they intersect in the table.

10.1.2.4. ACR/EULAR Rheumatoid Arthritis Remission

Two ACR/EULAR definitions of RA remission will be evaluated, a “Boolean-based definition” and an “index-based definition” (Felson et al. 2011).

- Boolean-based Definition of Remission: All 4 criteria below must be met:
 - TJC28 ≤ 1
 - SJC28 ≤ 1
 - hsCRP ≤ 1 mg/dL (10 mg/L)
 - patient global DAS on VAS (0 to 10.0) ≤ 1
- Index-based Definition of Remission (see Section 10.1.2.5):
 - SDAI ≤ 3.3

10.1.2.5. Simplified Disease Activity Index (SDAI)

The SDAI is a tool for measurement of disease activity in RA that integrates measures of physical examination, acute phase response, patient self-assessment, and evaluator assessment. The SDAI is calculated by adding together scores from the following assessments:

- number of swollen joints (0 to 28),
- number of tender joints (0 to 28),
- CRP in mg/dL (0.1 to 10.0),
- patient global DAS on VAS (0 to 10.0), and
- evaluator global health score on VAS (0 to 10.0) (Aletaha and Smolen 2005)

Disease remission according to ACR/EULAR index-based definition of remission is defined as an SDAI score of ≤3.3 (Felson et al. 2011).

10.1.2.6. Clinical Disease Activity Index (CDAI)

The CDAI is similar to the SDAI, but it allows for immediate scoring because it does not use a laboratory result. The CDAI is calculated by adding together scores from the following assessments:

- number of swollen joints (0 to 28),
- number of tender joints (0 to 28),
- patient global DAS on VAS (0 to 10.0), and
- evaluator global health score on VAS (0 to 10.0) (Aletaha and Smolen 2005)

Remission is defined as a CDAI score of ≤ 2.8 (Felson et al. 2011).

10.1.2.7. European League Against Rheumatism Responder Index

Assessments of patients with RA by EULAR response criteria will be used to categorize patients as nonresponders, moderate responders, good responders, or responders (moderate + good responders) according to [Table JAGS.10.2](#) (van Gestel et al. 1998).

Table JAGS.10.2. Categorization of Patients as Nonresponders, Moderate Responders, or Good Responders

Postbaseline Level of DAS28	Improvement Since Baseline in DAS28		
	>1.2	≤ 1.2 and > 0.6	≤ 0.6
DAS28 ≤ 3.2	Good response	Moderate response	No response
$3.2 < \text{DAS28} \leq 5.1$			
DAS28 > 5.1			

Abbreviation: DAS28 = Disease Activity Score modified to include the 28 diarthrodial joint count.

10.1.2.8. Van der Heijde Modified Total Sharp Score (mTSS)

Structural progression will be measured using the mTSS (van der Heijde 2000). During screening, all patients meeting entry criteria will have a single posteroanterior radiographic assessment of the left and right hand/wrist and a single dorsoplantar radiographic image taken of the left and right foot. Alternatively, images taken within 4 weeks prior to baseline may be used. Images taken at screening (or within 4 weeks of screening) will be reviewed centrally by qualified readers for evidence of erosive bony change(s). These initial radiographs will serve as the baseline radiographs for comparison throughout the study. If a patient undergoes rescreening, radiographic images acquired as part of initial screening and within 10 weeks of randomization may be used. Radiographs (that is, a single posteroanterior view of each hand and a single dorsolantar view of each foot) will be obtained at Weeks 16, 24, and 52.

For patients who discontinue prematurely from the study, x-rays will be performed at the early termination visit only if the previous x-rays were taken more than 12 weeks earlier.

The mTSS (van der Heijde 2000) quantifies the extent of bone erosions and joint space narrowing for 44 and 42 joints, respectively, of the hands/wrists and feet with higher scores

representing greater damage. X-rays will also be assessed for joint space narrowing and bone erosions.

Quality control will be performed for proper imaging examination technique prior to allocation of the images to the central readers. The independent read of x-ray images will be performed by 2 primary readers and 1 adjudicator, when necessary, based on predefined criteria. The 2 primary readers will each read 100% of the study patients. All time points will be displayed in random order for a given patient. The reader will have no knowledge of the true chronologic order, patient identity, or treatment group. To ensure a consistent read by the independent readers, an inter-/intra-reader variability assessment will be included in the central imaging core laboratory independent read. Repeat x-rays will be requested for all images obtained that are of poor quality as defined in the Image Acquisition Guidelines.

10.2. Health Outcome Measures

The health outcome measures listed below will be administered. Country-specific translations of each measure will be available in this study. Items being measured using a paper diary should be completed at the end of the patient's day (typically their bedtime).

- Morning Joint Stiffness Severity Numeric Rating Scale (NRS): The Morning Joint Stiffness NRS is a patient-administered, single-item, 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.” Patients rate their morning joint stiffness each day (using a paper diary) by selecting the one number that describes their overall level of joint stiffness at the time they woke up.
- Duration of morning joint stiffness:
 - Patient diary: The duration of morning joint stiffness is a patient-administered item that allows for the patients to enter the length of time in hours and minutes that their morning joint stiffness lasted each day.
 - Electronic PRO (tablet): The duration of morning joint stiffness is a patient-administered item that allows for the patients to enter the length of time in minutes (using 15-minute increments) that their morning joint stiffness lasted on the day prior to that visit (using an electronic patient-reported outcomes [ePRO] tablet).
- Recurrence of joint stiffness: The recurrence of joint stiffness throughout the day is a patient-administered single-item question assessed each day (using a paper diary) that asks “Did you have any joint stiffness throughout today, other than when you woke up?” (Yes/No)
- Tiredness Severity Numeric Rating Scale (Worst Tiredness NRS):
 - Patient diary: The Worst Tiredness NRS is a patient-administered, single-item, 11-point horizontal scale anchored at 0 and 10, with 0 representing “no tiredness” and 10 representing “as bad as you can imagine.” Patients rate their tiredness each day (using a paper diary) by selecting the one number that describes their worst level of tiredness.

- Electronic PRO (tablet): The Worst Tiredness NRS is a patient-administered, single-item, 11-point horizontal scale anchored at 0 and 10, with 0 representing “no tiredness” and 10 representing “as bad as you can imagine.” Patients rate their tiredness by selecting the one number that describes their worst level of tiredness during the past 24 hours.
- Severity of joint pain Numeric Rating Scale (Worst Joint Pain NRS):
 - Patient diary: The Worst Joint Pain NRS is a patient-administered, single-item, 11-point horizontal scale anchored at 0 and 10, with 0 representing “no pain” and 10 representing “pain as bad as you can imagine.” Patients rate their joint pain each day (using a paper diary) by selecting the one number that describes their worst level of joint pain.
 - Electronic PRO (tablet): The Worst Joint Pain NRS is a patient-administered, single-item, 11-point horizontal scale anchored at 0 and 10, with 0 representing “no pain” and 10 representing “joint pain as bad as you can imagine.” Patients rate their joint pain by selecting the one number that describes their worst level of joint pain during the past 24 hours.
- HAQ-DI: The HAQ-DI is a patient-reported questionnaire that is commonly used in RA to measure disease-associated disability (assessment of physical function). It consists of 24 questions referring to 8 domains: dressing/grooming, arising, eating, walking, hygiene, reach, grip, and activities (Fries et al. 1980, 1982; Ramey et al. 1996). The disability section of the questionnaire scores the patient’s self-perception on the degree of difficulty (0 = without any difficulty, 1 = with some difficulty, 2 = with much difficulty, and 3 = unable to do) when dressing and grooming, arising, eating, walking, hygiene, reach, grip, and performing other daily activities. The reported use of special aids or devices and/or the need for assistance of another person to perform these activities is also assessed. The scores for each of the functional domains will be averaged to calculate the functional disability index.
- FACIT-F scale: The FACIT-F 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-F uses 0 (“not at all”) to 4 (“very much”) numeric rating scales to assess fatigue and its impact in the past 7 days. Scores range from 0 to 52 with higher scores indicating less fatigue.
- WPAI-RA: The WPAI-RA questionnaire was developed to measure the effect of general health and symptom severity on work productivity and regular activities. It contains 6 items covering overall work productivity (health), overall work productivity (symptom), impairment of regular activities (health), and impairment of regular activities (symptom). Scores are calculated as impairment percentages (Reilly et al. 1993).
- EQ-5D-5L: The EQ-5D-5L is a standardized measure of health status that provides a simple, generic measure of health for clinical and economic appraisal. The EQ-5D-5L consists of 2 components: a descriptive system of the respondent’s health and a rating of his or her current health state using a 0- 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/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 a cardinal score. The VAS records the respondent's self-rated health on a vertical VAS in which 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 a value (also called weights) to each of the levels in each dimension (Brooks 1996; EuroQol Group 2011 [WWW]; Herdman et al. 2011).

- **Quick Inventory of Depressive Symptomatology Self-Rated-16 (QIDS-SR₁₆):** The QIDS-SR₁₆ is a 16-item, self-report 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*, 4th Edition (APA 1994). Patients are 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. Additional information and the QIDS-SR₁₆ questions may be found on the University of Pittsburgh Epidemiology Data Center web site (Inventory of Depressive Symptomatology/Quick Inventory of Depressive Symptomatology [WWW]) (Rush et al. 2003; Trivedi et al. 2004).
- **Healthcare resource utilization:** The healthcare resource utilization data will be collected by site staff regarding the number of visits to medical care providers such as general practitioners, specialists, physical or occupational therapists, and other non-physical care providers for services outside of the clinical study; emergency room admissions; hospital admissions; and concomitant medications related to the treatment of RA. These data will be collected to support economic evaluations of treatment.

10.3. Safety Evaluations

Investigators will be responsible for monitoring the safety of patients who enter 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 will be responsible for the appropriate medical care of patients during the study.

The investigator remains responsible for following, through an appropriate health care option, AEs that are serious or that cause the patient to discontinue before completing the study. The patient should be followed until the event is resolved or explained. Frequency of follow-up evaluation will be left to the discretion of the investigator.

10.3.1. Adverse Events (AEs)

Lilly has standards for reporting AEs that are to be followed regardless of applicable regulatory requirements that may be less stringent.

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

Cases of pregnancy that occur during maternal or paternal exposures to investigational product or drug delivery system should be reported. Data on fetal outcome and breastfeeding are collected for regulatory reporting and drug safety evaluation.

Study site personnel will record the occurrence and nature of each patient's preexisting conditions, including clinically significant signs and symptoms of the disease under treatment in the study.

After the ICF is signed, site personnel will record any change in the condition(s) and the occurrence and nature of any AEs. All AEs related to protocol procedures will be reported to Lilly or its designee.

In addition, all AEs occurring after the patient receives the first dose of investigational product must be reported to Lilly or its designee via electronic entry.

Any clinically significant findings from ECGs, laboratory tests, vital sign measurements, or other procedures that result in a diagnosis should be reported to Lilly or its designee.

Investigators will be instructed to report to Lilly or its designee their assessment of the potential relatedness of each AE to protocol procedure, studied disease state, investigational product, and/or drug delivery system via electronic entry.

Study site personnel must alert Lilly or its designee within 24 hours of the investigator **unblinding** a patient's treatment group assignment for any reason.

If a patient's dosage is reduced or treatment is discontinued as a result of an AE, study site personnel must clearly report to Lilly or its designee via electronic entry the circumstances and data leading to any such dosage reduction or discontinuation of treatment.

10.3.1.1. Adverse Events of Special Interest

Adverse events of special interest (AESIs) will include:

- severe or opportunistic infections
- myelosuppressive events of anemia, leukopenia, neutropenia, lymphopenia, and thrombocytopenia
- thrombocytosis
- elevations in ALT or AST (>3 times ULN) with total bilirubin (>2 times ULN)

Patients with these laboratory value specified events will be identified using the same criteria for the interruption of investigational product with the exception of anemia, which will be identified using the same criteria for the discontinuation of investigational product, and thrombocytosis, which will be defined as a platelet count >600,000/ μ L.

10.3.1.2. Serious Adverse Events (SAEs)

Serious adverse event collection will begin after the patient has signed informed consent and has received investigational product. If a patient experiences an SAE after signing informed consent

but prior to receiving investigational product, the event will NOT be collected unless the investigator feels the event may have been caused by a protocol procedure.

Previously planned (prior to signing the ICF) surgeries should not be reported as SAEs unless the underlying medical condition has worsened during the course of the study.

Study site personnel must alert Lilly or its designee of any **serious** adverse event (SAE) within 24 hours of investigator awareness of the event via a sponsor-approved method. Alerts issued via telephone are to be immediately followed with official notification on study-specific SAE forms. 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
- considered significant by the investigator for any other reason

In addition, any patient who is permanently discontinued from study drug for an AE or abnormal laboratory result should have the AE or abnormal laboratory value reported as an SAE (reported as ‘Other Reason Serious’).

Important medical events that may not result in death, be life-threatening, or require hospitalization may be considered serious adverse drug events when, based on appropriate medical judgment, they may jeopardize the patient and may require medical or surgical intervention to prevent one of the outcomes listed in this definition.

Serious adverse event occurring after a patient has taken the last dose of investigational product will be collected in the pharmacovigilance system and the clinical data collection database for 28 days after the last dose of investigational product, regardless of the investigator’s opinion of causation. Thereafter, SAEs are not required to be reported unless the investigator feels the events were related to investigational product, drug delivery system, or protocol procedure.

Information on SAEs expected in the study population independent of drug exposure and that will be assessed by the sponsor in aggregate periodically during the course of the trial may be found in the IB. In the RA population, the occurrence of malignancies, major cerebrocardiovascular events (including death, myocardial infarction, hospitalization for unstable angina, hospitalization for heart failure, serious arrhythmia, resuscitated sudden death, cardiogenic shock due to myocardial infarction, coronary revascularization procedure; neurologic-stroke; and peripheral vascular events), and serious infections is reasonably anticipated because of the age of the population, comorbid conditions, disease state, and concomitant medications.

10.3.1.2.1. Suspected Unexpected Serious Adverse Reactions (SUSARs)

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.

Lilly has procedures, which are consistent with global regulations and the associated detailed guidances that will be followed for the recording and expedited reporting of SUSARs.

10.3.2. Other Safety Measures

10.3.2.1. Measures Evaluated During Screening Period

Physical examination. One complete physical examination (excluding pelvic and rectal examinations) will be performed at screening. This examination will determine whether the patient meets the criteria required to participate in the study and will also serve as a monitor for preexisting conditions and as a baseline for TEAE assessment. Body weight and height will also be recorded.

Electrocardiograms. A single 12-lead ECG will be obtained at screening to determine whether the patient meets entry criteria for the study. Electrocardiograms (ECGs) will be locally (machine) read and interpreted by a qualified physician (the investigator or qualified designee) at the site.

Chest x-ray and tuberculosis testing. A posterior-anterior view chest x-ray will be obtained locally, unless results from a chest x-ray obtained within 6 months prior to the study are available. The chest x-ray will be reviewed by the investigator or his/her designee to exclude patients with active TB infection.

In addition, patients will be tested at screening and entry for evidence of active or latent TB as described in exclusion criteria, Section [8.2.2](#).

10.3.2.2. Measures Evaluated During Screening and Treatment Periods

Vital signs and physical characteristics. Vital signs (sitting blood pressure and heart rate) will be measured at times indicated in the Study Schedule ([Attachment 1](#)). 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. Any clinically significant findings that result in a diagnosis should be captured on the CRF and reported as an AE. Additional measurements of vital signs may be performed at the discretion of the investigator. Physical characteristics including weight and waist circumference will be measured at times indicated in the Study Schedule ([Attachment 1](#)).

Standard laboratory tests. Hematology, clinical chemistry, urinalysis, lipid profile, eGFR, iron studies, hsCRP, and ESR will be measured at times indicated in the Study Schedule ([Attachment 1](#)).

Hepatitis B virus (HBV) DNA monitoring: Patients who were hepatitis B core antibody (HBcAb) positive at screening will require HBV DNA monitoring as indicated in the study schedule ([Attachment 1](#)). If an HBV DNA result of “target detected” ≥ 29 IU/mL, then the patient should be referred to a hepatology specialist immediately. In selected cases, investigators may temporarily continue study drug 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 evaluation of individual patient risks and benefits.

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

- If a result of “target detected” is obtained at any time during the study with a value of ≥ 29 IU/mL, refer to [Table JAGS.8.2](#)
- If a single result of “target detected” is obtained with a value of < 29 IU/mL, 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 two or more test results of “target detected” with a value of < 29 IU/mL at any time, HBV testing should be conducted approximately once per month for the remainder of the study and referral to a hepatologist is recommended.

The additions to the Study Schedule and table of Clinical Laboratory Tests are shown in [Attachment 1](#) and [Attachment 2](#).

10.3.3. Safety Monitoring

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

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

Liver function monitoring will occur frequently throughout the study. If a study patient experiences elevated ALT or AST $\geq 3 \times$ ULN or elevated total bilirubin $\geq 2 \times$ ULN, clinical and laboratory monitoring should be initiated by the investigator. Details for hepatic monitoring depend on the severity and persistence of observed laboratory test abnormalities. To ensure patient safety and comply with regulatory guidance, the investigator is to consult with the Lilly

designated medical monitor regarding collection of specific recommended clinical information and follow-up laboratory tests (see [Attachment 3](#)).

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, in addition, the following:

- regular monitoring of lipid levels
- adjudication of major adverse cardiovascular events (all deaths and nonfatal myocardial infarctions, hospitalization for unstable angina, hospitalization for heart failure, coronary interventions [such as coronary artery bypass graft or percutaneous coronary intervention], stroke, and transient ischemic attack) by a Cardiovascular Safety Committee at regular intervals
- estimation of relative rates of occurrence of select AEs using observations obtained from a matched, observational, comparator cohort of RA patients using existing RA registry/registries.

10.3.4. Complaint Handling

Lilly collects product complaints on investigational products and drug delivery systems used in clinical studies to ensure the safety of study participants, monitor quality, and facilitate process and product improvements. Complaints related to unblinded concomitant drugs or drug delivery systems are reported directly to the manufacturers of those drugs/devices in accordance with the package inserts. For blinded studies, all product complaints associated with material packaged, labeled, and released by Lilly or its delegate will be reported.

The investigator or his/her designee is responsible for handling the following aspects of the product complaint process in accordance with the instructions provided for this study:

- recording a complete description of the product complaint reported and any associated AEs using the study-specific complaint forms provided for this purpose
- faxing the completed product complaint form within 24 hours to Lilly or its designee

If the investigator is asked to return the product for investigation, he/she will return a copy of the product complaint form with the product.

10.4. Sample Collection and Testing

[Attachment 1](#) lists the schedule for sample collections in this study, and [Attachment 2](#) lists the specific tests that will be performed for this study.

10.4.1. Samples for Standard Laboratory Testing

Standard laboratory tests, including chemistry, hematology, and urinalysis panels, will be performed. Specifically, hsCRP results will be blinded after Visit 2. Serum and/or urine pregnancy tests will be performed at each visit as described in the study schedule ([Attachment 1](#)). Blood and urine samples will be collected for other laboratory tests at the times specified in the study schedule ([Attachment 1](#)). Blood will be collected by venipuncture. [Attachment 2](#) lists the specific tests that will be performed for this study.

Routine and other clinical laboratory tests will be analyzed by a central laboratory selected by Lilly.

Investigators must document their review of each laboratory safety report.

Samples collected for specified laboratory tests will be destroyed within 60 days of receipt of confirmed test results. Tests are run and confirmed promptly whenever scientifically appropriate. When scientific circumstances warrant, however, it is acceptable to retain samples to batch the tests run or to retain the samples until the end of the study to confirm that the results are valid. Certain samples may be retained for a longer period, if necessary, to comply with applicable laws, regulations, or laboratory certification standards.

10.4.2. Exploratory Stored Samples

10.4.2.1. Pharmacogenetics

There is growing evidence that genetic variation may impact a patient's response to therapy. Variable response to therapy may be due to genetic determinants that affect drug absorption, distribution, metabolism, and excretion; the mechanism of action of the drug; the disease etiology; and/or the molecular subtype of the disease being treated. Therefore, where local regulations allow, blood samples will be collected for pharmacogenetic analysis, as noted in the Study Schedule ([Attachment 1](#)).

Samples will be stored, and analysis may be performed on genetic variants thought to play a role in active RA including, but not limited to, the cluster of immune response genes on chromosome 6, such as *HLA-DRB1*, *HLA-DQA1*, *HLA-DPA1*, *HLA-DPBI*, TNF- α , and TNF receptors; RANK-Ligand; the JAK signaling pathways; MTX metabolizing genes such as methylene tetrahydrofolate reductase and other folate reductases; PTPN22 and related signaling molecules; IL-6 pathway members; and the STAT signaling family members, including STAT4, to evaluate their association with observed response to baricitinib.

In the event of an unexpected AE or the observation of unusual response, the samples may be genotyped and analysis may be performed to evaluate a genetic association with response to baricitinib. These investigations may be limited to a focused candidate gene study or, if appropriate, genome-wide association studies may be performed to identify regions of the genome associated with the variability observed in drug response. Samples will only be used for investigations related to disease and drug or class of drugs under study in the context of this clinical program. They will **not** be used for broad exploratory unspecified disease or population genetic analysis.

The sample will be identified by the patient number (coded) and stored for up to a maximum of 15 years after the last patient visit for the study at a facility selected by the sponsor. The sample and any data generated from it can be linked back to the patient only by investigator site personnel. The duration allows the Sponsor to respond to regulatory requests related to the study drug.

10.4.2.2. Nonpharmacogenetic/Biomarker Stored Samples

Collection of samples for nonpharmacogenetic biomarker research is a part of this study where local regulations allow. Serum, plasma, whole blood RNA, and urine samples will be collected at the times specified in the Study Schedule ([Attachment 1](#)).

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

Samples will be identified by the patient_number (coded) and stored for up to a maximum of 15 years after the last patient visit for the study at a facility selected by the sponsor.

10.5. Appropriateness of Measurements

All of the clinical and safety assessments made in this study are standard, widely used, and generally recognized as reliable, accurate, and relevant.

11. 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 a start-up training session to instruct the investigators and study coordinators. This session 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

In addition, Lilly or its representatives may 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.

To ensure the safety of participants in the study and to ensure accurate, complete, and reliable data, the investigator will keep records of laboratory tests, clinical notes, and patient medical records in the patient files as original source documents for the study. If requested, the investigator will provide the sponsor, applicable regulatory agencies, and applicable ERBs with direct access to original source documents.

11.1. Data Capture System

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

Electronic physician-reported outcome measures will be directly entered by the physician or designee (Physician global assessment of disease activity) or by the blinded joint assessor (TJC/SJC) into the electronic system.

Case report form data will be encoded and stored in a clinical trial database. Data managed by a central vendor, such as laboratory test data or x-ray data, will be stored electronically in the central vendor's database system. Data will subsequently be transferred from the central vendor to the Lilly generic laboratories system.

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

12. Sample Size and Statistical Methods

12.1. Determination of Sample Size

Approximately 288 patients (144 in baricitinib, 144 in placebo) will provide:

- 99% power to detect a difference between the baricitinib and placebo treatment groups in ACR20 response rate (60% versus 35%) at Week 12 based on a 2-sided chi-square test at a significance level of 0.05.

The sample size and power estimates were obtained using nQuery[®] Advisor 7.0.

12.2. Statistical and Analytical Plans

12.2.1. General Considerations

All statistical tests of treatment effects will be performed at 2-sided significance levels of 0.05, unless otherwise stated. Treatment comparisons of discrete efficacy variables will be made using a logistic regression analysis with country, joint erosion status (1-2 joint erosions plus seropositivity vs at least 3 joint erosions) and treatment group in the model. The proportions and 95% CI will be reported. Treatment comparisons of continuous efficacy variables will be made using analysis of covariance (ANCOVA) with country, treatment group, joint erosion status and baseline value in the model. Type III sums of squares for the least-squares (LS) means will be used for the statistical comparison; the 95% CI will also be reported.

When a restricted maximum likelihood (REML)-based mixed model for repeated measures (MMRM) is used, the model will include the fixed, categorical effects of treatment, country, joint erosion status, visit, treatment-by-visit-interaction, as well as the continuous, fixed covariates of baseline score and baseline score-by-visit-interaction. An unstructured (co)variance structure will be used to model the within-patient errors. If this analysis fails to converge, other structures will be tested. Type III sums of squares for the LS means will be used for the statistical comparison; the 95% CI will also be reported. Treatment group comparisons at Week 12 and all other visits will be tested. Further details on the MMRM will be described in the statistical analysis plan (SAP).

Fisher's exact test will be used for all AEs, discontinuation, and other categorical safety data. Continuous vital signs and other continuous safety variables including laboratory parameters will be analyzed using ANCOVA with treatment and baseline value in the model.

Analysis Population:

Unless otherwise specified, the efficacy and safety analyses will be conducted on a modified intent-to-treat basis (mITT) (Gillings and Koch 1991). This analysis set includes all data from all randomized patients who received at least 1 dose of study drug. In addition, the primary and secondary efficacy analyses for the double-blind period will be repeated in a subset of mITT analysis set, the per protocol (PP) set, which is defined as all randomized patients who are compliant with treatment, who do not have significant protocol violations, and whose investigator site does not have significant GCP issues that require a report to the regulatory

agencies. Analysis of structural progression (van der Heijde mTSS) will be conducted on the mITT population using patients with available baseline and at least 1 postbaseline value.

Sensitivity analyses of the primary endpoints may be conducted and will be described in the SAP.

Missing Data Imputation:

In accordance with precedent set with other Phase 3 RA trials (Keystone et al. 2004, 2008, 2009; Cohen et al. 2006; Smolen et al. 2008, 2009), the following methods for imputation of missing data will be used:

1. Nonresponder imputation (NRI): All patients who discontinue the study or the study treatment at any time for any reason will be defined as nonresponders for the NRI analysis for categorical variables, such as ACR20/50/70, from the time of discontinuation and onward. Patients who receive rescue therapy at Week 16 will be analyzed as nonresponders after Week 16 and onward. Randomized patients without available data at a postbaseline visit will be defined as nonresponders for the NRI analysis at that visit.
2. Linear extrapolation method: The linear extrapolation method will be used for analysis of the structural progression endpoint (van der Heijde mTSS) to impute missing data at time points when analyses are conducted (including the analyses at Week 24 and Week 52). For patients who discontinue the study or the study treatment or for patients who miss a radiograph for any reason, baseline data and the most recent radiographic data prior to discontinuation or the missed radiograph, adjusted for time, will be used for linear extrapolation to impute missing data at subsequent time points. For patients who receive rescue therapy starting at Week 16 or at any time point thereafter, baseline data and the most recent radiographic data prior to initiation of rescue therapy, adjusted for time, will be used for linear extrapolation. All patients originally randomized to placebo will be switched to active treatment at Week 24; thus, there will be no observed data for the placebo arm for the structural comparison at subsequent time points. Therefore, baseline data and the most recent radiographic data prior to initiation of the new therapy, adjusted for time, will be used for linear extrapolation at Week 52. The linear extrapolation method has been established as an appropriate missing data imputation method in other Phase 3 RA trials (Keystone et al. 2004, 2008, 2009; Cohen et al. 2006; Smolen et al. 2008, 2009). Missing postbaseline values will be imputed only if both a baseline value and a postbaseline value from another time point are available, as long as the patient was on the same treatment at each applicable time point.
3. Multiple imputation: Other methods than linear extrapolation that include multiple imputation will be employed for analysis of the structural progression endpoint (van der Heijde mTSS) to impute missing data at Week 24. Details will be provided in the SAP.
4. Modified baseline observation carried forward (mBOCF): The mBOCF method will be used for the analysis of secondary continuous endpoints (unless otherwise stated). For patients who discontinue the study or the study treatment because of an AE, including death, the baseline observation will be carried forward to the corresponding endpoint for evaluation. For patients who discontinue the study for reason(s) other than an AE, the last nonmissing postbaseline observation prior to discontinuation will be carried forward

to the corresponding endpoint for evaluation. For patients who receive rescue therapy starting at Week 16, the last nonmissing observation at or before rescue will be carried forward to subsequent time points for evaluation.

5. Modified last observation carried forward (mLOCF): For all continuous measures including safety analyses, the mLOCF will be a general approach to impute missing data unless otherwise specified. For patients who receive rescue therapy starting at Week 16, the last nonmissing observation at or before rescue will be carried forward to subsequent time points for evaluation. For all other patients discontinuing from the study or the study treatment for any reason, the last nonmissing postbaseline observation before discontinuation will be carried forward to subsequent time points for evaluation. After mLOCF imputation, data from patients with nonmissing baseline and postbaseline observations will be included in the analyses. The mLOCF will be used as a sensitivity method to the mBOCF method with respect to the secondary endpoints.
6. Other methods for data imputation for analysis of endpoints other than mTSS may be used and will be described in the SAP.

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 SAP and the clinical study report (CSR). Additional exploratory analyses of the data will be conducted as deemed appropriate.

12.2.2. Patient Disposition

The number of randomized patients along with mITT and PP set populations will be summarized by treatment group. Frequency counts and percentages in different study parts will be presented for each treatment group. All patients who discontinue from the study or the study treatment will be identified, and the extent of their participation in the study will be reported along with their reason for discontinuation. Reasons for discontinuation from the study will be summarized by treatment group.

12.2.3. Patient Characteristics

Demographic and baseline characteristics will be summarized descriptively by treatment group. Descriptive statistics including number of patients, mean, standard deviation (SD), 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.

12.2.4. Concomitant Therapy

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

12.2.5. Treatment Compliance

Compliance to study treatment for baricitinib and placebo will be assessed through counts of returned study drug tablets. For baricitinib or placebo, a patient will be considered significantly noncompliant if he or she misses more than 20% of the prescribed doses during the study, unless the patient's study drug is withheld by the investigator for safety reasons. Similarly, a patient will be considered significantly noncompliant if he or she is judged by the investigator to have intentionally or repeatedly taken more than the prescribed amount of study drug. Patients who are significantly noncompliant will be excluded from the PP set.

Patient compliance will be further defined in the SAP.

12.2.6. Primary Outcome and Methodology

A logistic regression model with treatment, country and joint erosion status in the model will be used to test the treatment difference between baricitinib and placebo in the proportion of patients achieving ACR20 response at Week 12. The NRI method as described above will be used to impute missing data.

12.2.7. Efficacy Analyses

Secondary analyses:

1. Comparison between baricitinib and placebo in change from baseline to Week 12 in HAQ-DI score: An ANCOVA model with treatment, country, joint erosion status, and baseline score in the model will be used to test the treatment difference between baricitinib and placebo in change from baseline to Week 12 in HAQ-DI. The mBOCF and mLOCF approaches as described above will be used to impute missing data.
2. Comparison between baricitinib and placebo in change from baseline to Week 12 in DAS28-hsCRP: An ANCOVA model with treatment, country, joint erosion status, and baseline score in the model will be used to test the treatment difference between baricitinib and placebo in change from baseline to Week 12 in DAS28-hsCRP. The mBOCF and mLOCF approaches as described above will be used to impute missing data.
3. Comparison between baricitinib and placebo in proportion of patients achieving an SDAI score ≤ 3.3 at Week 12: A logistic regression model with treatment, country, and joint erosion status in the model will be used to test the treatment difference between baricitinib and placebo in the proportion of patients achieving an SDAI score ≤ 3.3 response at Week 12. The NRI method as described above will be used to impute missing data.
4. Comparison between baricitinib and placebo in mean duration of morning joint stiffness in the 7 days prior to Week 12: An ANCOVA model with treatment, country, joint erosion status, and baseline score (Day 1 on the diaries) in the model will be used to test the treatment difference between baricitinib and placebo in mean duration of morning joint stiffness in the 7 days prior to Week 12. Refer to the SAP for more details.
5. Comparison between baricitinib and placebo in mean severity of morning joint stiffness in the 7 days prior to Week 12: An ANCOVA model with treatment, country, joint

erosion status, and baseline score (Day 1 on the diaries) in the model will be used to test the treatment difference between baricitinib and placebo in mean severity of morning joint stiffness in the 7 days prior to Week 12. Refer to the SAP for more details.

6. Comparison between baricitinib and placebo in mean Worst Tiredness NRS in the 7 days prior to Week 12: An ANCOVA model with treatment, country, joint erosion status, and baseline score (Day 1 on the diaries) in the model will be used to test the treatment difference between baricitinib and placebo in mean Worst Tiredness NRS in the 7 days prior to Week 12. Refer to the SAP for more details.
7. Comparison between baricitinib and placebo in mean Worst Joint Pain NRS in the 7 days prior to Week 12: An ANCOVA model with treatment, country, joint erosion status, and baseline score (Day 1 on the diaries) in the model will be used to test the treatment difference between baricitinib and placebo in mean Worst Joint Pain NRS in the 7 days prior to Week 12. Refer to the SAP for more details.

Other efficacy and health outcome analyses:

Besides the analyses listed in the objective section, the following analyses will be performed:

- change from baseline in mTSS
- proportion of patients with mTSS change ≤ 0
- change from baseline in joint space narrowing and bone erosion scores
- proportion of patients achieving DAS28-hsCRP ≤ 3.2 and < 2.6
- proportion of patients achieving HAQ-DI improvement ≥ 0.22 and ≥ 0.3
- proportions of patients achieving ACR20 (other than primary), ACR50, and ACR70
- percentage change from baseline in individual components of the ACR Core Set
- change from baseline in DAS28-hsCRP
- change from baseline in DAS28-ESR
- proportion of patients achieving DAS28-ESR ≤ 3.2 and < 2.6
- change from baseline in HAQ-DI
- change from baseline in CDAI score
- change from baseline in SDAI score
- proportion of patients achieving CDAI score ≤ 2.8
- proportion of patients achieving SDAI score ≤ 3.3 (ACR/EULAR remission according to the Index-based definition)
- proportion of patients achieving ACR/EULAR remission according to the Boolean-based definition
- proportion of patients achieving moderate response and good response based on EULAR response criteria
- hybrid ACR (bounded) response measure
- change from baseline in fatigue score in the FACIT-F
- change from baseline in health states, VAS score, and index score of EQ-5D-5L
- change from baseline in absenteeism, presenteeism, work productivity loss, and activity impairment scores of WPAI-RA
- individual items and total score in healthcare resource utilization

- change from baseline in duration of morning joint stiffness assessed at the visit using an ePRO tablet
- change from baseline in Worst Tiredness NRS assessed at the visit using an ePRO tablet
- change from baseline in Worst Joint Pain NRS assessed at the visit using an ePRO tablet.

Approaches similar to the primary and above secondary analyses will be used to analyze other efficacy and health outcome measures. Analysis methods will consist of ANCOVA and logistic regression, as described above, using mBOCF, linear extrapolation, NRI, and mLOCF missing data imputation methods, where applicable. In addition, continuous secondary parameters that are collected at repeated visits throughout the double-blind period may be analyzed using a MMRM approach as described in Section 12.2.1. The MMRM will be considered as a sensitivity approach to ANCOVA results.

The Study JAGS data of mTSS, joint space narrowing and bone erosion scores will be pooled with JADV data for further analyses.

Data collected after initiation of rescue therapy will be summarized as appropriate. Further details will be provided in the SAP.

12.2.8. Health Outcome Analyses

The health outcomes will be analyzed using methods described for continuous or categorical data as described for efficacy measures in Section 12.2.7. More detailed analytical methods will be described in the SAP.

12.2.9. Safety Analyses

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

For the purposes of calculating changes from baseline during Part A and Part B treatment, baseline may be considered as the original baseline for Part A – prior to any study drug received – or at the visits that mark the transition between treatments received, depending on the purpose of the summary or comparison being made. Summaries within Parts A and B will be conducted by treatment group used within each respective part. Additional summaries by treatment group that continue unchanged across multiple parts of the study will be conducted, as appropriate.

Treatment-emergent adverse events are defined as AEs that first occurred or worsened in severity after the first dose of study treatment. AEs will be considered treatment emergent for nonresponders assigned to baricitinib if the AEs first occurred or worsened in severity after the visit at which rescue therapy is assigned. The number of TEAEs as well as the number and percentage of patients who experienced at least 1 TEAE will be summarized using Medical Dictionary for Regulatory Activities (MedDRA) for each system organ class (SOC) (or a body system) and each preferred term by treatment group. TEAEs will also be summarized by relationship to treatment and by severity within each treatment group. Serious adverse events (SAEs) and AEs that lead to study drug discontinuation will also be summarized by treatment

group. Fisher's exact test will be used to perform the statistical comparisons between baricitinib and placebo.

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

Observed values and changes from baseline (predose or screening if missing) for vital signs and physical characteristics will be descriptively summarized by treatment group and time point. Change from baseline to postbaseline in vital signs, QIDS-SR₁₆, and body weight will be analyzed using ANCOVA with treatment and baseline value in the model.

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

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

12.2.10. Subgroup Analyses

Subgroups analyses comparing baricitinib to placebo will be performed using ACR20, ACR50, HAQ-DI and DAS28-hsCRP at Week 12, Week 24, and mTSS at Week 24. Subgroups to be evaluated will include country, renal function, background therapy, joint erosion status, gender, age, race, etc. Any other additional subgroups will be defined in the SAP.

For ACR20 and ACR50 analyses at Week 12 and Week 24, a logistic regression model will be used with treatment, subgroup, and treatment by subgroup interaction included as factors.

For the change from baseline to Weeks 12 and Week 24 in HAQ-DI score and DAS28-hsCRP and to Week 24 in mTSS, an ANCOVA model will be used with treatment, baseline score, subgroup, and treatment by subgroup interaction included as factors.

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.

12.2.11. Planned Analyses

Two database locks are planned for the study. The first database lock and a blinded review of data will occur after all patients complete the Week 24 study visit. The final database lock and unblinded analysis will occur at the end of the study.

Unblinding details will be specified in the unblinding plan section of the SAP.

13. Informed Consent, Ethical Review, and Regulatory Considerations

13.1. Informed Consent

The investigator is responsible for ensuring that the patient understands the potential risks and benefits of participating in the study, including answering any questions the patient may have throughout the study and sharing in a timely manner any new information that may be relevant to the patient's willingness to continue his or her participation in the trial.

The ICF will be used to explain the potential risks and benefits of study participation to the patient in simple terms before the patient is entered into the study and to document that the patient is satisfied with his or her understanding of the risks and benefits of participating in the study and desires to participate in the study.

The investigator is responsible for ensuring that informed consent is given by each patient. This includes obtaining the appropriate signatures and dates on the ICF prior to the performance of any protocol procedures and prior to the administration of investigational product.

13.2. Ethical Review

Lilly or its representatives must approve all ICFs before they are submitted to the ERB and are used at investigative sites(s). All ICFs must be compliant with the International Conference on Harmonisation (ICH) guideline on GCP.

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). The ERB(s) will review the protocol as required.

Any member of the ERB who is directly affiliated with this study as an investigator or as site personnel must abstain from the ERB's vote on the approval of the protocol.

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

- the current IB or package labeling and updates during the course of the study
- ICF
- relevant curricula vitae

13.3. Regulatory Considerations

This study will be conducted in accordance with:

1. consensus ethics principles derived from international ethics guidelines, including the Declaration of Helsinki and Council for International Organizations of Medical Sciences International Ethical Guidelines
2. the ICH GCP Guideline [E6]
3. applicable laws and regulations

The investigator or designee will promptly submit the protocol to the applicable ERB(s).

An identification code assigned by the investigator to each patient will be used in lieu of the patient's name to protect the patient's identity when reporting AEs and/or other trial-related data.

13.3.1. Investigator Information

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

13.3.2. 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.

13.3.3. Final Report Signature

The clinical study report coordinating investigator will sign the final clinical study report 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.

Lilly will select a qualified investigator from among investigators participating in the design, conduct, and/or analysis of the study to serve as the clinical study report coordinating investigator.

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

14. References

- Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO 3rd, Birnbaum NS, Burmester GR, Bykerk VP, Cohen MD, Combe B, Costenbader KH, Dougados M, Emery P, Ferraccioli G, Hazes JM, Hobbs K, Huizinga TW, Kavanaugh A, Kay J, Kvien TK, Laing T, Mease P, Ménard HA, Moreland LW, Naden RL, Pincus T, Smolen JS, Stanislawska-Biernat E, Symmons D, Tak PP, Upchurch KS, Vencovský J, Wolfe F, Hawker G. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Arthritis Rheum.* 2010;62(9):2569-2581.
- Aletaha D, Smolen J. The Simplified Disease Activity Index (SDAI) and the Clinical Disease Activity Index (CDAI): a review of their usefulness and validity in rheumatoid arthritis. *Clin Exp Rheumatol.* 2005;23(5 suppl 39):S100-S108.
- American College of Rheumatology Committee to Reevaluate Improvement Criteria. A proposed revision to the ACR20: the hybrid measure of American College of Rheumatology response. *Arthritis Rheum.* 2007;57(2):193-202.
- [APA] American Psychiatric Association's *Diagnostic and Statistical Manual of Mental Disorders*, 4th Edition. Washington, DC; American Psychiatric Association. 1994.
- Brooks R. EuroQol: The current state of play. *Health Policy.* 1996;37:53-72.
- Cella D, Webster K. Linking outcomes management to quality-of-life measurement. *Oncology (Willeston Park).* 1997;11(11A):232-235.
- Cohen SB, Emery P, Greenwald MW, Dougados M, Furie RA, Genovese MC, Keystone EC, Loveless JE, Burmester GR, Cravets MW, Hessey EW, Shaw T, Totoritis MC. Rituximab for rheumatoid arthritis refractory to anti-tumor necrosis factor therapy. Results of a multicenter, randomized, double-blind, placebo-controlled phase III trial evaluating primary efficacy and safety at twenty-four weeks. *Arthritis Rheum.* 2006;54(9):2793-2806.
- Colmegna I, Ohata BR, Menard HA. Current understanding of rheumatoid arthritis therapy. *Clin Pharmacol Ther.* 2012;91(4):607-620.
- Curtis JR, Patkar N, Xie A, Martin C, Allison JJ, Saaq M, Shatin D, Saaq KG. Risk of serious bacterial infections among rheumatoid arthritis patients exposed to tumor necrosis factor alpha antagonists. *Arthritis Rheum.* 2007;56(4):1125-1133.
- The EuroQol Group. EQ-5D-5L User Guide. Version 1.0. April 2011. Available at: http://www.euroqol.org/fileadmin/user_upload/Documenten/PDF/Folders_Flyers/UserGuide_EQ-5D-5L.pdf. Accessed July 1 2012.
- Felson DT, Smolen JS, Wells G, Zhang B, van Tuyl LH, Funovits J, Aletaha D, Allaart CF, Bathon J, Bombardieri S, Brooks P, Brown A, Matucci-Cerinic M, Choi H, Combe B, de Wit M, Dougados M, Emery P, Furst D, Gomez-Teino J, Hawker G, Keystone E, Khanna D, Kirwan J, Kvien TK, Landewé R, Listing J, Michaud K, Martin-Mola E, Montie P, Pincus T, Richards P, Siegel JN, Simon LS, Sokka T, Strand V, Tugwell P, Tyndall A, van der Heijde D, Verstappen S, White B, Wolfe F, Zink A, Boers M. American College of Rheumatology/European League against Rheumatism provisional definition of remission in rheumatoid arthritis for clinical trials. *Ann Rheum Dis.* 2011;70:404-413.

- Fleischmann RM. Is there a need for new therapies for rheumatoid arthritis? *J Rheum.* 2005;32(suppl 73):3S-7S.
- 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.
- Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. *Arthritis Rheum.* 1980;23(2):137-145.
- Fries JF, Spitz PW, Young DY. The dimensions of health outcomes: the health assessment questionnaire, disability and pain scales. *J Rheumatol.* 1982;9(5):789-793.
- Gillings D, Koch G. The application of the principle of intention-to-treat to the analysis of clinical trials. *Drug Inf J.* 1991;25:411-424.
- Herdman M, Gudex C, Lloyd A, Janssen MF, Kind P, Parkin D, Bonnel G, Badia X. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res.* 2011;20(10):1727-1736.
- Keystone EC, Kavanaugh AF, Sharp JT, Tannenbaum H, Hua Y, Teoh LS, Fischkoff SA, Chartash EK. Radiographic, clinical, and functional outcomes of treatment with adalimumab (a human anti-tumor necrosis factor monoclonal antibody) in patients with active rheumatoid arthritis receiving concomitant methotrexate therapy: a randomized, placebo-controlled, 52-week trial. *Arthritis Rheum.* 2004;50(5):1400-1411.
- Keystone E, Heijde D, Mason D Jr, Landewé R, Vollenhoven RV, Combe B, Emery P, Strand V, Mease P, Desai C, Pavelka K. Certolizumab pegol plus methotrexate is significantly more effective than placebo plus methotrexate in active rheumatoid arthritis: findings of a fifty-two-week, phase III, multicenter, randomized, double-blind, placebo-controlled, parallel-group study. *Arthritis Rheum.* 2008;58(11):3319-3329.
- Keystone E, Emery P, Peterfy CG, Tak PP, Cohen S, Genovese MC, Dougados M, Burmester GR, Greenwald M, Kvien TK, Williams S, Hagerty D, Cravets MW, Shaw T. Rituximab inhibits structural joint damage in patients with rheumatoid arthritis with an inadequate response to tumour necrosis factor inhibitor therapies. *Ann Rheum Dis.* 2009;68(2):216-221.
- Kitas GD, Gabriel SE. Cardiovascular disease in rheumatoid arthritis: state of the art and future perspectives. *Ann Rheum Dis.* 2011;70(1):8-14. Erratum in *Ann Rheum Dis.* 2011;70(8):1520.
- Klareskog L, Catrina AI, Paget S. Rheumatoid arthritis. *Lancet.* 2009;373(9664):659-672.
- Kremer JM, Bloom BJ, Breedveld FC, Coombs JH, Fletcher MP, Gruben D, Krishnaswami S, Burgos-Vargas R, Wilkinson B, Zerbini CA, Zwillich SH. The safety and efficacy of a JAK inhibitor in patients with active rheumatoid arthritis: Results of a double-blind, placebo-controlled phase IIa trial of three dosage levels of CP-690,550 versus placebo. *Arthritis Rheum.* 2009;60(7):1895-1905.
- Ramey DR, Fries JF, Singh G. The Health Assessment Questionnaire 1995: status and review. In: Spiker B, editor. Quality of life and pharmacoeconomics in clinical trials. 2nd ed. Philadelphia: Lippincott-Raven, 1996: p. 227-237.

- Reilly MC, Zbrozek AS, Dukes EM. The validity and reproducibility of a work productivity and activity impairment instrument. *Pharmacoeconomics*. 1993;4(5):353-365.
- Rubbert-Roth A, Finckh A. Treatment options in patients with rheumatoid arthritis failing initial TNF inhibitor therapy: a critical review. *Arthritis Res Ther*. 2009;11(suppl 1):S1.
- Rush AJ, Trivedi MH, Ibrahim HM, Carmody TJ, Arnow B, Klein DN, Markowitz JC, Ninan PT, Kornstein S, Manber R, Thase ME, Kocsis JH, Keller MB. The 16-Item quick inventory of depressive symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. *Biological Psychiatry*. 2003;54(50):573-583. Erratum in *Biol Psychiatry*. 2003;54(5):585.
- Smolen J, Landewé RB, Mease P, Brzezicki J, Mason D, Luijstens K, van Vollenhoven RF, Kavanaugh A, Schiff M, Burmester GR, Strand V, Vencovský J, van der Heijde D. Efficacy and safety of certolizumab pegol plus methotrexate in active rheumatoid arthritis: the RAPID 2 study. A randomised controlled trial. *Ann Rheum Dis*. 2009;68(6):797-804.
- Smolen JS, Aletaha D, Bijlsma JW, Breedveld FC, Boumpas D, Burmester G, Combe B, Cutolo M, de Wit M, Dougados M, Emery P, Gibofsky A, Gomez-Reino JJ, Haraoui B, Keystone EC, Kvien TK, McInnes I, Martin-Mola E, Montecucco C, Schoels M, van der Heijde D, T2T Expert Committee. Treating rheumatoid arthritis to target: recommendations of an international task force. *Ann Rheum Dis*. 2010;69(4):631-637.
- Smolen JS, Beaulieu A, Rubbert-Roth A, Ramos-Remus C, Rovensky J, Alecock E, Woodworth T, Alten R, OPTION investigators. Effect of interleukin-6 receptor inhibition with tocilizumab in patients with rheumatoid arthritis (OPTION study): a double-blind, placebo-controlled, randomised trial. *Lancet*. 2008;371(9617):987-997.
- Smolen JS, Breedveld FC, Eberl G, Jones I, Leeming M, Wylie GL, Kirkpatrick J. Validity and reliability of the twenty-eight-joint count for the assessment of rheumatoid arthritis activity. *Arthritis Rheum*. 1995;38(1):38-43.
- Smolen J, Steiner G. Therapeutic strategies for rheumatoid arthritis. *Nat Rev Drug Discov*. 2003;2(6):473-488.
- Trivedi MH, Rush AJ, Ibrahim HM, Carmody TJ, Biggs MM, Suppes T, Crismon ML, Shores-Wilson K, Toprac MG, Dennehy EB, Witte B, Kashner TM. The Inventory of Depressive Symptomatology, Clinician Rating (IDS-C) and Self-Report (IDS-SR), and the Quick Inventory of Depressive Symptomatology, Clinician Rating (QIDS-C) and Self-Report (QIDS-SR) in public sector patients with mood disorders: a psychometric evaluation. *Psychol Med*. 2004;34(1):73-82.
- University of Pittsburgh Epidemiology Data Center; IDS/QIDS web site. Available at: <http://www.ids-qids.org>. Accessed June 13, 2012.
- Vander Cruyssen B, Van Looy S, Wyns B, Westhovens R, Durez P, Van den Bosch F, Veys EM, Mielants H, De Clerck L, Peretz A, Malaise M, Verbruggen L, Vastesaegeer N, Geldhof A, Boullart L, De Keyser F. DAS28 best reflects the physician's clinical judgment of response to infliximab therapy in rheumatoid arthritis patients: validation of the DAS28 score in patients under infliximab treatment. *Arthritis Res Ther*. 2005;7(5):R1063-1071.

van der Heijde D. How to read radiographs according to the Sharp/van der Heijde method. *J Rheumatol.* 2000; 27(1):261-263.

van Gestel AM, Haagsma CJ, van Riel PL. Validation of rheumatoid arthritis improvement criteria that include simplified joint counts. *Arthritis Rheum.* 1998;41(10):1845-1850.

Williams WV, Scherle P, Shi J, Newton R, McKeever E, Fridman J, Vaddi K, Levy R, Moreland L. Initial efficacy of INCB018424, a selective Janus kinase 1 & 2 (JAK 1&2) inhibitor in rheumatoid arthritis (RA). *Ann Rheum Dis.* 2008;67(suppl II):62.

Attachment 1. Protocol JAGS Study Schedule

Study Schedule, Protocol I4V-CR-JAGS

	Screening		Blinded, Placebo- Controlled Treatment Part A										Open-label Part B					Follow-up Part Ca
Visit	1	1a	2	3	4	5	6	7	8	9	10	11	12	13	14	15	ET	801
Week of Treatment			Base-line 0	1	2	4	8	12	14	16	20	24	28	32	40	52	Any week	
Days from Visit 2	-42 to -3 days		0	7 ± 2 days	14 ± 2 days	28 ± 2 days	56 ± 3 days	84 ± 3 days	98 ± 3 days	112 ± 3 days	140 ± 3 days	168 ± 3 days	196 ± 3 days	224 ± 5 days	280 ± 5 days	365 ± 5 days	Any day	28 ± 5 days after last dose
Informed consent	X																	
Clinical assessments:																		
History	X																	
Physical examination	X																	
Symptom-directed physical examination ^b			X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Previous therapy/previous RA therapy	X																	
Patient demographics	X																	
Height	X																	
Weight	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Waist circumference	X							X			X				X			
Vital signs (BP and pulse)	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Habits ^c	X																	
Electrocardiogram ^d	X																	
Inclusion/exclusion review for entry	X																	
Inclusion/exclusion review for enrollment			X															

Visit	Screening		Blinded, Placebo- Controlled Treatment Part A										Open-label Part B				ET	Follow-up Part Ca
	1	1a	2	3	4	5	6	7	8	9	10	11	12	13	14	15		
Week of Treatment			Base-line 0	1	2	4	8	12	14	16	20	24	28	32	40	52	Any week	
Days from Visit 2	-42 to -3 days		0	7 ± 2 days	14 ± 2 days	28 ± 2 days	56 ± 3 days	84 ± 3 days	98 ± 3 days	112 ± 3 days	140 ± 3 days	168 ± 3 days	196 ± 3 days	224 ± 5 days	280 ± 5 days	365 ± 5 days	Any day	28 ± 5 days after last dose
Administer PPD or QuantiFERON®-TB Gold or T SPOT® TB teste	X																	
Read PPD or QuantiFERON-TB Golde		X																
Chest x-rayf	X																	
Preexisting conditions	X		X															
Adverse events			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
Randomization			X															
IWRS	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Investigational product dispensed			X ^g			X	X	X		X	X	X	X	X	X			
Investigational products returned and assess compliance				X ^h	X ^h	X	X	X	X ^h	X	X	X	X	X	X	X	X	
Tender/swollen joint count (68/66 joints)	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Patient's assessment of pain visual analog scale			X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Patient global assessment of disease activity			X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

	Screening		Blinded, Placebo- Controlled Treatment Part A										Open-label Part B					Follow-up Part Ca
Visit	1	1a	2	3	4	5	6	7	8	9	10	11	12	13	14	15	ET	801
Week of Treatment			Base- line 0	1	2	4	8	12	14	16	20	24	28	32	40	52	Any week	
Days from Visit 2	-42 to -3 days		0	7 ± 2 days	14 ± 2 days	28 ± 2 days	56 ± 3 days	84 ± 3 days	98 ± 3 days	112 ± 3 days	140 ± 3 days	168 ± 3 days	196 ± 3 days	224 ± 5 days	280 ± 5 days	365 ± 5 days	Any day	28 ± 5 days after last dose
Physician global assessment of disease activity			X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Hand and foot x-rays ⁱ	X ⁱ									X		X				X	X ⁱ	
HAQ-DI			X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Morning joint stiffness - diary			X ^k															
Morning joint stiffness - tablet ^l			X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
FACIT-F			X			X	X	X		X	X	X	X	X	X	X	X	X
EQ-5D-5L			X			X	X	X		X	X	X	X	X	X	X	X	X
Worst tiredness NRS-diary			X ^k															
Worst tiredness NRS-tablet ^l			X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
WPAI-RA			X		X	X	X	X		X	X	X	X	X	X	X	X	X
Worst joint pain NRS-diary			X ^k															
Worst joint pain NRS-tablet ^l			X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
QIDS-SR16			X					X				X				X	X	X
Healthcare resource use			X			X	X	X		X	X	X	X	X	X	X	X	X
Laboratory tests																		
TSH	X																	
HIV	X																	
Hepatitis C antibody testing	X																	
Hepatitis B testing (HBsAg, HBcAb, HBsAb)	X																	

Visit	Screening		Blinded, Placebo- Controlled Treatment Part A										Open-label Part B				ET	Follow-up Part Ca
	1	1a	2	3	4	5	6	7	8	9	10	11	12	13	14	15		
Week of Treatment			Base-line 0	1	2	4	8	12	14	16	20	24	28	32	40	52	Any week	
Days from Visit 2	-42 to -3 days		0	7 ± 2 days	14 ± 2 days	28 ± 2 days	56 ± 3 days	84 ± 3 days	98 ± 3 days	112 ± 3 days	140 ± 3 days	168 ± 3 days	196 ± 3 days	224 ± 5 days	280 ± 5 days	365 ± 5 days	Any day	28 ± 5 days after last dose
Serum pregnancy test ^m	X		X															
Urine pregnancy test ^m			X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
FSH ⁿ	X																	
Clinical chemistry ^o	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Lipid panel ^p			X					X				X				X		
Lipid NMR ^q			X					X										
Hematology	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
HBV DNA ^r								X				X		X	X	X	X ^s	
Iron studies (iron, TIBC, and ferritin)	X		X					X				X				X		
Urinalysis	X		X					X				X		X		X	X	
Rheumatoid factor	X		X					X				X				X	X	
hsCRP	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Erythrocyte sedimentation rate ^t	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Anti-CCP	X		X					X				X				X	X	
Immunoglobulin levels (IgG, IgA, IgM)	X		X			X		X				X				X	X	
Lymphocyte subsets (T, B, NK, and T-cell subsets)	X		X			X		X				X				X	X	
Exploratory storage samples (serum and plasma)			X	X	X	X		X				X				X		

Visit	Screening		Blinded, Placebo- Controlled Treatment Part A										Open-label Part B				ET	Follow-up Part Ca
	1	1a	2	3	4	5	6	7	8	9	10	11	12	13	14	15		
Week of Treatment			Base-line 0	1	2	4	8	12	14	16	20	24	28	32	40	52	Any week	
Days from Visit 2	-42 to -3 days		0	7 ± 2 days	14 ± 2 days	28 ± 2 days	56 ± 3 days	84 ± 3 days	98 ± 3 days	112 ± 3 days	140 ± 3 days	168 ± 3 days	196 ± 3 days	224 ± 5 days	280 ± 5 days	365 ± 5 days	Any day	28 ± 5 days after last dose
Exploratory storage samples (urine)			X	X	X	X		X				X					X	
Exploratory storage samples (RNA)			X	X	X	X		X				X					X	
DNA collection			X									X						

Abbreviations: Anti-CCP = anticyclic citrullinated peptide; BP = blood pressure; DNA = deoxyribonucleic acid; eGFR = estimated glomerular filtration rate; EQ-5D-5L = European Quality of Life-5 Dimensions-5 Level; ET = early termination; FACIT-F = Functional Assessment of Chronic Illness Therapy Fatigue scale; FSH = follicle-stimulating hormone; HAQ-DI = Health Assessment Questionnaire-Disability Index; 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; hsCRP = high-sensitivity C-reactive protein; Ig = immunoglobulin; IWRS = interactive web-response system; NK = natural killer; NMR = nuclear magnetic resonance; NRS = numeric rating scale; PPD = purified protein derivative; QIDS-SR16 = Quick Inventory of Depressive Symptomatology-Self-Rated (16 items); RA = rheumatoid arthritis; RNA = ribonucleic acid; TB = tuberculosis; TIBC = total iron binding capacity; TSH = thyroid stimulating hormone; WPAI-RA = Work Productivity and Activity Impairment-Rheumatoid Arthritis.

- Those patients who complete the study and do not enroll in extension Study JADY should return for a safety follow-up visit 28 days after their last dose of study drug. Those patients who do enroll in Study JADY do not need to return for a follow-up visit.
- Symptom-directed physical examinations may be conducted at the investigator's discretion at any visit and any time a patient presents with physical complaints.
- Include recording of habits such as caffeine, alcohol, and tobacco consumption.
- Electrocardiograms will be performed locally and will be locally (machine) read.

- e 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 or T-SPOT.TB test (if available). If the QuantiFERON-TB Gold or T-SPOT.TB test results are not negative, the patient will be considered to have latent TB. If the QuantiFERON-TB Gold or T-SPOT.TB test is available and in the judgment of the investigator, preferred as an alternative to the PPD skin test for the evaluation of TB infection, it may be used instead of the PPD TB test. If the QuantiFERON-TB Gold or T-SPOT.TB test is positive, the patient will be considered to have latent TB. If the test is not negative, the test may be repeated once within 2 approximately weeks of the initial value. If the repeat test results are again not negative, the patient will be considered to have latent TB.
- f A chest x-ray will be taken locally at screening unless one has been obtained within the past 6 months (provided the x-ray and/or report are available for review).
- g At Visit 2, the patient will take his/her oral investigational product at the clinic under the supervision of study staff.
- h At Visits 3, 4, and 8, patients should bring their investigational product with them to the study visit, but no tablet counts need to be performed.
- i All the radiographic images of the hands and feet will be submitted to the central reader for reviewing. At screening, radiographic images of the hands and feet acquired within 4 weeks prior to study entry may be submitted to the central reader to confirm eligibility. If a patient undergoes rescreening, radiographic images acquired as part of initial screening and within 10 weeks of randomization may be used.
- j X-rays will only be taken at early termination if the most recent x-ray was more than 12 weeks earlier.
- k Morning joint stiffness, Worst Tiredness, and Worst Joint Pain assessments will be collected using paper patient diary through Week 12 to assess duration, severity of morning joint stiffness, recurrence of stiffness during the day, worst tiredness and worst joint pain.
- l Morning joint stiffness duration, Worst Tiredness, and Worst Joint Pain assessments will be collected using an electronic patient-reported outcomes [ePRO] tablet at each visit.
- m For all women of childbearing potential, a serum pregnancy test (central laboratory) will be performed at Visit 1 and Visit 2. Urine pregnancy tests (local laboratory) will also be performed at Visit 2 and at all subsequent study visits.
- n To confirm postmenopausal status for women ≥ 40 and < 60 years of age who have had a cessation of menses, an FSH test will be performed. Nonchildbearing potential will be defined as an FSH ≥ 40 mIU/mL and a cessation of menses for at least 12 months.
- o Clinical chemistry will include the following values calculated from serum creatinine: calculated creatinine clearance (Cockcroft-Gault equation) and eGFR (calculated using the Modification of Diet in Renal Disease isotope dilution mass spectrometry traceable method).
- p 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, this will not be considered a protocol violation.
- q Lipid NMR tests will not be performed at sites in China. Other participated countries will have this tests.
- r HBV DNA monitoring will be performed in patients who tested positive for HBcAb at screening.
- s To be performed only if previous test was more than 3 months earlier.
- t Erythrocyte sedimentation rate will be performed locally using a kit provided by the sponsor.

Attachment 2. Protocol JAGS Clinical Laboratory Tests

Clinical Laboratory Tests

Hematology^{a,b}

Hemoglobin
 Hematocrit
 Erythrocyte count (RBCs)
 Absolute reticulocyte count
 Mean cell volume
 Mean cell hemoglobin
 Mean cell hemoglobin concentration
 Leukocytes (WBCs)
 Absolute count of
 Neutrophils, segmented
 Neutrophils, juvenile (bands)
 Lymphocytes
 Monocytes
 Eosinophils
 Basophils
 Platelets
 Cell morphology
 CBC, including differential and blood smear

Urinalysis^{a,b}

Color
 Specific gravity
 pH
 Protein
 Glucose
 Ketones
 Blood
 Bilirubin
 Urobilinogen
 Leukocyte esterase
 Nitrite
 Microscopic examination of sedimentⁱ

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

 Stored serum, plasma, urine, mRNA and DNA samples
 for possible exploratory biomarker analyses

Clinical Chemistry^{a,b}

Serum concentrations of
 Sodium
 Potassium
 Total bilirubin
 Direct bilirubin
 Alkaline phosphatase
 Alanine aminotransaminase/serum glutamic
 pyruvic transaminase (ALT/SGPT)
 Aspartate aminotransferase/serum glutamic
 oxaloacetic transaminase (AST/SGOT)
 Blood urea nitrogen
 Creatinine
 Estimated glomerular filtration rate^c
 Calculated creatinine clearance^d
 Uric acid
 Calcium
 Glucose^e
 Albumin
 Creatine kinase
 Total protein

Lipid profile including^e:

total cholesterol,
 HDL-C, LDL-C, and triglycerides
 ApoA1, ApoB

Lipoprotein subfractions (NMR)

Pregnancy Test (females only)^{b,f}

FSH (females only)^g

Other Tests

Hepatitis B surface antigen
 Hepatitis B surface antibody
 Hepatitis B core antibody
 Hepatitis C antibody^h
 Human immunodeficiency virus
 Iron studies (iron, TIBC, and ferritin)
 Thyroid-stimulating hormone
 PPD or QuantiFERON[®] -TB Gold^j or
 T-SPOT[®].TB
 Rheumatoid factor

High-sensitivity C-reactive protein
 ESR^k
 Anticyclic citrullinated peptide
 Immunoglobulins (IgG, IgA, IgM)
 HBV DNA

Footnotes appear on the following page.

Abbreviations: ApoA1 = apolipoprotein A1; ApoB = apolipoprotein B; CBC = complete blood count; CPK = creatine phosphokinase; DNA = deoxyribonucleic acid; ESR = erythrocyte sedimentation rate; FSH = follicle-stimulating hormone; HBV DNA = hepatitis B virus deoxyribonucleic acid; HDL-C = high-density lipoprotein cholesterol; Ig = immunoglobulin; LDL-C = low-density lipoprotein cholesterol; mRNA = messenger ribonucleic acid; NK = natural killer; NMR = nuclear magnetic resonance; PPD = purified protein derivative; RBC = red blood cell; TIBC = total iron binding capacity; WBC = white blood cell.

- a Assayed/calculated by a sponsor-designated laboratory.
- b Unscheduled blood chemistry (including CPK), hematology, and urinalysis panels may be performed at the discretion of the investigator. If tests are done to evaluate laboratory results to resume study drug, samples must be assayed centrally.
- c Estimated glomerular filtration rate for serum creatinine calculated by the central laboratory using the Modification of Diet in Renal Disease isotope dilution mass spectrometry traceable method.
- d The calculated creatinine clearance will be determined by the central laboratory each time a serum creatinine sample is collected using the Cockcroft-Gault equation: Men: $([140 - \text{age}] \times [\text{weight in kg}]) / (72 \times [\text{serum creatinine in mg/dL}])$ or $([140 - \text{age}] \times [\text{weight in kg}]) / (0.814 \times [\text{serum creatinine in } \mu\text{mol/L}])$. Women: $(\text{above formula}) \times 0.85$.
- e Fasting laboratory values for glucose and lipids will be required at baseline, Week 12, Week 24, and Week 52. If a patient attends these visits in a nonfasting state, this will not be considered a protocol violation. These tests may be performed nonfasting at all other visits.
- f For all women of childbearing potential, a serum pregnancy test (central laboratory) will be performed at Visit 1 and both urine (local laboratory) and serum pregnancy (central laboratory) tests will be performed at Visit 2 to determine study eligibility. Urine pregnancy tests (local laboratory) will also be performed at each subsequent study visit per study schedule.
- g To confirm postmenopausal status for women ≥ 40 and < 60 years of age, an FSH test will be performed. Nonchildbearing potential is defined as an FSH ≥ 40 mIU/mL and a cessation of menses for at least 12 months.
- h A positive hepatitis C antibody result will be confirmed with a positive hepatitis C virus result.
- i Microscopic examination of sediment will be performed only if abnormalities are noted on the routine urinalysis.
- j Test will be required at Visit 1 only to determine eligibility of patient for the study. In countries where the QuantiFERON-TB Gold or T-SPOT.TB test is available and, in the judgment of the investigator, medically preferred as an alternative to the PPD test for the evaluation of mycobacterium tuberculosis (MTB) infection, it may be used instead of the PPD test (positive tests excluded) and may be read locally.
- k ESR analysis will be performed by the local laboratory.

Attachment 3. Protocol JAGS 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 designated medical monitor.

Hepatic Monitoring Tests

Hepatic hematology^a

Hemoglobin
Hematocrit
RBC
WBC
Neutrophils, segmented
Lymphocytes
Monocytes
Eosinophils
Basophils
Platelets

Hepatic chemistry^a

Total bilirubin
Direct bilirubin
Alkaline phosphatase
ALT
AST
GGT
CPK

Haptoglobin^a

Hepatic coagulation^a

Prothrombin time
Prothrombin time, INR

Hepatic serologies^{a,b}

Hepatitis A antibody, total
Hepatitis A antibody, IgM
Hepatitis B surface antigen
Hepatitis B surface antibody
Hepatitis B core antibody
Hepatitis C antibody
Hepatitis E antibody, IgG
Hepatitis E antibody, IgM

Anti-nuclear antibody^a

Anti-smooth muscle antibody^a

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

^a Assayed by Lilly-designated or local laboratory.

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

**Attachment 4. Protocol Amendment I4V-CR-JAGS (c)
Summary A Randomized, Double-Blind, Placebo-
Controlled, Phase 3 Study Evaluating the Efficacy and
Safety of Baricitinib in Patients with Moderately to
Severely Active Rheumatoid Arthritis Who Have Had an
Inadequate Response to Methotrexate Therapy**

Overview

Protocol I4V-CR-JAGS has been amended. The new protocol is indicated by amendment (c) and will be used to conduct the study in place of any preceding version of the protocol.

The overall changes and rationale for the changes made to this protocol are as follows:

This protocol amendment clarifies the timepoint of each database lock (DBL), the blindness of each DBL, and the analysis after each DBL.

- The relevant change is in the Section [12.2.11](#).

Besides, the duration for Study JADY is also updated.

- The relevant change is in the Section [7.1](#).

Revised Protocol Sections

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

2. Synopsis

Objectives:

The secondary objectives of the study are to evaluate the efficacy of baricitinib versus placebo as assessed by:

...

- mean Worst Joint Pain NRS in the 7 days prior to Week 12-compared to placebo as collected in diaries

The exploratory objectives in the study include:

...

- change from baseline to Weeks 12, 24 and 52 in Worst Joint Pain NRS assessed at the visit using an ePRO tablet.

Criteria for Evaluation:

Health Outcomes

The following health outcome measures will be administered in this study:

- duration and severity of morning joint stiffness
- recurrence of joint stiffness during the day
- tiredness severity numeric rating scale (Worst Tiredness NRS)
- pain severity numeric rating scale (Worst Joint Pain NRS)
- FACIT-F
- WPAI-RA
- EQ-5D-5L
- Quick Inventory of Depressive Symptomatology Self-Rated-16 (QIDS-SR₁₆)
- healthcare resource utilization

Analyses methods:

Secondary analyses:

...

- Comparison between baricitinib and placebo in mean “Worst Joint Pain” NRS in the 7 days prior to Week 12: An ANCOVA model with treatment, country, joint erosion status, and baseline score (Day 1 on the diaries) in the model will be used to test the treatment difference between baricitinib and placebo in mean “Worst Joint Pain” NRS.

Besides the analyses listed in the objective section, the following analyses will be performed as well.

...

- change from baseline in Worst Joint Pain NRS assessed at the visit using an ePRO tablet.

6. Objective

6.2. Secondary Objectives

The secondary objectives of the study are to evaluate the efficacy of baricitinib versus placebo as assessed by:

...

- mean Worst Joint Pain NRS in the 7 days prior to Week 12 as collected in diaries

6.3. Exploratory Objective

The exploratory objectives in the study include efficacy evaluation of baricitinib as assessed by:

...

- change from baseline to Weeks 12, 24 and 52 in Worst Joint Pain NRS assessed at the visit using an ePRO tablet.

7. Investigational Plan

7.1. Summary of Study Design

Planned enrollment is approximately 288 randomized patients: 144 to receive baricitinib, and 144 to receive placebo. After the Week 52 study visit, eligible patients may proceed to a separate extension study (Study JADY) lasting for up to ~~2 years~~49 months or to the posttreatment follow-up period of this study (Part C).

10. Efficacy, Health Outcome Measures, Safety Evaluations, Sample Collection and Testing, and Appropriateness of Measurements

10.2. Health Outcome Measures

- Severity of joint pain Numeric Rating Scale (Worst Joint Pain NRS):
 - Patient diary: The Worst Joint Pain NRS is a patient-administered, single-item, 11-point horizontal scale anchored at 0 and 10, with 0 representing “no pain” and 10 representing “pain as bad as you can imagine.” Patients rate their joint pain each day (using a paper diary) by selecting the one number that describes their worst level of joint pain.
 - Electronic PRO (tablet): The Worst Joint Pain NRS is a patient-administered, single-item, 11-point horizontal scale anchored at 0 and 10, with 0 representing “no pain” and 10 representing “joint pain as bad as you can imagine.” Patients rate their joint pain by selecting the one number that describes their worst level of joint pain during the past 24 hours.

12. Sample Size and Statistical Methods

12.2. Statistical and Analytical Plans

12.2.7. Efficacy Analyses

Secondary analyses:

7. Comparison between baricitinib and placebo in mean Worst Joint Pain NRS in the 7 days prior to Week 12: An ANCOVA model with treatment, country, joint erosion status, and baseline score (Day 1 on the diaries) in the model will be used to test the treatment difference between baricitinib and placebo in mean Worst Joint Pain NRS in the 7 days prior to Week 12. Refer to the SAP for more details.

Other efficacy and health outcome analyses:

Besides the analyses listed in the objective section, the following analyses will be performed:

...

- change from baseline in Worst Joint Pain NRS assessed at the visit using an ePRO tablet.

12.2.11. Planned Analyses

~~The analysis of primary and secondary endpoints will occur when all patients complete the Week 24 study visit. A final analysis will occur when all patients complete the study. Two database locks are planned for the study. The first database lock and a blinded review of data will occur after all patients complete the Week 24 study visit. The final database lock and unblinded analysis will occur at the end of the study.~~

Attachment 1. Protocol JAGS Study Schedule

Study Schedule, Protocol I4V-CR-JAGS (c)

	Screening		Blinded, Placebo- Controlled Treatment Part A										Open-label Part B				ET	Follow-up Part Ca
Visit	1	1a	2	3	4	5	6	7	8	9	10	11	12	13	14	15		801
Week of Treatment			Base-line 0	1	2	4	8	12	14	16	20	24	28	32	40	52	Any week	
Days from Visit 2	-42 to -3 days		0	7 ± 2 days	14 ± 2 days	28 ± 2 days	56 ± 3 days	84 ± 3 days	98 ± 3 days	112 ± 3 days	140 ± 3 days	168 ± 3 days	196 ± 3 days	224 ± 5 days	280 ± 5 days	365 ± 5 days	Any day	28 ± 5 days after last dose
Informed consent	X																	
Clinical assessments:																		
History	X																	
Physical examination	X																	
Symptom-directed physical examination ^b			X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Previous therapy/previous RA therapy	X																	
Patient demographics	X																	
Height	X																	
Weight	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Waist circumference	X							X				X				X		
Vital signs (BP and pulse)	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Habits ^c	X																	
Electrocardiogram ^d	X																	
Inclusion/exclusion review for entry	X																	
Inclusion/exclusion review for			X															

	Screening		Blinded, Placebo- Controlled Treatment Part A										Open-label Part B					Follow-up Part Ca
Visit	1	1a	2	3	4	5	6	7	8	9	10	11	12	13	14	15	ET	801
Week of Treatment			Base-line 0	1	2	4	8	12	14	16	20	24	28	32	40	52	Any week	
Days from Visit 2	-42 to -3 days		0	7 ± 2 days	14 ± 2 days	28 ± 2 days	56 ± 3 days	84 ± 3 days	98 ± 3 days	112 ± 3 days	140 ± 3 days	168 ± 3 days	196 ± 3 days	224 ± 5 days	280 ± 5 days	365 ± 5 days	Any day	28 ± 5 days after last dose
enrollment																		
Administer PPD or QuantiFERON®-TB Gold or T SPOT® TB teste	X																	
Read PPD or QuantiFERON-TB Golde		X																
Chest x-rayf	X																	
Preexisting conditions	X		X															
Adverse events			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
Randomization			X															
IWRS	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Investigational product dispensed			X ^g			X	X	X		X	X	X	X	X	X			
Investigational products returned and assess compliance				X ^h	X ^h	X	X	X	X ^h	X	X	X	X	X	X	X	X	
Tender/swollen joint count (68/66 joints)	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Patient's assessment of pain visual analog scale			X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Patient global assessment of			X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

Visit	Screening		Blinded, Placebo- Controlled Treatment Part A										Open-label Part B				ET	Follow-up Part Ca
	1	1a	2	3	4	5	6	7	8	9	10	11	12	13	14	15		
Week of Treatment			Base-line 0	1	2	4	8	12	14	16	20	24	28	32	40	52	Any week	
Days from Visit 2	-42 to -3 days		0	7 ± 2 days	14 ± 2 days	28 ± 2 days	56 ± 3 days	84 ± 3 days	98 ± 3 days	112 ± 3 days	140 ± 3 days	168 ± 3 days	196 ± 3 days	224 ± 5 days	280 ± 5 days	365 ± 5 days	Any day	28 ± 5 days after last dose
disease activity																		
Physician global assessment of disease activity			X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Hand and foot x-rays ⁱ	X ⁱ									X	X				X	X ⁱ		
HAQ-DI			X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Morning joint stiffness - diary			X ^k															
Morning joint stiffness - tablet ^l			X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
FACIT-F			X			X	X	X		X	X	X	X	X	X	X	X	X
EQ-5D-5L			X			X	X	X		X	X	X	X	X	X	X	X	X
Worst tiredness NRS-diary			X ^k															
Worst tiredness NRS-tablet ^l			X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
WPAI-RA			X		X	X	X	X		X	X	X	X	X	X	X	X	X
Worst joint pain NRS-diary			X ^k															
Worst joint pain NRS-tablet ^l			X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
QIDS-SR16			X					X				X				X	X	X
Healthcare resource use			X			X	X	X		X	X	X	X	X	X	X	X	X
Laboratory tests																		
TSH	X																	
HIV	X																	
Hepatitis C antibody testing	X																	

	Screening		Blinded, Placebo- Controlled Treatment Part A										Open-label Part B					Follow-up Part Ca
Visit	1	1a	2	3	4	5	6	7	8	9	10	11	12	13	14	15	ET	801
Week of Treatment			Base-line 0	1	2	4	8	12	14	16	20	24	28	32	40	52	Any week	
Days from Visit 2	-42 to -3 days		0	7 ± 2 days	14 ± 2 days	28 ± 2 days	56 ± 3 days	84 ± 3 days	98 ± 3 days	112 ± 3 days	140 ± 3 days	168 ± 3 days	196 ± 3 days	224 ± 5 days	280 ± 5 days	365 ± 5 days	Any day	28 ± 5 days after last dose
Hepatitis B testing (HBsAg, HBcAb, HBsAb)	X																	
Serum pregnancy test ^m	X		X															
Urine pregnancy test ^m			X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
FSH ⁿ	X																	
Clinical chemistry ^o	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Lipid panel ^p			X					X				X				X		
Lipid NMR ^q			X					X										
Hematology	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
HBV DNA ^r								X				X		X	X	X	X ^s	X
Iron studies (iron, TIBC, and ferritin)	X		X					X				X				X		
Urinalysis	X		X					X				X		X		X	X	X
Rheumatoid factor	X		X					X				X				X	X	
hsCRP	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Erythrocyte sedimentation rate ^t	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Anti-CCP	X		X					X				X				X	X	
Immunoglobulin levels (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

	Screening		Blinded, Placebo- Controlled Treatment Part A										Open-label Part B					Follow-up Part Ca
Visit	1	1a	2	3	4	5	6	7	8	9	10	11	12	13	14	15	ET	801
Week of Treatment			Base-line 0	1	2	4	8	12	14	16	20	24	28	32	40	52	Any week	
Days from Visit 2	-42 to -3 days		0	7 ± 2 days	14 ± 2 days	28 ± 2 days	56 ± 3 days	84 ± 3 days	98 ± 3 days	112 ± 3 days	140 ± 3 days	168 ± 3 days	196 ± 3 days	224 ± 5 days	280 ± 5 days	365 ± 5 days	Any day	28 ± 5 days after last dose
Exploratory storage samples (serum and plasma)			X	X	X	X		X				X				X		
Exploratory storage samples (urine)			X	X	X	X		X				X				X		
Exploratory storage samples (RNA)			X	X	X	X		X				X				X		
DNA collection			X									X						

Abbreviations:

...

k Morning joint stiffness, Worst Tiredness, and Worst Joint Pain assessments will be collected using paper patient diary through Week 12 to assess duration, severity of morning joint stiffness, recurrence of stiffness during the day, worst tiredness and worst joint pain.

¹ Morning joint stiffness duration, Worst Tiredness, and Worst Joint Pain assessments will be collected using an electronic patient-reported outcomes [ePRO] tablet at each visit.

Leo Document ID = a7fca142-cda1-478c-a499-5a7a5bcc7c3d

Approver: PPD
Approval Date & Time: 07-Dec-2016 01:25:10 GMT
Signature meaning: Approved

Approver: PPD
Approval Date & Time: 07-Dec-2016 10:21:59 GMT
Signature meaning: Approved