Post-authorisation Safety Study (PASS) Information

Title	Study I4V-MC-B004: A Retrospective Cohort Study to Assess the Long-Term Safety of Baricitinib Compared with Other Therapies Used in the Treatment of Adults with Moderate-to-Severe Rheumatoid Arthritis in the Course of Routine Clinical Care
Version identifier	1.0
Date of last version	02 Nov 2018
EU PAS Register No:	EUPAS25145
Active substance	Baricitinib; ATC code L04AA37
Medicinal product(s):	Olumiant 2-mg and 4-mg film-coated tablets
Product reference:	EU/1/16/1170
Procedure number:	EMEA/H/C/004085
Marketing authorisation holder(s)	Eli Lilly Nederland B.V, Papendorpseweg 83, 3528BJ Utrecht, The Netherlands
Joint PASS	No
Research question and objectives	The goal of this study is to monitor the incidence and nature of key serious infections, MACE, VTE, and malignancies amongst patients exposed long-term to baricitinib compared to patients treated long-term with bDMARDs or cDMARDs. This goal will be achieved through the following specific objectives: 1) To assess and compare the risk of the following aggregate outcomes: serious infections (including herpes zoster) and opportunistic infections (including tuberculosis, *Candida* infections, and progressive multifocal leukoencephalopathy [PML]), major adverse cardiovascular events (MACE), malignancies (including lymphoma and typically virus-induced malignancies such as cervical and many oropharyngeal cancers), and venous thromboembolism (VTE), among patients with long-term exposure to baricitinib compared to similar patients with RA with long-term exposure to other indicated medications. 2) To describe the incidence rates of the following individual outcomes: lymphoma; herpes zoster; opportunistic infections such as tuberculosis, *Candida*, and PML; rhabdomyolysis; myelosuppression (agranulocytosis); hyperlipidaemia (hypercholesterolaemia, hypertriglyceridaemia); gastrointestinal perforations; and evidence of drug-induced liver injury. A secondary objective is to describe the incidence of the above outcomes in very elderly patients (aged ≥75 years old).
Country(-ies) of study	United States
Author	PPD Eli Lilly and Company Lilly Corporate Canter Indianapolis, IN 46285 United States Email: PPD Telephone: PPD

Approval Date: 06-Dec-2018 GMT

Marketing Authorisation Holder

Marketing authorisation holder (MAH)	Eli Lilly Nederland BV, Papendorpseweg 83, 3528BJ Utrecht, The Netherlands
MAH contact person	PPD Eli Lilly and Company Lilly Corporate Center Indianapolis, IN 46285 United States Email: PPD Telephone:PPD

1. Table of Contents

Section **Page**

List of Tables

List of Figures

List of Annexes

2. List of Abbreviations

Term	Definition
AE	adverse event
ATC	anatomical therapeutic chemical
bDMARD	biologic disease-modifying anti-rheumatic drug
BMI	body mass index
cDMARD	conventional disease-modifying anti-rheumatic drug
CFR	Code of Federal Regulations
CI	confidence interval
CPT	current procedural terminology
DMARD	disease-modifying anti-rheumatic drug
ENCePP	European Network of Centres for Pharmacoepidemiology and Pharmacovigilance
ERB	ethical review board
EU PAS	European Union electronic Register of Post-Authorisation Studies
EU RMP	European Union Risk Management Plan
HCPCS	Health Care Common Procedure Coding System
HIPAA	Health Insurance Portability and Accountability Act
HIRD	HealthCore Integrated Research Database℠
HR	hazard ratio
ICD	International Classification of Disease
IRB	Institutional Review Board
IT	Information technology
JAK	Janus kinase
MACE	major adverse cardiovascular events
MICE	multiple imputation by chained equations
MedDRA	Medical Dictionary for Regulatory Activities

MI	myocardial infarction
NDC	National Drug Code
NDI	National Death Index
NMSC	nonmelanoma skin cancer
NNH	number needed to harm
PASS	postauthorisation safety study
PHI	Protected Health Information
PML	progressive multifocal leukoencephalopathy
RA	rheumatoid arthritis
SAP	statistical analysis plan
US	United States
VTE	venous thromboembolism
WHO	World Health Organization

3. Responsible Parties

Eli Lilly Principal Investigator
PPD
Eli Lilly and Company
Lilly Corporate Center
Indianapolis, IN 46285
United States
Email: PPD
Telephone: PPD

4. Abstract

Study I4V-MC-B004: A Retrospective Cohort Study to Assess the Long-Term Safety of Baricitinib Compared with Other Therapies Used in the Treatment of Adults with Moderate-to-Severe Rheumatoid Arthritis in the Course of Routine Clinical Care.

Version: 1.0

Main author: PPD, Eli Lilly and Company

Rationale and Background

Baricitinib is a Janus kinase (JAK) 1/JAK2 selective inhibitor recently approved in Europe and countries in other regions for the treatment of moderate-to-severe rheumatoid arthritis (RA). Data from clinical trials in patients with RA have been evaluated and demonstrate that baricitinib is effective and generally well tolerated; however, the long-term safety profile among patients with RA in routine clinical practice has not been characterised. The purpose of this study is to assess the long-term safety of baricitinib compared with other systemic therapies used in the treatment of adults with moderate-to-severe RA in the course of routine clinical care.

Study I4V-MC-B004 (B004) is intended as a stand-alone evaluation of the safety profile of baricitinib. Observational studies are subject to potential biases, such as misclassification of outcomes or selection bias, which can produce erroneous results and/or affect the generalisability of the study results. An appropriate study design and analysis plan are able to address many potential sources of bias, but confirmation of the study results through replication is a generally accepted indicator of robust, reliable results. Confirmation of the study results from the I4V-MC-B003 Corrona registry, a prospective cohort study with medical (rheumatologist) confirmation of outcomes, with similar results from this study, executed in a different population (i.e., a different data source) will validate the results as well as the generalisability of findings to patients with rheumatoid arthritis outside of the immediate study populations. Such replication represents standard practice within epidemiological research when results may inform substantial decisions (Peng et al. 2006).

Note: This protocol is intended to validate results obtained from the Corrona registry study (I4V-MC-B003); therefore, it replicates the study objectives and analyses of that study as feasible, given the differences in data sources.

Research Question and Objectives

The goal of this study is to monitor the incidence and nature of key serious infections, major adverse cardiovascular events (MACE), venous thromboembolism (VTE), and malignancies amongst patients exposed long-term to baricitinib compared to patients treated long-term with bDMARDs or cDMARDs. This goal will be achieved through the following specific objectives:

1. To assess and compare the risk of the following aggregate outcomes:

- serious infections (including herpes zoster) and opportunistic infections (including tuberculosis, *Candida* infections, and progressive multifocal leukoencephalopathy [PML]),
- major adverse cardiovascular events (MACE),
- malignancies (including lymphoma and typically virus-induced malignancies such as cervical and many oropharyngeal cancers), and
- venous thromboembolism (VTE)

among patients with long-term exposure to baricitinib compared to patients with long-term exposure to other medications indicated for moderate-to-severe RA.

2. To describe the incidence rates of the following individual outcomes: lymphoma; herpes zoster; opportunistic infections such as tuberculosis, *Candida*, and PML; rhabdomyolysis; myelosuppression (agranulocytosis); hyperlipidaemia (hypercholesterolaemia, hypertriglyceridaemia); gastrointestinal perforations; and evidence of drug-induced liver injury.

As a secondary objective, the incidence of the above outcomes will be described in very elderly patients (aged ≥75 years).

Study Design

This study will use a retrospective cohort design to assess the safety of baricitinib among patients with RA using data from a United States (US) electronic health care database, the HealthCore Integrated Research Database℠, or "HIRD", with clinical information on patients with RA.

Study Population

This study will include adult patients with RA enrolled in the US health care database during the period 2018 to 2030. The study will include only new users defined as patients who were not exposed to the same disease-modifying antirheumatic drug (DMARD) within a 12-month period. Those with previous exposure to a non-baricitinib JAK inhibitor will be excluded from analyses of malignancy.

Variables

The following will be ascertained from the health care database: 1) exposure to baricitinib and other RA medications and targeted adverse events; and 2) potential confounders such as demographics, medical history, comorbidities, and health care resource utilisation. Select outcomes will be validated based on clinical data (e.g., medical record) and diagnostic test results from routine clinical care may also be incorporated.

Drug Exposure

Drugs used to treat patients with RA will be considered in 3 exposure groups or cohorts: conventional DMARDs (cDMARDs), biologic medications, and baricitinib. Follow-up time will begin at treatment initiation and continue until an incident event, initiation of another medication, end of study period, disenrolment from the database, or death. When a patient

switches medication, even within an exposure group, a new index date will be assigned. New and continuing user status will be defined at the start of each treatment episode.

Drug exposures will be classified differently for the evaluation of malignancy compared to the other outcomes listed in the objectives to accommodate the prolonged temporal relation expected between the diagnosis of a cancer and exposure to any putative causal agent. Specifically, for malignancy, assignment to exposure groups will be hierarchical: once exposure to a biologic medication occurs, subsequent person-time may not be attributed to the cDMARDs cohort. Similarly, once exposure to baricitinib occurs, time may not be attributed to other cohorts. For outcomes other than malignancy, an "as-treated" approach will be used such that person-time will accrue based simply on the treatment received.

Drugs in the same pharmacological class as baricitinib (JAK inhibitors) will be excluded from all exposure cohorts to prevent the possibility that a potential class effect, should any exist, could mask the existence of an increased risk associated with the use of baricitinib.

Data Source

Data for this study will come from an electronic healthcare database, the HealthCore Integrated Research Database, which contains patient-level administrative claims data and has the capacity to obtain clinical information (e.g., medical records) for the confirmation of selected outcomes. The data source will be confirmed prior to initiating analysis, based on the availability of important variables, for example, drug prescriptions, and size of the data source. Findings from this study are expected to be generalizable to other patient populations.

Study Size

This study anticipates including at least 4000 baricitinib-exposed patients and 4000 biologic DMARD (bDMARD) comparator patients. No minimum size is anticipated for the cDMARD cohort. The final study size will depend on the number of eligible patients available in the database between 2018 – 2030 who are available to contribute time at risk.

Control for Confounding

Calendar-specific propensity-score estimation and matching will be used to control for confounding. The propensity-score model will be finalised before initiating any comparative safety outcome analyses. Use of concomitant cDMARDs will be included as a covariate in statistical models for analyses of nonmalignancy outcomes.

Data Analysis

For all analyses, baricitinib will be the treatment of interest. Comparisons will be made with the bDMARD cohort and the cDMARD cohort where appropriate. Comparison with cDMARD users is intended to permit evaluation of risks associated with baricitinib that might not otherwise be detected by a comparison to biologic medications.

Analyses of Nonmelanoma Skin Cancer and Malignancy Excluding Nonmelanoma Skin Cancer

Baseline demographic and clinical characteristics for each exposure cohort will be examined, as well as the crude incidence rate, for all malignancies excluding nonmelanoma skin cancer (NMSC), and malignancies by type. Cox proportional hazards regression will be used to estimate the hazard ratio in the population matched using propensity scores. Sensitivity analyses are planned to examine the effect of unmeasured confounding, the association between the duration of baricitinib exposure and the risk of malignancy, and the effect of different latency periods. The rate of malignancies will also be evaluated in very elderly patients (aged ≥75 years).

Analyses of Serious Infections (Including Herpes Zoster), Opportunistic Infections (Including Tuberculosis, Candida *infections, and Progressive Multifocal Leukoencephalopathy), Major Adverse Cardiovascular Events, and Venous Thromboembolism*

Exposure time will be classified as baricitinib, cDMARD, or bDMARD medication use, based on the pattern of dispensing of a patient's prescription. Events occurring within 5 half-lives or 30 days, whichever is longer, after the end of a medication prescription will be attributed to that exposure, unless the patient begins a new medication within that time window. After propensity score matching, baseline characteristics will be examined for each exposure cohort and crude incidence rates will be calculated. Initial comparability between the exposure groups will be examined by standardised differences.

Cox proportional hazards regression will be used to compare the adverse event incidence rate between the baricitinib and biologic cohorts and the baricitinib and cDMARDs cohorts. Any variables that remain unbalanced after propensity-score matching, as determined based on review of standardized differences, may also be included in the regression model. Patients will be censored at the end of the study period, 5 half-lives or 30 days, whichever is longer, following discontinuation of a medication, disenrolment from the database, occurrence of an incident event, or death. Several sensitivity analyses are planned. The incidence rate of these outcomes will also be evaluated in very elderly patients (aged ≥75 years).

Descriptive Analyses of Outcomes

Overall incidence rates and rates over time will be calculated separately for outcomes included in the comparative analyses of aggregate events (e.g., tuberculosis in opportunistic infections) and for uncommon outcomes (e.g., agranulocytosis), as feasible. Outcomes that will be monitored and described in this way will include lymphoma; herpes zoster; opportunistic infections such as tuberculosis, *Candida*, and PML; rhabdomyolysis; agranulocytosis; hyperlipidaemia (hypercholesterolaemia, hypertriglyceridaemia); gastrointestinal perforations; and evidence of drug-induced liver injury.

5. Amendments and Updates

Not applicable.

6. Milestones

Milestone	Planned Date
Start of data collection/extraction	5 years after market availability in the US, estimated Q1 2024
End of data collection/extraction	11 years after market availability in the US, estimated Q1 2030
Registration in the EU PAS register	Prior to extraction of data
Study Progress report[a]	Included annually in baricitinib PBRER/PSUR after start of data collection
Interim report	Q1 2027 or after 1000 patients exposed to baricitinib have accrued in the data if this target has not been reached by then.
Final report of study results	Approximately 1 year after the end of data collection (i.e., date from when final analytic dataset is available), estimated Q1 2031

Abbreviations: EU PAS = European Union electronic Register of Post-Authorisation Studies; PBRER/PSUR = Periodic Benefit Risk Evaluation Report/Periodic Safety Update Report; US = United States.

[a] Annual progress reports in the PSUR/PBRER will provide updates, as available in the data, on the number of patients receiving treatment with baricitinib in the database and will include descriptive information on the aggregate safety outcomes of interest.

.

7. Rationale and Background

Rheumatoid arthritis (RA) is a chronic autoimmune inflammatory disease characterised by progressive joint destruction, systemic complications, and reduced survival (Smolen and Steiner 2003; Colmegna et al. 2012). It has a profoundly negative impact on the quality of life of those affected, particularly among those with moderate-to-severe disease (Choy and Panayi 2001; Allaire et al. 2009; Wasserman 2011).

Baricitinib is a Janus kinase (JAK)1/JAK2 selective inhibitor recently approved in Europe for the treatment of moderate-to-severe rheumatoid arthritis (RA). In clinical studies of patients with RA, baricitinib produced clinically meaningful improvements across all relevant domains of efficacy, including signs and symptoms, low disease activity and remission rates, physical function, and patient-reported outcomes, as well as inhibiting progressive radiographic joint damage. Data from clinical trials in patients with RA have been evaluated and demonstrate that baricitinib is effective and generally well tolerated; however, the long-term safety profile among patients with RA in routine clinical practice has not been characterised.

Rheumatoid arthritis is associated with a number of serious comorbidities (CDC 2015 [WWW]). Obesity and smoking are risk factors for RA, consequently, patients with RA have a higher prevalence of these risk factors than age-matched controls (Crowson et al. 2013; Chang et al. 2014). In addition, among patients with RA, there is a high prevalence of comorbidities including myocardial infarction (MI), stroke, malignancy, venous thromboembolism (VTE), infections, hypertension, and gastrointestinal ulcer (Matta et al. 2009; Chung et al. 2014; Dougados et al. 2014). The EU-RMP for baricitinib (v.6.0) describes one important identified risk (herpes zoster) and 9 important potential risks (malignancies [including lymphoma and typically virus-induced malignancies such as cervical and many oropharyngeal cancers], serious and opportunistic infections [including tuberculosis, *Candida* infections, progressive multifocal leukoencephalopathy (PML)], myelosuppression [agranulocytosis], myopathy including rhabdomyolysis, potential for drug-induced liver injury, gastrointestinal perforations, major adverse cardiovascular events (MACE), VTE, and foetal malformation following exposure in utero); this study will provide data from administrative claims on the incidence of these conditions among patients exposed to baricitinib and other RA therapies.

The long-term safety of biologics and other new therapies entering the market is of interest to rheumatologists, regulators, and health care professionals (Ramiro et al. 2017). The majority of clinical trials collect information about safety for 6 to 12 months in a comparative manner against placebo or active control, but the long-term safety of new therapies may remain unclear. This is especially true for outcomes with long latency periods or that occur infrequently such as cancer and cardiovascular disease, but could also include the occurrence of serious infections after prolonged exposure to medications, as may occur when treating a chronic disease such as RA.

Administrative claims data contain information on millions of patients, including patients with RA, and reflect routine clinical practice with medications prescribed at various doses, combinations, and used in diverse patient populations. These data are not subject to recall bias,

include patients who may not be referred to or choose to participate in clinical trials, and can be readily used to investigate potential safety signals, important risks, and missing information. Limitations, particularly related to uncertain diagnostic validity for outcomes and lack of detailed clinical information, can be addressed through linkage to electronic medical records to validate outcomes. This study will serve as an independent population to validate results from the Corrona Rheumatoid Arthritis registry study (I4V-MC-B003) and will thus include patients with RA who are being treated with an approved biologic or conventional agent.

This study is intended as a stand-alone evaluation of the safety profile of baricitinib. Observational studies are subject to potential biases, such as misclassification of outcomes or selection bias, which can produce erroneous results and/or affect the generalisability of the study results. An appropriate study design and analysis plan are able to address many potential sources of bias, but confirmation of the study results through replication is a generally accepted indicator of robust, reliable results. Confirmation of the study results from the I4V-MC-B003 Corrona registry, a prospective cohort study with medical (rheumatologist) confirmation of outcomes, with similar results from this study, executed in a different population (i.e., a different data source) will validate the results as well as the generalisability of findings to patients with RA outside of the immediate study populations. Such replication represents standard, if infrequently applied, practice within epidemiological research when results may inform substantial decisions (Peng et al. 2006).

This protocol outlines how data obtained from this database will be used to evaluate the safety of baricitinib in routine clinical practice.

8. Research Question and Objectives

The goal of this study is to monitor the incidence and nature of key serious infections, MACE, VTE, and malignancies amongst patients exposed long-term to baricitinib compared to patients treated long-term with bDMARDs or cDMARDs. This goal will be achieved through the following specific objectives:

1. To assess and compare risk of the following aggregate outcomes:
    - serious infections (including herpes zoster) and opportunistic infections (including tuberculosis, *Candida* infections, and progressive multifocal leukoencephalopathy [PML]),
    - major adverse cardiovascular events (MACE),
    - malignancies (including lymphoma and typically virus-induced malignancies such as cervical and many oropharyngeal cancers), and
    - venous thromboembolism (VTE)

    among patients with long-term exposure to baricitinib compared to patients with long-term exposure to other medications indicated for moderate-to-severe RA.
2. To describe the incidence rates of the following individual outcomes: lymphoma; herpes zoster; opportunistic infections such as tuberculosis, *Candida* infections, and PML; rhabdomyolysis; myelosuppression (agranulocytosis); hyperlipidaemia (hypercholesterolaemia, hypertriglyceridaemia); gastrointestinal perforations; and evidence of drug-induced liver injury.

As a secondary objective, the incidence of the above outcomes will be described in very elderly patients (aged ≥75 years).

9. Research Methods

9.1. Study Design

This retrospective cohort study will use data from an administrative claims database, the HealthCore Integrated Research Database or HIRD, and confirm selected outcomes based on clinical information. The data sources will include information on patient demographics, RA diagnosis, records of filled prescriptions or administrations of RA treatment, comorbidities, hospitalisations, and medication use, among others.

9.2. Setting

9.2.1. Study Population

The study population will consist of adult patients diagnosed with RA who, during the study period, are exposed to baricitinib or other approved disease-modifying antirheumatic drugs (DMARDs) used to treat RA. The claims-based definition of RA will be defined based on a combination of diagnostic codes and treatment with DMARDs. Detailed algorithms using diagnostic codes, procedures, and/or pharmacy codes to identify the study population will be outlined in a separate Statistical Analysis Plan (SAP). Among all patients exposed to baricitinib, those who meet the eligibility criteria will be identified and included in the analytic dataset. Only new users will be included in the study, defined as patients without prior exposure to DMARDs in the most recent 12-month period prior to index. Some additional or alternate exclusions may be employed for specific analyses.

Inclusion Criteria

1. The patient has 2 RA diagnosis codes from physician encounters on at least 2 separate visits (on different dates) and initiated baricitinib or a different DMARD* (date of treatment initiation defined as index date, described in detail and illustrated in Section 9.2.2),
2. The patient is aged at least 18 years or older on the index date,
3. The patient has continuous medical and prescription drug coverage for a specified minimum duration prior to the index date.

   Primary analyses will include only those patients with at least 12 months of enrolment, and medical and prescription drug coverage prior to index date. A sensitivity analysis using 6 months of enrolment is described in Section 9.7.8.

*Initiation of a DMARD is defined as dispensing of one of the following DMARDs without prior dispensing of the DMARD during the baseline period (at least 12 months, but all available data will be used). Eligible conventional DMARDs include methotrexate, hydroxychloroquine, leflunomide, and sulfasalazine. Eligible bDMARDs include abatacept, adalimumab, certolizumab pegol, etanercept, golimumab, infliximab, rituximab, sarilumab, and tocilizumab. Eligible targeted synthetic DMARDs include tofacitinib. Azathioprine,

cyclosporine, minocycline, and gold are not included due to infrequent use for RA treatment, based on the 2015 American College of Rheumatology Guideline for the treatment of RA (Singh et al. 2016).

Exclusion Criteria

1. The patient has ≥1 filled prescription for a JAK inhibitor prior to the index date.

All patients meeting these criteria will be included in the main analyses and an attrition table generated for each criterion. These inclusion and exclusion criteria are applicable to specific treatment periods and are not applied at the patient level.

For an example that illustrates how the age criterion is applied at the treatment level: a patient who has been enrolled in the health plan since age 16 years and initiates methotrexate monotherapy at age 17 years, adds etanercept at age 18 years, and switches to tofacitinib at age 19 years, will contribute approximately 1 year of person-time to the primary analyses since treatment with etanercept occurs after the patient reaches 18 years of age.

Additional inclusion/exclusion criteria may be applied to select outcomes, such as recurrent infection which requires a study population with history of hospitalised infection.

As the goal of this study is to investigate long-term safety among patients treated with baricitinib, long-term exposure will be based on the longest period of observation available within the study period, that is, a maximum of 11 years. Average enrolment in administrative claims data is usually approximately 3 years, as determined based on the length of time a subject is insured by a specific provider, but patients with chronic diseases such as RA tend to have longer enrolment than others. Based on information for biologic medications, the median time to discontinuation for second and subsequent-line therapies is 12 months (Ogale et al. 2011). Sensitivity analyses will be performed in which the inclusion/exclusion criteria, exposure ascertainment, and outcome identification will vary from that in the primary analyses (e.g., patients exposed to baricitinib and tofacitinib, both JAK inhibitors, will be evaluated to assess potential class effects). Sensitivity analyses are summarized in Section 9.7.8.

9.2.2. Duration of Study

The study will include eligible patients who are enrolled in the administrative claims database and receive baricitinib, cDMARDs or bDMARDs during the study period.

The study will use a new user design in which patients who initiate baricitinib or a different DMARD will be identified. The date of the first prescription fill or dispensing for a study drug will be defined as the "index date", indicating the start of follow-up. The baseline period will be defined as ≥12 months enrolment in the database before the index dispensing. All patients will be required to maintain continuous enrolment, defined as continuous medical and prescription drug coverage with a gap no longer than a prespecified length, that is, 2 days. The end of follow-up will be the earliest of any of the following: occurrence of a study endpoint (varies by study outcome), end of the initial treatment episode, switch to a different DMARD, health plan disenrolment, death or the calendar end of study period. The design is illustrated in Figure 1.

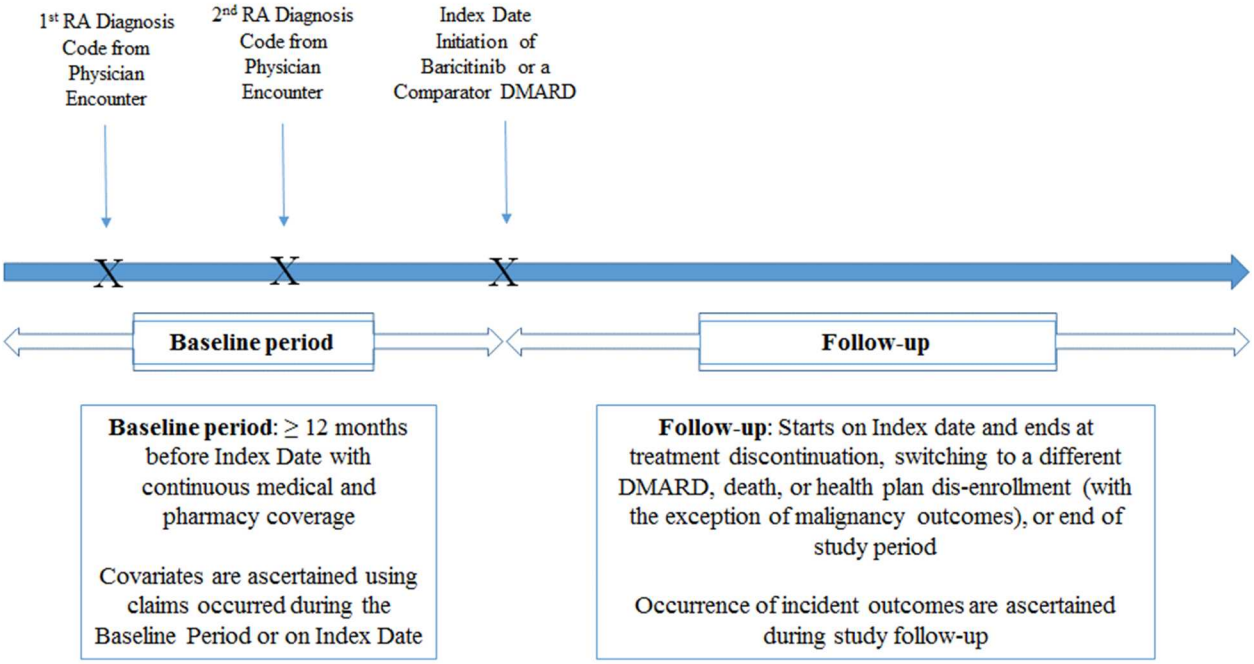

Figure 1. Schematic of study design.

9.2.3. Patient Subgroups of Special Interest

These data will also be used to monitor the incidence of primary outcomes in very elderly patients (aged ≥75 years).

9.3. Variables

Data on patient demographics (age, sex, geographic region): history of prior RA treatment, prior medical history (for example, comorbidities, prior VTE), hospitalisations, concomitant medication use, and health care resource utilisation will be assessed. The following sections detail the variables available for this study. All information will be obtained via administrative claims data (i.e., the HIRD data), with confirmation of selected outcomes based on additional clinical information, as feasible. Details of the clinical review and adjudication process will be described in a forthcoming Medical Record Plan.

Baseline variables will be ascertained from claims occurring in the baseline period (≥12 months) immediately prior to the index date when observation begins and patients begin contributing time-at-risk (e.g., exposed person-time).

9.3.1. Drug Exposure

Exposure to baricitinib or other medications indicated for the treatment of RA will be ascertained based on the National Drug Code or Generic Product Identifier for outpatient pharmacy dispensings and based on Health Care Common Procedure Coding System (HCPCS) for

injections or infusions that occur in a health care setting. Specific applicable codes will be detailed in a separate SAP. Exposure will be defined as a dispensing or administration of a study drug to a patient diagnosed with RA within the period of the study. Table 1 displays medications available at the time of this protocol development. Newly available RA medications will be included as they are approved.

Table 1. Eligible Medications for Identifying Patients Treated for Rheumatoid Arthritis

Conventional Disease-Modifying Antirheumatic Drugs	Biologic Disease-Modifying Antirheumatic Drugs		Targeted Synthetic Disease-Modifying Antirheumatic Drugs
Methotrexate	Abatacept	Golimumab	Tofacitinib
Sulfasalazine	Adalimumab	Infliximab	
Hydroxychloroquine	Certolizumab pegol	Rituximab	
Leflunomide	Etanercept	Sarilumab	
		Tocilizumab	

A patient will be defined as initiating a biosimilar only if the patient has never had prior exposure to the originator drug. Tofacitinib, a JAK inhibitor used to treat patients with moderate-to-severe RA, is excluded from all analyses except for sensitivity analyses to allow for identification of potential class effects. Patients may receive concomitant treatment with biologic medications and cDMARDs; this will be taken into account with time-dependent covariates included in the statistical models. Simultaneous use of biologic medications is expected to be rare because current treatment guidelines do not recommend the use of multiple biologic medications (Singh et al. 2016). The use of glucocorticoids will also be ascertained from the claims, separately for oral and injectable formulations.

The classification of drug exposure within this study for the evaluation of malignancy differs from the classification used for other outcomes to accommodate the long latency of malignant outcomes even after a causal exposure (see Sections 9.3.1.1 and 9.3.1.2, respectively). For malignancy, assignment to exposure groups will be hierarchical. For outcomes other than malignancy, an "as-treated" approach, in which person-time accrues simply based on the treatment received, without regard to a hierarchy, will be used instead.

9.3.1.1. Drug Exposure and Cohort Identification for Malignancy Analysis

For the primary outcome of malignancy, follow-up time will be categorised into 3 hierarchical and mutually exclusive exposure groups or cohorts: cDMARDs, bDMARDs, and baricitinib. To accommodate the long latency that would be expected for malignant outcomes that occur after an exposure, exposure assignment for malignancies will be hierarchical. Once exposure to a biologic medication occurs, time may not be attributed to the cDMARDs cohort and once exposure to baricitinib occurs, time may not be attributed to the other cohorts. This approach is

conservative because it will tend to attribute malignant events to baricitinib, regardless of subsequent exposures and of latent effects of past exposure.

- **cDMARDs cohort**: Biologic-naive cDMARD users with no prior exposure to a JAK inhibitor. Follow-up will begin at treatment initiation and continue until initiation of a biologic medication or baricitinib, end of the study period, disenrolment from the database, incident malignancy, or death; a new index date will be assigned if a patient initiates a biologic medication or baricitinib; new and continuing user status will be updated at each time point to ensure that patients who reinitiate use of a previously used cDMARD are censored.
- **bDMARD or biologic cohort**: Patients initiating a biologic medication with no previous exposure to a JAK inhibitor. Follow-up will begin at treatment initiation and will continue until the initiation of a JAK inhibitor, end of the study period, disenrolment from the database, incident malignancy, or death; a new index date will be assigned if a patient initiates baricitinib; new and continuing user status will be updated at each time point to ensure that patients who reinitiate the use of a previous bDMARD are censored.
- **Baricitinib cohort**: Patients initiating baricitinib with no prior exposure to another JAK inhibitor. Follow-up will begin at treatment initiation and will continue until initiation of another JAK inhibitor, end of the study period, disenrolment from the database, incident malignancy, or death; new and continuing user status will be updated at each time point to ensure that patients who reinitiate baricitinib are censored.

9.3.1.2. Drug Exposure and Cohort Identification for Nonmalignancy Analyses

Unlike malignancy, nonmalignant outcomes are expected to occur in closer temporal proximity to a putatively causal exposure, so it is reasonable for the classification of exposures to reflect changes in treatment regimens. Thus, 3 cohorts similar to those described in Section 9.3.1.1 will be created, but exposure will be classified using an as-treated approach. Using this approach, person-time will accrue based on the treatment received and will reflect actual use during each "medication episode".

- **cDMARDs cohort**: cDMARD users with no prior exposure to a biologic DMARD or a JAK inhibitor. Follow-up will begin at treatment initiation and continue until initiation of a biologic DMARD or a JAK inhibitor, discontinuation of all cDMARDs, end of the study period, disenrolment from the database, occurrence of an incident event, or death; a new index date will be assigned when a patient initiates a new cDMARD.
- **bDMARDs or Biologic cohort**: Patients initiating a biologic medication with no prior exposure to a JAK inhibitor. Follow-up will begin at treatment initiation and will continue until the initiation of a JAK inhibitor/tsDMARD, medication discontinuation, end of study period, disenrolment from the database, occurrence of an incident event, or death; a new index date will be assigned when a patient initiates a new bDMARD; concomitant use of a cDMARD will be assessed and included in the analysis as a time-dependent covariate.
- **Baricitinib cohort**: Patients initiating baricitinib with no prior exposure to another JAK inhibitor. Follow-up will begin at treatment initiation and will continue until initiation of

a bDMARD or other JAK inhibitor/tsDMARD, medication discontinuation, end of study period, disenrolment from the database, occurrence of an incident event, or death; concomitant use of a cDMARD will be assessed and included in the analysis as a time-dependent covariate.

A window corresponding to 5 half-lives or 30 days, whichever is longer, will be added to the end of each treatment period. This window is the ***extension period***. In case the standard interval between medication administration is longer than 5 half-lives (e.g., 183 days between rituximab infusions), the standard interval will be used. During this window, patients will continue to accrue time "at risk" for a brief period after the medication is stopped.

9.3.1.3. Example of Exposure Classification

Patients in the database are under routine clinical care and may add, discontinue, or switch medications during the course of follow-up. This results in switching within the bDMARD cohort and among the 3 medication cohorts. Detailed examples of the attribution of at-risk time are described below.

Each patient will be assigned an index date, when the patient begins contributing person-time for a time-to-event analysis. Information on relevant potential confounding factors is collected at this time, for example, history of infection at baseline. This information is used in the propensity-score matching described in Section 9.7.3 to help ensure that confounding factors are evenly distributed across the groups being compared. When a patient switches medication, a new index date will be assigned, and that patient will be rematched with another initiator from the comparison group. The alternative, for the original patient who switched medication to remain matched to a continuing user of the previous medication, is not appropriate as patients who switch therapies may have different baseline risk regarding the study outcomes than patients who continue their treatment. The use of concomitant cDMARDs will be included as a time-dependent covariate and patients who start or stop concomitant use will not have their cohort observation time censored.

Follow-up time will be measured from the index date to initiation of a medication in another exposure cohort, medication discontinuation, end of study period, disenrolment from the database, occurrence of an incident event (for time-to-event analysis), or death. As described earlier, an extension period corresponding to 5 half-lives will be added to the treatment period to extend the time at risk for that medication. An event occurring during the use of a medication or during the subsequent extension period would be assigned to the discontinued medication. However, if a new medication is initiated within the extension period, the time at risk for the previous medication will end when the new medication is initiated. To account for ambiguities in the attribution of events that occur shortly after a switch, a sensitivity analysis will be conducted in which such events are attributed to the previous medication. Detailed description of the sensitivity analysis can be found in Section 9.7.8.3.

An example of the follow-up time calculation for a patient who switches biologic medications is provided in Figure 2.

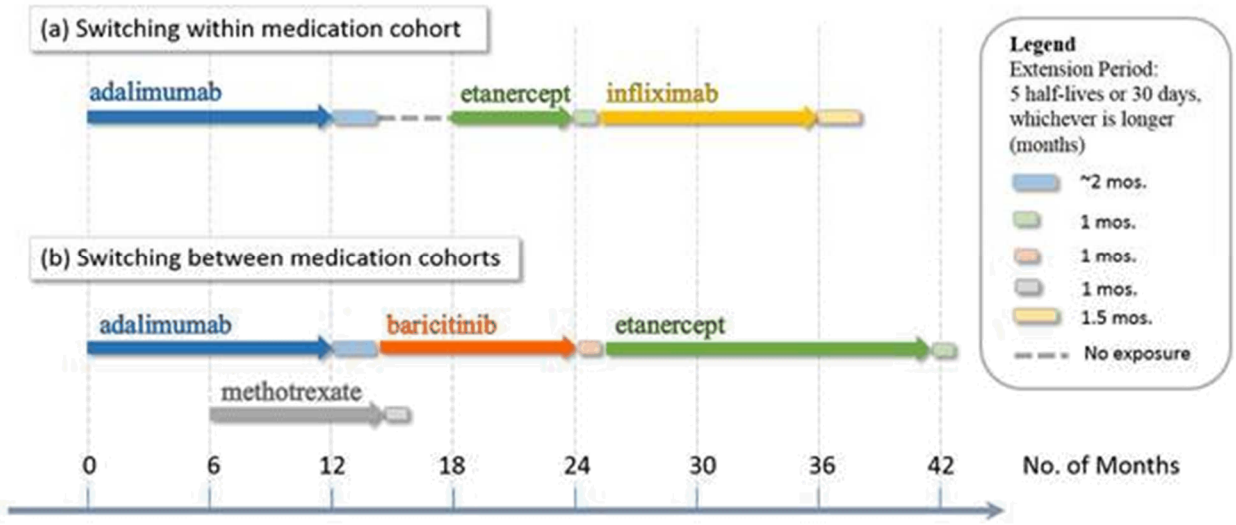

Figure 2. Example of exposure classification.

(a) Switching within medication cohorts

The example in Figure 2(a) illustrates a patient who:

- At the start of the registry, initiates adalimumab treatment and continues until Month 12.
  - Adalimumab time: 12 months + 5 half-lives (70 days or approximately 2 months for the purpose of this example) = 14 months
  - Biologic exposure cohort total time: 14 months
- At Month 18, the patient initiates etanercept and continues until Month 24.
  - Etanercept time: 6 months + 30 days (5 half-lives = 21 days < 30 days) = 7 months
  - Biologic exposure cohort total time: 14 + 7 = 21 months
- At Month 25, the patient initiates infliximab and continues until Month 36.
  - Infliximab time: 11 months + 5 half-lives (45 days or 1.5 months) = 12.5 months
  - Biologic exposure cohort total time: 21 + 12.5 = 33.5 months

By the end of this observation period, this patient will contribute 3 treatment episodes to bDMARDs and 33.5 months of person-time in total to the biologic group.

(b) Switching between medication cohorts

Switching between medication cohorts may also occur and will be managed similarly to the within-medication cohort switch described above. This scenario is illustrated in Figure 2(b).

For this example, assume an analysis for a nonmalignancy outcome so exposures are assigned "as treated" rather than following a hierarchy. Just as for switching within a cohort, this patient:

- At the start of the registry initiates treatment with adalimumab and continues treatment until Month 12.

  - Adalimumab time: 12 months + 5 half-lives (70 days or approximately 2 months for the purpose of this example) = 14 months
  - Biologic exposure cohort: 14 months
  - However, at month 6 the patient initiates concomitant methotrexate and continues treatment until Month 15. Concomitant methotrexate will be included in the analysis as a time-dependent covariate.
- At Month 15, the patient switches to baricitinib and is now included in the baricitinib cohort until Month 24.
  - Baricitinib exposure cohort: 9 months + 30 days (5 half-lives = 2.7 days << 30 days) = 10 months
- At Month 25, the patient initiates etanercept and continues on treatment until the end of the observation period at Month 42.
  - Etanercept time: 17 months + 30 days (5 half-lives = 21 days < 30 days) = 18 months
  - Biologic exposure cohort: 14 + 18 = 32 months

By the end of this observation period, this patient will contribute 2 treatment episodes to bDMARDs and 1 treatment episode to baricitinib.

Had this been an intent-to-treat analysis for malignancy, the patient would have remained in the baricitinib exposure group and would not have been eligible to contribute person-time to the biologic exposure cohort upon switching to etanercept at Month 25. After initiating baricitinib in Month 15, all subsequent person-time would have been contributed to the baricitinib exposure cohort (i.e., "ever exposed" to baricitinib for malignancy outcomes).

9.3.1.3.1. Medication Restarts

In accord with the new user study design (Lund et al. 2015), only patients who newly initiate treatment will be included in the primary analyses. If the number of new users is not sufficient, however, to adequately power the analyses, patients who restart a previous medication, that is, prevalent users, may be included in sensitivity analyses. In this case, additional details will need to be considered, as described below. Sufficient numbers will be evaluated with regard to 80% power to detect at least a two-fold increase in risk among the baricitinib exposure cohort compared to the comparator cohort.

Patients in routine care may stop and restart medications at the discretion of their physician. For instance, patients may be required to discontinue their biologic DMARD temporarily to receive a live vaccine (e.g., herpes zoster vaccine). Another possible cause of periods during which dispensings/administrations of medications are not observed is that patients may be hospitalised and receive medications in the inpatient setting. Medication restarts will not affect the analysis of malignancy outcomes because the analytic approach for malignancies considers "ever exposure" versus "never exposure" to baricitinib. However, medication restarters will be excluded from analyses of nonmalignancy outcomes which employ the new user design to avoid selection bias related to discontinuation and prevalent use of study medications.

Patients who restart medications may have a different baseline risk for the study outcomes than patients who continue treatment or initiate a new treatment. Those who restart a previous treatment may be at reduced risk of adverse events (AEs) compared to others, especially AEs likely to lead to discontinuation, such as serious AEs. Including patients who restart medications is a violation of the new user study design, but it is also not appropriate to simply exclude restarters from an analysis or censor them at the time of discontinuation of their previous treatment episode since restarting a medication may be correlated with an outcome. The regression analyses proposed in the protocol attempt to address this problem by assigning new index dates for patients who restart and rematching them to restarters in the comparison cohort, but in the absence of standard approaches, managing these comparisons appropriately is challenging. If restarters represent only a small fraction of total medication users or if the proportion of restarters is very different across the groups being compared, it may not be possible to match them in the analysis. Any unmatched restarters would be dropped from the analysis. In this scenario, restarters would simply be described by treatment group.

It is also important to compare restarters and non-restarters with regard to the distribution of risk factors for the outcome of interest. If only modest differences exist between these subgroups, no distinction may be needed and the analysis may proceed without the need to consider restarter status. If large differences exist in the distribution of risk factors, though, the total number and relative proportion of restarters across treatment groups will need to be evaluated to determine whether it is feasible to proceed with an analysis that includes both. If it is not feasible, restarters, expected to be the smaller of the two groups, will be described. This exploration of restarters and their potential effect on results will be evaluated in sensitivity analyses only; primary analyses will focus only on incident users.

In real-world practice, nonadherence may be a problem in that patients may not take the prescribed amount of medication at the desired intervals. To distinguish patients who continue their treatment despite gaps from patients who discontinue their medication and restart, the following approach will be applied:

> *Definition of discontinuation*: If the gap between the end of the days-supply of one dispensing and the next dispensing is more than 60 days, the next dispensing will not be incorporated into the treatment period and the patient will be considered to have discontinued treatment.

For medications that are administered at a physician office/infusion centre/hospital, the recommended interval will be used instead of days-supply. For example, the recommended interval between infliximab infusions is 56 days. When constructing the treatment period, if the gap between visits for infusions is longer than 56+60 = 116 days, the second visit will not be incorporated into the treatment period.

Finally, the proportion of patients who have additional dispensings after discontinuation will be assessed and the distribution of the time interval between the date of the last dispensing of a treatment period and the subsequent dispensing of the same medication will be examined. Prior

to initiation of comparative analyses, information from this assessment will be used to confirm the appropriateness of the above definition.

9.3.2. Outcomes

The primary outcomes in this study will be evaluated in analyses comparing the occurrence of events in baricitinib-treated patients to the occurrence in the biologic exposure group and to the occurrence in the cDMARD group: serious infections, MACE, malignancy excluding nonmelanoma skin cancer (NMSC), and VTE. Individual outcomes will be described, but may not be evaluated separately in comparative analyses if the number of events would not provide sufficient statistical power to detect a difference between comparison groups: lymphoma; herpes zoster; opportunistic infections such as tuberculosis, *Candida* infections, and PML; rhabdomyolysis; agranulocytosis; hyperlipidaemia; gastrointestinal perforations; and evidence of drug-induced liver injury. If sufficient number of events are identified for an individual outcome, a formal statistical comparison will be performed. A threshold of an estimated 80% power to detect a two-fold difference in risk will be used to determine whether or not to perform such comparison.

Administrative claims data will be used to identify potential incident cases for each outcome based on International Classification of Disease, 10th Revision (ICD-10) diagnosis and procedure, Current Procedural Terminology (CPT), HCPCS codes, and National Drug Code (NDC) codes. Validated, well-established, claims-based algorithms will be used to identify the study outcomes where available. Alternatively, in the absence of such algorithms, medical records will be used to confirm outcomes, such as VTE, that do not have validated claims-based algorithms or are focal safety concerns (further details are provided under specific outcomes below). When necessary, validation of algorithms will be based on a simple random sample of at least 100 patients identified from the study data using the selected case definition. This will provide sufficient sample size to estimate the performance of the algorithm. The following sections describe each outcome. The codes and algorithms that will be used to identify the outcomes will be detailed in a separate SAP.

9.3.2.1. Malignancy, Excluding Nonmelanoma Skin Cancer

The occurrence of malignancy will be captured based on ICD diagnostic and procedure codes and will include melanoma, solid tumours (lung, cancer, others), and haematologic cancer (lymphoma, others). Validation of incident malignancies will be conducted through linkage to electronic medical records and information included there, such as from pathology reports, and notes from oncologists or from primary care physicians. Nonmelanoma skin cancers are excluded from the analysis to ensure that a signal for malignancies excluding NMSC can be detected, should one exist.

9.3.2.2. Nonmelanoma Skin Cancer

The occurrence of NMSC will be captured based on ICD diagnostic and procedure codes. Validation of NMSCs will follow the same process as for the malignancies described above.

9.3.2.3. Serious Infections

The occurrence of serious infections will be captured based on ICD diagnostic and procedure codes and will include infections requiring hospitalisation and infections requiring parenteral or intravenous antibiotics or equivalent therapy. Validation of incident serious infections will be conducted as deemed appropriate through linkage to electronic medical records and information included there. For example, serious infections in aggregate will not be validated except potentially as a sample for evaluation of validity, but any analyses of specific serious infections (e.g., tuberculosis) will be confirmed with medical record review.

Discussion on how these infections will be analysed is provided in Section 9.7.7.1.

9.3.2.4. Opportunistic Infections

The occurrence of opportunistic infections will be captured as feasible based on ICD diagnostic and procedure codes available in administrative claims data and will be considered as serious infections as defined in Section 9.3.2.3. Prespecified opportunistic infections of interest include infections due to *Cryptococcus neoformans*, *Histoplasma capsulatum*, *Mycobacterium tuberculosis*, *Pneumocystis jirovecii*, and other microorganisms and invasive infections due to *Listeria monocytogenes*, *Listeria*, *Salmonella*, and *Candida*. Details for the analysis of opportunistic infections is provided in Section 9.7.7.2.

9.3.2.5. Major Adverse Cardiovascular Events

Major adverse cardiovascular events are composite cardiovascular endpoints that include fatal and nonfatal MI, fatal and nonfatal ischaemic stroke, and cardiovascular death. Incident MACE, and acute MI as a separate outcome, will be captured based on ICD diagnostic and procedure codes available in administrative claims data. Linkage to the National Death Index or other data, for example, Social Security Administration data, will be performed to capture deaths (all-cause mortality) that occurred before patients were admitted to the emergency room or hospital.

9.3.2.6. Venous Thromboembolism

The occurrence of venous thromboembolic events is captured based on ICD diagnostic and procedure codes available in administrative claims data and includes both pulmonary embolism and deep vein thrombosis. An algorithm based on the validated definition of VTE developed by the Sentinel system will be evaluated for ascertainment of VTE in administrative claims data. The algorithm will be validated in these data through review of medical charts as described in 9.3.2.

9.3.2.7. Evidence of Drug-Induced Liver Injury

The occurrence of events indicative of hepatic injury (i.e., any hepatic event that is serious or requires liver biopsy) will be defined based on ICD diagnostic codes in administrative claims data, with further confirmation evaluated based on the availability of additional clinical information.

9.3.2.8. Rhabdomyolysis

The occurrence of rhabdomyolysis is captured as reported through ICD diagnostic codes. Rhabdomyolysis will be confirmed through review of additional clinical information from available medical charts.

9.3.2.9. Hyperlipidaemia (Including Hypercholesterolaemia and Hypertriglyceridaemia)

The occurrence of hyperlipidaemia will be captured as feasible, based on ICD diagnostic codes found in administrative claims data or new dispensing of antihyperlipidemic medication.

9.3.2.10. Myelosuppression (Agranulocytosis)

The occurrence of events indicative of agranulocytosis will be captured based on ICD diagnostic and procedure codes found in administrative claims data.

9.3.2.11. Gastrointestinal Perforation

The occurrence of events indicative of gastrointestinal perforation will be identified based on an algorithm defined by hospital admissions, and ICD-10-CM/PCS diagnosis and procedure codes recorded in administrative claims data.

9.3.3. Covariates

In addition to information on safety outcomes of interest for this study, information about a variety of other categories will be evaluated including demographics; medical history and comorbidities; and RA disease treatment. The covariates listed in Table 2 will be considered in the analyses for their potential to confound the association between exposure to a medication indicated for RA and the outcomes under investigation. Further explanation is provided in Section 9.7.

Table 2. Baseline Covariates for Consideration in Each Outcome-Specific Analysis

Outcome	Baseline Covariates for Consideration in the Propensity-Score Model
Malignancy (excluding NMSC); NMSC	Age, sex, modified Charlson Comorbidity score, previous biologic medication use in the year prior to index date, health care resource utilisation
Serious and opportunistic infection	Age, sex, diabetes mellitus, chronic lung disease, liver disorder, ischaemic heart disease, previous infections or antibiotic use, glucocorticoid use, previous cDMARD or bDMARD use in the year prior to index date, health care resource utilisation
Major adverse cardiovascular event	Age, sex, history of cardiovascular disease (MI, stroke, unstable angina, hospitalised congestive heart failure, ventricular arrhythmia, cardiovascular revascularisation procedure, coronary artery disease, and transient ischaemic attack), diabetes mellitus, current hypertension, history of hypertension, dyslipidaemia, use of prescription aspirin, use of lipid-lowering agents or antiplatelet agents in the year prior to index date, health care resource utilisation
Venous thromboembolism	Age, sex, history of cancer, cardiovascular disease (hospitalised congestive heart failure, ventricular or atrial arrhythmia), diabetes mellitus, previous VTE, or recent pregnancy; recent hospitalisation, surgery or trauma; use of prescription aspirin, anticoagulants, glucocorticoids, methotrexate; oral

	contraceptives or hormone replacement therapy, previous biologic medication use in the year prior to index date, health care resource utilisation

Abbreviations: bDMARD = Biologic disease-modifying antirheumatic drug; cDMARD = conventional disease-modifying antirheumatic drug; MI = myocardial infarction; NMSC = nonmelanoma skin cancer; RA = rheumatoid arthritis; VTE = venous thromboembolism.

Some covariates included in Study I4V-MC-B003 (the Corrona registry) cannot be ascertained in the claims data, including race, body mass index (BMI), alcohol use, smoking, education, RA disease activity, and disease duration. This limitation is discussed in Section 9.9. While these data cannot be captured, statistical models will adjust for proxies of some of these covariates where available. For example, adjusting for cDMARD use and oral or injectable glucocorticoids which are proxies of disease activity. Quantitative bias analysis methods, as described in Section 9.7.8 will be considered to quantify the potential impact of unmeasured confounding by these factors and by the use of claims-based algorithms, given that these highly specific algorithms have low sensitivity (Desai et al. 2016).

9.4. Data Sources

The study will be conducted using an administrative claims database, the HealthCore Integrated Research Database℠ (HIRD), a large health care database maintained by HealthCore for use in health outcomes and pharmacoepidemiologic research. The HIRD includes longitudinal medical and pharmacy claims data from health plan members across the United States (US). Member enrolment, medical care (professional and facility claims), outpatient prescription drug use, and health care utilisation may be tracked for health plan members in the database dating back to January 2006, with diagnoses recorded in the ICD-10, since October 2015, with dispensing of self-administered medications recorded in the NDC and with medications administered at physician office, hospitals, and outpatient infusion centered recorded in CPT code/HCPCS code. The HealthCore Integrated Research Environment has the ability to link the claims data in the HIRD to complementary data sources, including inpatient and outpatient medical records and national vital statistics records such as the National Death Index (NDI).

The accrual of baricitinib users in the HIRD database will be evaluated regularly. If the projected accrual of baricitinib patients lags or uptake suggests that the HIRD is not the most appropriate choice for the planned analyses, the HIRD may be replaced or possibly supplemented with additional data sources. Such a change in the data source would be documented through an amendment to the protocol. The final selection of the data source(s) will be based on various metrics including timeliness, baricitinib uptake, and capacity for medical record retrieval and linkage to NDI.

9.5. Study Size

Sample size and statistical power calculations were performed using the background incidence rate of MACE from the Corrona Rheumatoid Arthritis Registry. The study size was selected to ensure adequate ability to detect a 1.5-fold increase in the risk of MACE among the

baricitinib-treated cohort relative to patients in the biologic medication cohort. The following assumptions were made for the calculations:

Power and sample size calculations for the primary aggregate endpoints were based on an approach that compares 2 survival curves (Log-rank test) (Lakatos et al. 1988). Based on the background rates observed in claims data for each aggregate outcome, the number of treatment episodes needed to achieve 80% power to detect a hazard ratio (HR) of 1.5 is provided in Table 3. Assuming an equal number of patients in baricitinib and bDMARD cohorts, and a onesided Type I error of 0.05, no more than the indicated number of patients would be required in each cohort. With the inclusion of at least 4000 baricitinib-treated patients and 4000 bDMARD-treated patients, the study will meet each of the sample sizes required to address the outcomes listed in Table 3. Each patient may contribute one or more treatment episodes. The sample size is provided based on an average length of exposed follow-up time of 3 years. Prior to initiation of the analysis, the background incidence rate of MACE will be confirmed using the most recent information. Although the average enrolment duration in an administrative claims database is only a few years, enrolment of new patients exposed to the relevant medication groups is expected to continue throughout the study period.

Table 3. Sample Size Required for 80% Statistical Power to Detect a Hazard Ratio of 1.5, by Aggregate Outcome

Outcome	Incidence Rate[a] (per 100 person-years)	Sample Size[b] (Treatment Episodes)
MACE	2.52	1061
Malignancy	0.72	3625
Serious Infections	1.87	1417
VTE	0.98	2672

a Based on an average 3 years of exposed follow-up time.

b Incidence rates among patients with RA in claims data are from cited rates: MACE (Solomon et al. 2013), malignancy (Askling et al. 2016), serious infection (Accortt et al. 2018), and VTE (Maro et al. 2018).

In order to estimate the duration of the observation period for the study overall, the market uptake of tofacitinib, another JAK inhibitor approved for the treatment of moderate-to-severe RA, was estimated in the HIRD. Tofacitinib was approved in November 2012 and by November 3, 2017, a total of 2011 tofacitinib users were identified who also had ICD diagnoses codes for RA. The number of new users during the study period continued to increase annually as of the date the estimates were obtained. Should the rate of uptake remain unchanged, the total number of tofacitinib new users is expected to exceed 5000 by Year 10 since market availability. Based on this experience with tofacitinib, this study has planned to use 10 years of data following market availability of baricitinib in the US. However, given uncertainties in the uptake of baricitinib, which may accrue faster or slower, a ratio of many to one matching may be

considered to mitigate any risk to the study power in the event that accrual of new baricitinib users lags. Inclusion of additional databases may also be considered, as described in Section 9.4. Either modification may also be considered in the event that the average follow-up is less than the 3 years assumed for Table 3 and suggests there will be insufficient patients to achieve 80% statistical power.

9.6. Data Management

Data will be managed according to the standard procedures required by Eli Lilly and Company ("Lilly") and the study partner with access to the health care data. Specifically, Lilly anticipates that this will include maintenance of datasets and analytic programs on a secure server belonging to the data partner. Procedures for acquisition and abstraction of medical record data will be described in a separate document. Full details relating to data security and quality assurance procedures will be provided in the SAP.

9.6.1. Data Collection and Retention

There is no active enrolment or active follow-up of study subjects, and no data are collected directly from individuals. HealthCore maintains Data Sharing Agreements (DSAs) and Business Associate Agreements with all covered entities who provide data to the HIRD. HealthCore's access, use, and disclosure of Protected Health Information (PHI) are in compliance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule (45 Code of Federal Regulations [CFR] Part 160 and Subparts A and E of Part 164). HealthCore does not access, use, or disclose identifiable PHI unless under a specific waiver of authorization (e.g., a HIPAA Waiver of Authorization from an Institutional Review Board [IRB]). HealthCore accesses the data in a manner that complies with federal and state laws and regulations, including those related to the privacy and security of individually identifiable health information. Datasets and analytic programmes will be kept on a secure server and archived per HealthCore record retention procedures.

Refer to Section 9.3 for information about data to be included in an analytic dataset.

9.6.1.1. Missing Data

Imputation will be considered only for variables that would be used to adjust for potential confounding ("adjusting variables") or for generating propensity scores. Imputation will not be used to account for missing information about exposure to DMARDs, including baricitinib, or about the safety outcomes of interest as this would not be appropriate for the main dependent and independent variables in the Cox regression models. If missing data for a covariate needed for the control of confounding exceed 15% of the cohort size, imputation of the missing values for the adjusting variables will be considered before modelling the data. If imputation is deemed necessary, multiple imputation by chained equations (MICE) will be considered (Royston 2004). This is a multiple multivariate imputation method that is described by van Buuren et al. (1999) and is implemented in Stata (StataCorp LP, College Station, TX) in a series of macros (Royston and White 2011). Other methods may be considered as needed.

9.7. Data Analysis

Analyses will be conducted separately for each outcome and will include descriptive analyses, comparative analyses (where appropriate), and the relevant sensitivity analyses. Propensity scores will be used to address imbalances in the potential confounding factors across the comparison groups that may confound the association between treatment and study outcomes. For all analyses, baricitinib will be the treatment of interest. The cDMARD and biologic cohorts will serve as the reference groups. Comparison with cDMARD users is intended to permit evaluation of any potential risks associated with baricitinib, which might not otherwise be detected by comparison to other biologic medications.

Final analyses will begin at the end of the study period, to coincide approximately with the end of the observation period for the Corrona registry study (I4V-MC-B003).

9.7.1. Analysis Population

The analysis population for all outcomes includes all patients present in the data who meet the eligibility criteria and are members of the drug-exposure groups defined in Section 9.3.1. Subgroup analyses will be performed on patients of special interest if there is sufficient sample size. Sample size and statistical power for subgroup analyses may be limited.

9.7.2. Background and Rationale for Propensity Scores

Drug exposure in pharmacoepidemiologic studies does not occur at random and is the result of patient-, physician-, and health care system-related factors. When these factors are associated with the outcome of interest, comparisons of different drug-exposure cohorts will be confounded due to channelling bias. Propensity scores address the imbalance across drug-exposure cohorts by providing a mechanism to compare patients with concordant baseline risk but discordant exposure (Schneeweiss 2007). For clarity, covariates included in the propensity-score models are also referred to as confounders because they confound the association between exposure and outcome.

9.7.3. Propensity Score Definition and Estimation

A propensity score is an estimate of the probability that a patient receives a particular treatment, conditional on measured characteristics at the time a treatment decision is made (Rosenbaum and Rubin 1983). For this study, a patient's propensity score will reflect the predicted probability of exposure to a medication given a patient's characteristics at the index date. Propensity scores will be estimated using logistic regression models predicting the probability of baricitinib exposure compared with exposures in the other groups (cDMARD users and biologic medication users). These models will be constructed separately for each primary comparison (Section 9.3.2). The models will include variables that are known risk factors for the outcomes of interest and are also associated with systemic treatments for RA. Covariates considered for inclusion in the propensity-score models are provided in Table 2. The inclusion of interaction and nonlinear terms will be guided by clinical judgement. Evaluation of the propensity-score models is discussed in Section 9.7.5.

Newly marketed drugs often experience changes in prescribing patterns over time so that a patient characteristic that was once associated with treatment selection becomes more or less relevant over the drug's life cycle. The time between marketing authorisation and the stabilisation of use patterns, market share, and insurance coverage is a particularly dynamic time in the life cycle of a drug (Schneeweiss et al. 2011). To account for this, a calendar-specific propensity score will be employed as described by Mack et al. (2013) and Seeger et al. (2003). These methods estimate the propensity-score models and match patients within blocks of calendar time to account for temporal variation in prescribing patterns.

9.7.4. Using the Propensity Score to Address Channelling Bias

Each patient in the study will have at least one estimated propensity score that represents the probability of exposure at the index date, given baseline characteristics (new medication starts will be considered a new index date). Matching and stratification on the propensity score is relatively straightforward with 2 exposure groups, but becomes increasingly complex as the number of exposure groups increases. Multiple exposure groups and the limited registry sample available for this study means that matching may result in a high number of unmatched patients and stratification may result in strata with few or no patients. Therefore, this study will examine pairwise comparisons of exposure cohorts: the baricitinib cohort versus the biologic cohort and the baricitinib cohort versus the cDMARDs cohort. For each comparison, the baricitinib exposure cohorts will be matched separately to the comparator exposure cohorts.

Matching will be performed using an objective algorithm and will be discussed further in the SAP. The effectiveness of the matching will be evaluated and the propensity-score model will be adjusted as appropriate. More information on the evaluation of the propensity score is provided in Section 9.7.5. The propensity-score model and matching will be finalised before initiating any safety outcome analyses. If the number of matched patients is prohibitively small, limiting the possibility of conducting a comparative analysis, other applications of the propensity score, such as matching within strata of propensity scores will be considered. One-to-many matching will also be considered if deemed necessary to increase sample size.

9.7.5. Evaluation of the Propensity-Score Model and Stratification

Before initiating any outcome analyses, the ability of the propensity-score stratification to balance the distribution of baseline confounders and reduce channelling bias will be evaluated.

The appropriateness of the propensity-score modelling is judged on whether balance on pretreatment characteristics is achieved between the treatment and reference groups (D'Agostino and D'Agostino 2007; Rubin 2007; Spreeuwenberg et al. 2010). Standardised differences will be used to assess differences between the cohorts across all measured baseline covariates before and after propensity score matching. As a general rule, standardised differences greater than 0.10 indicate an imbalance that may require further investigation (Austin and Mamdani 2006; Austin 2011). Higher-level terms or interactions may be considered when a variable is unbalanced across the baricitinib and reference cohorts or when informed by clinical judgement (e.g., an interaction between age and sex for MACE outcomes).

9.7.6. Malignancy Analyses

Analyses will be performed for malignancy, excluding nonmelanoma skin cancer, and nonmelanoma skin cancer. Malignancies have a long latency period; consequently, they are not easily attributed to a specific exposure. To account for this ambiguity, malignancy analyses will consider the risk of malignancy associated with ever use of baricitinib. These analyses will include only patients initiating treatment, that is, "new users". The incidence of malignancy outcomes will be monitored in very elderly patients (aged ≥75 years).

Proportional hazards of the exposure groups will be evaluated using Kaplan–Meier survival curves. If the proportional hazards assumption is violated, other models that relax the proportional hazards assumption will be considered.

9.7.6.1. Malignancy Excluding Nonmelanoma Skin Cancer

The outcome for malignancy analysis will include all definite malignancies excluding NMSC. All patients with an active malignancy at enrolment will be excluded from these analyses.

Within the propensity score matched population(s), a number of descriptive statistics and rates will be generated to understand the registry data before comparative analyses begin:

- Number of people with past baricitinib or biologic use at baseline
- Baseline demographic and clinical characteristics and standardised differences for the baricitinib, biologic, and cDMARDs cohorts (all patients, matched patients, and unmatched patients)
- Description of RA treatments in each cohort and median duration of exposure of each treatment
- History of malignancy at baseline, based on available data
- Baseline demographic and clinical characteristics for all patients excluded due to history of malignancy at baseline
- Pattern of medication use post-index date for the baricitinib, biologic, and cDMARDs cohorts (all patients, matched patients, and unmatched patients)
- Distribution of survival time for all malignancy outcomes excluding NMSC and malignancy by type for the baricitinib, biologic, and cDMARDs cohorts (all patients, matched patients, and unmatched patients)

After the specified descriptive statistics are calculated, calendar-specific propensity score matching will be used to match patients between cohorts as described in Section 9.7.4. Descriptive statistics requiring matched cohorts will then be conducted. No comparative analyses will begin until finalisation of the exposure cohorts and propensity-score models are achieved.

Cox proportional hazards regression models will be used to estimate the HRs and 95% confidence intervals (CIs) of incident malignancies, excluding NMSC, amongst patients in the baricitinib cohort versus the biologic cohort and the baricitinib cohort versus the cDMARDs cohort (null hypothesis [H_0]: HR = 1). The model will contain the exposure cohort variable, any variables that remain unbalanced after propensity score matching (Table 2), and a time-

dependent variable for any concomitant RA treatment. A sandwich variance estimator will be used to account for the matched data, and methods used to account for within-patient correlation will be examined. Model diagnostics will be performed to identify any influential observations. These analyses may be modified to focus on the class of JAK inhibitors rather than specifically baricitinib, if evaluating "ever exposure" to baricitinib without other JAK inhibitor use is not feasible due to small numbers. However, in the scenario where few patients exist with exposure to baricitinib and no exposure to other JAK inhibitors, descriptive statistics will be provided separately for patients exposed singly to each JAK inhibitor (i.e., those exposed only to baricitinib and those exposed only to each other JAK inhibitor). Sensitivity analyses will be performed accordingly and are discussed in Section 9.7.8.1.

9.7.6.2. Nonmelanoma Skin Cancer

The outcome for this analysis is definite NMSC as defined in Section 9.3.2.2. All patients with an active malignancy at baseline will be excluded from these analyses. In addition to the descriptive statistics and crude rates described in Section 9.7.7, analysis of NMSC will include the following:

- History of NMSC at baseline
- Baseline demographic and clinical characteristics for all patients excluded due to active malignancy
- Distribution of survival time post-index date until NMSC events for the baricitinib, biologic, and cDMARDs cohorts (all patients, matched patients, and unmatched patients)
- Pattern of medication use post index date for the baricitinib, biologic, and cDMARDs cohorts (all patients, matched patients, and unmatched patients)

After descriptive statistics are calculated, calendar-specific propensity scores will be used to match patients between cohorts as described in Section 9.7.4. No comparative analyses will begin until finalisation of the exposure cohorts and propensity-score models are achieved.

Cox proportional hazards regression models will be used to estimate HRs and 95% CIs of newly diagnosed NMSCs among patients in the baricitinib cohort versus the biologic medications cohort and the baricitinib cohort versus the cDMARDs cohort. The model will contain exposure cohort, any variables that remain unbalanced after propensity-score matching (Table 2), and a time-dependent variable to indicate any concomitant RA treatment. A sandwich variance estimator will be used to account for the matched data, and methods to account for within-patient correlation will be examined. Patients will be censored upon development of NMSC; switch to a medication in an alternate exposure cohort; 5 half-lives following discontinuation of a medication; or at the end of the study period, disenrolment from the database, or death. Model diagnostics will be performed to identify any influential observations. These analyses may be modified to focus on the class of JAK inhibitors rather than specifically baricitinib, if evaluating "ever exposure" to baricitinib without also including other JAK inhibitor use is not feasible due to small numbers. However, in the scenario where few patients exist with exposure to baricitinib and no exposure to other JAK inhibitors, descriptive statistics will be provided separately for

patients exposed singly to each JAK inhibitor (i.e., those exposed only to baricitinib and those exposed only to each other JAK inhibitor).

9.7.7. Nonmalignancy Analyses

Analyses will be performed for serious infections, opportunistic infections (including tuberculosis), MACE, and VTE. These analyses will include only patients initiating treatment, that is, "new users." The incidence of non-malignancy outcomes will be monitored in very elderly patients (aged ≥75 years).

Before beginning comparative analyses, a number of descriptive statistics and crude rates will be generated to understand the registry data:

- Baseline demographic and clinical characteristics and standardised differences for the cDMARDs cohort, biologic cohort, and baricitinib cohort (all patients, matched patients, and unmatched patients)
- Prevalence of the outcomes at baseline
- Baseline demographic and clinical characteristics for all patients excluded due to prevalent secondary outcome
- Distribution of follow-up time for the cDMARDs cohort, biologic cohort, and baricitinib cohort (all patients, matched patients, and unmatched patients)
- Distribution of baseline demographic and clinical characteristics and standardised differences for matched patients in the cDMARDs cohort, biologic cohort, and baricitinib cohort by exposure duration
- The number of new medication starts for matched patients within the cDMARDs cohort, biologic cohort, and baricitinib cohort

Comparative analyses will be implemented using calendar-specific propensity-score matching to control for confounding. Model diagnostics will be performed to identify any influential observations. Sensitivity analyses will be performed accordingly. No comparative analyses will begin until finalisation of the exposure cohorts and propensity-score models is achieved. Details of outcome-specific analyses are presented below. Proportional hazards of the exposure groups will be evaluated using Kaplan–Meier survival curves. If the proportional hazards assumption is violated, other models that relax the proportional hazards assumption will be considered.

9.7.7.1. Serious Infections

The outcome of interest in this analysis is first serious infection as defined in Section 9.3.2.3.

In addition to the descriptive statistics and crude rates described in Section 9.7.7, analyses of serious infections will include:

- Pattern of medication use post-index date for the baricitinib, biologic, and cDMARDs cohorts (all patients, matched patients, and unmatched patients)
- Distribution of survival time until first serious infection for the cDMARDs cohort, biologic cohort, and baricitinib cohort (all patients, matched patients, and unmatched patients)

- Crude rate of first serious infection and first serious infection by site of infection and for the prespecified infections (Section 9.3.2.3) per 100 patient-years for the cDMARDs cohort, biologic cohort, and baricitinib cohort (all patients, matched patients, and unmatched patients); within cohorts of matched patients, stratified by concomitant DMARD use
- The distribution of the number of serious infections per patient
- The incidence of serious infections prior to use of baricitinib and after commencing baricitinib treatment

Any patient with an existing serious infection at baseline will be excluded from all analyses, including baseline descriptive statistics, crude rates, and comparative analyses. Propensity scores will be used to match patients from different exposure cohorts as described in Section 9.7.4. Cox proportional hazards regression will be used to compare the rates of first serious infection of the baricitinib and cDMARDs cohorts and the baricitinib and biologic cohorts. All models will include the exposure cohort, concomitant cDMARD use, concomitant glucocorticoid use, and any variables that remain unbalanced after propensity-score matching. A sandwich variance estimator will be used to account for the matched data, and methods used to account for within-patient correlation will be examined. Patients will be censored at the first serious infection, switch to a medication in an alternate exposure cohort, 5 half-lives following discontinuation of a medication, end of the study period, disenrolment from the database, or death.

Another analysis that includes all serious infections will also be performed and is described in Section 9.7.8.5.

9.7.7.2. Opportunistic Infections

The outcome of interest for this analysis is first serious opportunistic infection and first infection irrespective of seriousness for a parallel analysis as defined in Section 9.3.2.4.

In addition to the descriptive statistics and crude rates described in Section 9.7.7, analyses for opportunistic infections will include the following for first serious opportunistic infection and first opportunistic infection regardless of seriousness:

- Pattern of medication use post index date for the baricitinib, biologic, and cDMARDs cohorts (all patients, matched patients, and unmatched patients)
- Distribution of survival time until first opportunistic infection for the cDMARDs cohort, biologic cohort, and baricitinib cohort (all patients, matched patients, and unmatched patients)
- Crude rate of the first opportunistic infection, the first opportunistic infection by type, and the first prespecified opportunistic infection (Section 9.3.2.3) per 100 patient-years for the cDMARDs cohort, biologic cohort, and baricitinib cohort (all patients, matched patients, and unmatched patients), stratified by concomitant DMARD use
- The distribution of the number of opportunistic infections per patient
- The distribution of serious and nonserious opportunistic infections

- The incidence of serious infections, including first and subsequent infections, both prior to use of baricitinib and after commencing baricitinib treatment

Any patient with an existing opportunistic infection at baseline will be excluded from all analyses, including baseline descriptive statistics and crude rates. Propensity scores will be used to match patients as described in Section 9.7.4. Cox proportional hazards regression will be used to compare the rates of first opportunistic infection of the baricitinib and cDMARDs cohorts and the baricitinib and nonbiologic cohorts. All models will include the exposure cohort, concomitant cDMARD use, and any variables that remain unbalanced after propensity-score matching. A sandwich variance estimator will be used to account for the matched data, and methods used to account for within-patient correlation will be examined. Patients will be censored at the first opportunistic infection, switch to a medication in an alternate exposure cohort, 5 half-lives following discontinuation of a medication, at the end of the study period, disenrolment from the database, or death. These analyses will be performed for first serious opportunistic infections and repeated for first opportunistic infections regardless of seriousness.

Proportional hazards of the exposure groups will be evaluated using Kaplan–Meier survival curves. If the proportional hazards assumption is violated, other models that relax the proportional hazards assumption will be considered. Another analysis that includes subsequent opportunistic infections will also be performed and is described in Section 9.7.8.5.

9.7.7.3. Major Adverse Cardiovascular Events

The outcome for this analysis is all incident MACE (Section 9.3.2.5). Because of the potential cardioprotective effect of methotrexate (Marks and Edwards 2012), all comparative analyses for MACE will include a time-dependent covariate to indicate concomitant treatment with methotrexate.

In addition to the descriptive statistics and crude rates described in Section 9.7.7, analyses for MACE will include the following:

- Pattern of medication use post index date for the baricitinib, biologic, and cDMARDs cohorts (all patients, matched patients, and unmatched patients)
- Distribution of survival time until first MACE for the cDMARDs cohort, biologic cohort, and baricitinib cohort (all patients, matched patients, and unmatched patients)
- Crude rate per 100 patient-years of first MACE as a component outcome and by individual event for the biologic cohort and the baricitinib cohort (all patients, matched patients, and unmatched patients) and within cohorts of matched patients, stratified by concomitant DMARD use. For analyses of individual events, fatal MI or stroke will be included both with fatal cardiovascular events and MI/stroke.
- The incidence of nonfatal MACE, both prior to use of baricitinib and after commencing baricitinib treatment

Propensity scores will be used to match patients from different exposure cohorts as described in Section 9.7.4. Cox proportional hazards regression will be used to compare the rates of MACE between the baricitinib and cDMARDs cohorts and the baricitinib and biologic cohorts. All models will include the exposure cohort, concomitant cDMARD use, and any variables that

remain unbalanced after propensity-score matching. A sandwich variance estimator will be used to account for the matched data and methods used to address within-patient correlation will be examined. Patients will be censored at the occurrence of MACE, switch to a medication in an alternate exposure cohort, 5 half-lives following discontinuation of a medication, the end of the study period, disenrolment from the database, or death.

Proportional hazards of the exposure groups will be evaluated using Kaplan–Meier survival curves. If the proportional hazards assumption is violated, other models that relax the proportional hazards assumption will be considered.

9.7.7.4. Venous Thromboembolic Events

The outcome for this analysis is all incident VTEs (Section 9.3.2.6). Because of the recognised risk of VTEs associated with methotrexate (e.g., methotrexate SmPC 2016), analyses will be conducted with inclusion of a time-dependent covariate for concomitant treatment with methotrexate, including for methotrexate dose as appropriate. In addition, because other specific medications (e.g., Cimzia® SmPC 2017; Humira® SmPC 2017) may also have different baseline risk of VTEs, an assessment of incidence amongst users of specific medications used to treat RA will also be conducted. Overall, analyses will focus on the comparative risk of VTEs among baricitinib users versus patients receiving standard treatment with biologic medications.

In addition to the descriptive statistics and crude rates described in Section 9.7.7, analyses for VTEs will include the following:

- Pattern of medication use post index date for the baricitinib, biologic, and cDMARDs cohorts (all patients, matched patients, and unmatched patients)
- Distribution of survival time until first VTE for the cDMARDs cohort, biologic cohort, and baricitinib cohort (all patients, matched patients, and unmatched patients)
- Crude rate of first VTE per 100 patient-years for the biologic cohort and the baricitinib cohort (all patients, matched patients, and unmatched patients) and within cohorts of matched patients, stratified by concomitant methotrexate use

Propensity scores will be used to match patients from different exposure cohorts as described in Section 9.7.4. Cox proportional hazards regression will be used to compare the rates of VTEs of the baricitinib and cDMARDs cohorts (as feasible given the number of events observed) and the baricitinib and biologic cohorts. All models will include the exposure cohort, concomitant cDMARD use, and any variables that remain unbalanced after propensity-score matching. A sandwich variance estimator will be used to account for the matched data, and methods used to address within-patient correlation will be examined. Patients will be censored at the occurrence of VTE, switch to a medication in an alternate exposure cohort, at 5 half-lives following discontinuation of a medication, end of the study period, death or disenrolment.

Proportional hazards of the exposure groups will be evaluated using Kaplan–Meier survival curves. If the proportional hazards assumption is violated, other models that relax the proportional hazards assumption will be considered.

9.7.7.5. Analysis of Individual Outcomes

Several outcomes include aggregated events that are of interest on their own, but that are not expected to occur with sufficient frequency to permit comparative analyses to be conducted. Other infrequent outcomes are not part of the aggregate categories but may also be of interest. Specifically, these outcomes are those listed in the second objective: lymphoma; herpes zoster; opportunistic infections such as tuberculosis, *Candida*, and PML; rhabdomyolysis; agranulocytosis; hyperlipidaemia; gastrointestinal perforations; and evidence of drug-induced liver injury.

Descriptive analyses will be conducted for these outcomes and will include the following:

- Distribution of survival time until first occurrence of the event of interest in the cDMARDs cohort, biologic cohort, and baricitinib cohort (all patients, matched patients, and unmatched patients)
- Crude rate of the first occurrence of the safety event per 100 patient-years for the cDMARDs cohort, biologic cohort, and baricitinib cohort (all patients, matched patients, and unmatched patients)
- Distribution of number of safety events per patient
- Incidence of nonlethal, potentially recurrent events of interest prior to use of baricitinib and after commencing baricitinib treatment

For specific opportunistic infections, herpes zoster, and gastrointestinal perforation, it may be useful to examine crude incidence rates stratifying by concomitant use of glucocorticoids. Other stratifications may also be informative and may be included as appropriate. In addition, for certain outcomes, incidence rates may vary by use of specific medications or classes of medications, for example, gastrointestinal perforations among interleukin-6 inhibitors such as tocilizumab and sarilumab. Thus, it will be useful to estimate crude incidence rates for each bDMARD or subgroup of bDMARDs as feasible considering the number of patients exposed to each medication and the suitability of each outcome. (Ascertainment of some events is likely to be subject to detection bias, and the ability to determine whether differences in rates are due to use of a particular medication or due to differences in clinical surveillance may be limited.) Finally, if the incidence of any outcome in this category occurs more frequently than anticipated, comparative analyses may be considered contingent upon having sufficient statistical power to permit differences in the incidence rates between the baricitinib and the comparator cohort to be detected.

9.7.7.6. Patient Subgroups of Special Interest

The incidence and nature of protocol-defined AEs amongst the very elderly (aged ≥75 years) will be monitored. Baseline characteristics will be described for enrolled patients over the age of 75 years.

9.7.8. Sensitivity Analysis

Several sensitivity analyses will be performed to examine the impact of assumptions on study conclusions. An underlying assumption for all of the analyses presented in this protocol is the absence of unmeasured confounding. It is possible that some of the potential confounding

variables may not be available within the data source. To address this issue, formal quantitative bias analysis methods, such as 1) a rule-out approach, as presented by Delaney and Seeger (2013) and/or 2) probabilistic bias analysis in which estimates are corrected based on plausible distributions of the unmeasured confounder through assessment of simulated datasets as presented by Lash et al. (2014) will be used for comparative analyses to quantify the effect that an unmeasured confounder would have on study results.

All primary analyses will be based on a minimum enrolment duration of 12 months, with continuous medical and prescription drug coverage, prior to index date. In the event that the sample size in the primary analysis of any outcome is not sufficient to provide at least 80% statistical power, an additional sensitivity analysis will be conducted based on 6 months of enrolment and coverage prior to index date.

Additional sensitivity analyses are presented below by outcome.

9.7.8.1. Malignancy, Excluding Nonmelanoma Skin Cancer

9.7.8.1.1. Assessment of the Association between Duration of Baricitinib Exposure and Malignancy Incidence

This sensitivity analysis will consider the risk of malignancy associated with cumulative duration of baricitinib exposure and will be conducted regardless of results from the main malignancy analysis. Cumulative duration of baricitinib exposure will be captured from physician enrolment or follow-up forms and will be calculated for each person in the baricitinib cohort. Exposure time will commence upon baricitinib initiation and will continue until the drug discontinuation date, initiation of another JAK inhibitor, the development of a malignancy, death, disenrolment from the database, or the end of the study follow-up period. Analyses will include crude incidence rates of malignancies excluding NMSC among quintiles of baricitinib exposure. Additionally, a logistic regression will be performed to assess the association of duration of baricitinib exposure with malignancy. The model will include duration of exposure and important confounding variables (Table 2). This analysis may be modified to focus on the class of JAK inhibitors rather than specifically baricitinib, if evaluating "ever exposure" to baricitinib without also including other JAK inhibitor use is not feasible due to small numbers. However, in the scenario where few patients exist with exposure to baricitinib and no exposure to other JAK inhibitors, descriptive statistics will be provided for patients exposed only to baricitinib.

9.7.8.1.2. Assessment of the Association between JAK Inhibitors and Malignancy

If the primary analysis reveals a significant result in favour of either baricitinib or the comparator, a sensitivity analysis that considers JAK inhibitors (e.g., baricitinib and tofacitinib) as a class will be performed. Exposure cohorts will be constructed as described in Section 9.3.1.1 and analysis will proceed as outlined in Section 9.3.1.2. For this analysis, patients treated with JAK inhibitors other than baricitinib will be included. Descriptive statistics will be provided for both baricitinib and other JAK inhibitors, based on 'ever' use of either medication group (or on exclusive use of a medication, should there be >5 exclusive users in either group).

9.7.8.1.3. Assessment of the Association between Baricitinib Exposure and Malignancy, within Periods of Time After Initiating Medication

A sensitivity analysis will be conducted to evaluate the malignancy events that occur only within the first 12 months after initiating medication. This analysis, and others such as within the first 24 months, will ***include*** only those malignancies that occur within the specified period after the index date. The pattern of results from these analyses will aid in the identification of detection bias and potential clustering of events after initiating therapy. Analyses will proceed as described in Section 9.7.6.

9.7.8.1.4. Assessment of the Association between Baricitinib Exposure and Malignancy, Allowing for Various Latency Periods

A sensitivity analysis will be conducted, regardless of the results of the primary analysis, to explore latency periods for the development of malignancies. This analysis will ***exclude*** malignancies that occur within 12 months after treatment initiation. Cancer develops over a lengthy period and the inclusion of events that cannot plausibly be related to a medication exposure will tend to bias results towards the null and reduce the ability to detect an effect, if one were truly to exist. In addition to the first 12 months, the first 24 months, first 48 months, or longer after treatment initiation will also be considered, depending on the sample size available for analysis. By evaluating results of analyses with multiple exclusion windows, these analyses will aid in the identification of events with longer latency.

9.7.8.2. Nonmelanoma Skin Cancer

If the primary NMSC analysis reveals a significant result in favour of either baricitinib or the comparator, a sensitivity analysis that considers the JAK inhibitors baricitinib, tofacitinib, and any others as a class will be performed. Exposure cohorts will be constructed as described in Section 9.3.1.1, and analysis will proceed as outlined in Section 9.7.6.2.

9.7.8.3. Nonmalignancy Outcomes, All

An analysis will be conducted in which a different rule will be used to assign events among patients who switch to a new medication before the at-risk window of the previous medication had ended or within the 30-day window after switching from the first medication (whichever is longer). In this sensitivity analysis, any event that occurs during the at-risk window of the previous medication will be attributed to the previous medication. This analysis will apply to patients switching within the bDMARD cohort or between the bDMARD and baricitinib cohorts. These analysis will only include treatment episodes that reflect incident use and will not include episodes of medication restarts.

9.7.8.4. Serious Infection

9.7.8.4.1. Serious Infections Including Recurrent Infections

An analysis that considers all serious infections, including any recurrent infections, will also be conducted as a sensitivity analysis. Propensity scores will be calculated, and patients will be matched on propensity score as described in Section 9.7.4. Generalised estimating equation negative binomial regression models with a log link will be used to estimate the relative rate and

95% CI for all serious infections between the baricitinib and cDMARDs cohorts and the baricitinib and biologic cohorts. The within-patient association will be accounted for by assuming a first-order autoregressive correlation structure. Any variables that remain unbalanced after propensity-score matching will be included in the model.

9.7.8.4.2. Including Patients with History of Serious Infection

A sensitivity analysis will be conducted including patients with history of serious infection, defined as having one or more hospitalised infections during the baseline period.

9.7.8.5. Opportunistic Infections

9.7.8.5.1. Including Recurrent Opportunistic Infections

An analysis that considers all serious opportunistic infections, including recurrent opportunistic infections, will also be conducted as a sensitivity analysis. Propensity scores will be calculated, and patients will be matched on propensity score as described in Section 9.7.4. Generalised estimating equation negative binomial regression models with a log link will be used to estimate the relative rate and 95% CI for all opportunistic infections between the baricitinib and cDMARDs cohorts and the baricitinib and biologic cohorts. The within-patient association will be accounted for by assuming a first-order autoregressive correlation structure. Any variables that remain unbalanced after propensity-score matching will be included in the model.

9.7.8.5.2. Including Patients with History of Opportunistic Infection

A sensitivity analysis will be conducted including patients with history of opportunistic infection, defined as having one or more hospitalized opportunistic infections during the baseline period.

9.7.8.6. Major Adverse Cardiovascular Events

9.7.8.6.1. Intent-to-Treat Analysis of the Association between Baricitinib Exposure and MACE

A sensitivity analysis to evaluate the occurrence of MACE as a result of atherosclerosis will be conducted. Like malignancy, these events are likely to have a long period of clinical latency before detection and will not be easily attributable to a specific exposure. To account for this ambiguity, this analysis will extend the at-risk period beyond the treatment window. In other words, this will estimate the risk associated with ever use of baricitinib using an intent-to-treat assignment of exposure, as described in Section 9.3.1.1. Analyses will otherwise proceed as described in Section 9.7.7.3.

9.7.8.6.2. Assessment of the Association between Duration of Baricitinib Exposure and MACE Incidence

This sensitivity analysis will consider the risk of MACE associated with cumulative duration of baricitinib exposure and will be conducted regardless of results from the main MACE analysis. Cumulative duration of baricitinib exposure will be captured from physician enrolment or follow-up forms and will be calculated for each person in the baricitinib cohort. Exposure time

will commence upon baricitinib initiation and will continue until the drug discontinuation date, initiation of another JAK inhibitor, the occurrence of MACE, death, disenrolment from the database, or the end of the study follow-up period. Analyses will include crude incidence rates of MACE among quintiles of baricitinib exposure. Additionally, a logistic regression will be performed to assess the association of duration of baricitinib exposure with MACE. The model will include duration of exposure and important confounding variables (Table 2).

9.7.8.7. Venous Thromboembolic Events

An analysis that excludes medications with recognised risk of VTEs according to information in the US prescribing information from the comparator group will also be conducted as a sensitivity analysis. Analyses will proceed as described in Section 9.7.7.4.

9.8. Quality Control

The research team will document the progress and scientific and quality review of all study activities and deliverables (e.g., protocol, data management, data analysis, reports, manuscripts, etc.) in a Quality Log. The Quality Log provides documentation of the major study tasks related to a specific study activity performed by HealthCore to develop and execute the requirements of the protocol. In addition, the Quality Log documents the quality assurance measures performed for each study activity during the conduct of the study. This is necessary to ensure that such communications are appropriately documented, that the most up-to-date versions of relevant documents are readily identifiable, and that affected documents are clearly tracked.

All programming required for study database extraction and creation of the analytic datasets from the HIRD will be performed in accordance with HealthCore Programming Standards. The HealthCore Programming Standards are a set of documents describing data extraction methods that are referenced in HealthCore Standard Operating Procedures and provide a guideline for basic, frequently used terms and definitions and respective coding information to maintain operational consistency. Data validation will occur throughout the data management and analysis process. Data quality checks include, but are not limited to, programming checks by an individual who is not a main programmer for the study, internal dataset consistency, and checks to ensure that Protocol criteria were met. If validation checks are not satisfied then an examination of the problem will be performed on the dataset or datasets in question and the problem resolved. All data validation, quality checks, and resolution of issues identified will be documented in the Project Log.

To help ensure consistency of clinical data collection from medical records, the medical record vendor will be trained on the study's process for medical record acquisition as agreed by HealthCore and Eli Lilly and Company and approved by an IRB. As part of the training, a pilot phase will be conducted to review a sample of charts (i.e., 5 to 10 charts) to help ensure that the trained professionals are accurately collecting the data. If any themes are identified across the trained professionals that are resulting in errors in medical record acquisition, retraining may be requested by HealthCore. Throughout the entire medical record acquisition phase, the vendor will keep a question log that will be exchanged with HealthCore. This log will allow the vendor to ask any questions that may not have been raised during the training and give HealthCore the

opportunity to provide standard answers that will then be shared with the vendor's trained professionals. HealthCore will perform quality checks on the clinical data obtained from the medical records, including any assessments of the clinical data by clinical consultants with expertise in rheumatology, and resolve any errors or discrepancies. HealthCore will integrate the clinical data with the claims analytical file from HIRD and perform quality checks to ensure that all the variables for analysis are correctly included.

Additional details of the quality control process for data collection, analysis, and reporting will be captured in the SAP.

9.9. Limitations of the Research Methods

The current study will use data from a large claims database, the HIRD, linked to electronic medical records to evaluate the safety of patients with RA treated with baricitinib. Because features of RA, such as disease severity or duration, may themselves influence the frequency of some or all of the outcomes of interest for this study, patients treated for RA with biologic medications will be the most appropriate comparison group since RA patients treated with biologics are likely those with moderate-to-severe RA. The use of concomitant cDMARD and other medication received (e.g., intramuscular glucocorticoids), proxies for disease severity or treatment response, will also be adjusted for in the analyses.

9.9.1. Channelling Bias in Observational Studies

Drug exposure in pharmacoepidemiologic studies does not occur completely at random and is a result of patient, physician, and system-related factors. When these factors are associated with the outcome of interest, comparisons of different drug-exposure cohorts will be confounded due to channelling bias. The present study addresses this limitation by applying propensity-score matching, as appropriate for each outcome. Propensity scores address this imbalance by providing a mechanism to compare patients with concordant baseline risk but discordant exposure (Schneeweiss 2007). Calendar-specific matching will be implemented to account for changes in treatment patterns that commonly occur after a new drug enters the marketplace; however, propensity scores are only able to adjust for measured confounders. The possibility of unmeasured confounding and the possible influence on study results will be considered in the final report.

9.9.2. Unmeasured Confounding

Administrative claims lack clinical data on potentially important confounders, such as BMI, disease duration and measures of inflammation and RA severity. For example, patients with consistently high levels of disease activity are managed more aggressively and are also at greater risk of developing untoward cardiovascular events due to high levels of chronic inflammation and use of systemic glucocorticoids. To address this issue, baricitinib-exposed patients will be compared to those exposed to other bDMARDs to increase the comparability of the comparison groups in the study design and to match on propensity score in the data analysis. In addition, concomitant use of cDMARDs and glucocorticoids, which are proxies of disease activity and disease flare, will be adjusted for in the analysis. It is also not possible to ascertain behavioural

and demographic factors from the claims database, such as smoking, alcohol use, and physical activity. The distribution of claims in the PS-matched analytical cohort that pertains to these factors will be evaluated, such as claims indicative of smoking cessation consultation. Quantitative bias analysis methods may also be used to quantify the potential impact of unmeasured confounding due to these factors as described in Section 9.7.8.

9.9.3. Assessing Risk of Outcomes with Long Latency Periods

Outcomes with long latency periods such as malignancy and MACE resulting from atherosclerosis are not easily attributed to a particular drug exposure. To account for this ambiguity, the analysis for malignancy and the sensitivity analysis for MACE consider the risk associated with the use of baricitinib, regardless of subsequent medication changes. Although this approach is considered to be conservative from the standpoint that attribution of events (MACE or malignancy) to baricitinib will not be missed, it ignores the duration of baricitinib exposure and any potential effect of exposure(s) to other systemic medications.

Typically, the risk of malignancy increases with increasing exposure to a known risk factor. This would also be true for MACE related to an proatherosclerotic exposure. If treatment with baricitinib were a risk factor for either outcome, combining patients with varying durations of exposure would tend to bias the measure of association towards the null. To address this issue, an analysis that examines the duration of baricitinib use will be performed.

Another challenge of studying malignancy is the effect of screening on the incidence rate. If the analysis reveals an association between exposure to baricitinib and malignancy or exposure to a JAK inhibitor and malignancy, a sensitivity analysis will be conducted. This will eliminate malignancy cases within the first 12 months of follow-up due to the low likelihood of biologic plausibility and the possibility of detection bias resulting from increased surveillance upon switching to a new medication. Only new initiators at baseline will be considered for the analysis in order to appropriately exclude malignancy cases that occur within 12 months of drug initiation.

In addition, the class of JAK inhibitors is fairly new, and questions remain regarding the risk of baricitinib relative to other JAK medications. If these analyses are able to consider the risk of malignancy or MACE associated with baricitinib, rather than the class of JAK inhibitors, a sensitivity analysis that compares baricitinib to other JAK inhibitors will be conducted. This sensitivity analysis will test the hypothesis that the effects of baricitinib and other JAK inhibitors on the incidence of malignancy/MACE are similar.

Finally, there are several limitations specific to the use of administrative claims data. Because the average length of enrolment in a commercial health plan is relatively short, the data source is less than optimal when used to examine potential causal relationships in which the outcome (malignancy or MACE) takes years to develop. Regardless, patients with longer observation periods will exist in the data to permit an estimate of such an association, should one exist. Another such limitation for analysis of malignancy is that it may not be plausible to distinguish between primary and recurrent cancers of the same type (location).

9.9.4. Assessing Loss to Follow-Up, Adverse Events, and Death

During the course of the observational period, it is expected that some patients will be lost to follow-up as a result of disenrolment from the health plan.

Although an outcome may be missed as a result of a drop out, it is expected that this would occur at random (e.g., change in employment) and should not be associated with exposure status. In addition, linkage to the National Death Index may be sought for patients disenrolled from the health plan to capture disenrolment that occurred as a result of death and the cause of death for these patients, who typically represent a minority of all those who disenroll.

9.9.5. Generalisability

The HealthCore HIRD database includes longitudinal medical and pharmacy claims data from a large number of health plan members across the US. Patients with RA included in the study include patients receiving RA care from rheumatologists and from other specialists (e.g., family practice). It also includes patients located in geographic areas with limited access to specialty care. As a result, the findings will complement those from analyses of the Corrona Registry that consist mainly of patients under rheumatologist care. Patients included in this study may not necessarily represent all adults with RA in the US. As enrolment in the health plans that contribute information to this data source is employment-based, the patients represented will tend to include healthier individuals who are able to remain in the workforce. In addition, owing to the availability of only the 2-mg dose of baricitinib in the US, results of this study may not generalise to the population of patients with RA who are treated with the 4-mg dose.

Although information on very elderly patients ≥75 years of age is available in the HIRD, the very elderly patients included may not be representative of the broader population of patients aged ≥75 years with RA. Very elderly patients in claims data will be those with supplemental insurance coverage or who remain employed and insured with medical and pharmacy benefits included. This group may represent people with higher income and/or education than those who are not included. However, this population may more closely represent very elderly patients who would receive treatment with baricitinib compared to other very elderly patients, for example, those present in Medicare data. Although approximately 12% of patients with RA in the HIRD were 75 years or older, the sample size will limit the analyses to descriptive results.

The large size of the HIRD claims database should help ensure that any results from this study apply to a large portion, if not the majority, of patients with RA. Examination of baseline characteristics and comparison with other data sources containing patients with RA may help to clarify the extent to which results from the registry have external validity. Regardless, findings from this study are expected to have internal validity and provide valuable information about the long-term safety of baricitinib.

9.10. Other Aspects

Not applicable.

10. Protection of Human Subjects

The current study applies epidemiologic research methodologies to medical and pharmacy claims data from a large insured population in the US augmented with information obtained from linked medical records and the NDI. The study is retrospective, analysing data from patient encounters that have already occurred in the health care system.

In order to validate the occurrence of cases identified from the claims data, PHI must be accessed from medical records. Medical records will also be obtained to confirm the occurrence of select outcomes. A HIPAA Waiver of Authorization will be submitted prior to any PHI being identified. As the IRB is independent, HealthCore cannot control the approval or whether there are conditions for the approval.

As described earlier in Section 9.6.1, there is no active enrolment or active follow-up of study subjects, and no data are collected directly from individuals. HealthCore maintains Data Sharing Agreements and Business Associate Agreements with all covered entities who provide data to the HIRD. HealthCore's access, use, and disclosure of PHI are in compliance with the HIPAA Privacy Rule [45 CFR Part 160 and Subparts A and E of Part 164]. HealthCore does not access, use, or disclose identifiable PHI unless under a specific waiver of authorization (e.g., a HIPAA Waiver of Authorization from an IRB). HealthCore accesses the data in a manner that complies with federal and state laws and regulations, including those related to the privacy and security of individually identifiable health information.

At no time during the conduct of this study will HealthCore provide to the sponsor information identifying patients or providers. De-identified aggregated results will be reported to the sponsor and the sponsor will not attempt to re-identify any patients or provider from aggregate data provided for the study.

11. Management and Reporting of Adverse Events/Adverse Reactions

Adverse Events

During the course of secondary use of data in observational research, information pertaining to AEs or suspected adverse drug reactions for an identifiable patient may be discovered during patient chart review. Researchers will include all protocol-defined AEs discovered in the individual patient record/chart associated with baricitinib in the study datasets. The protocol-defined AEs are specified in Section 9.3. Researchers will report any other suspected adverse drug reactions with the attribution explicitly stated in the individual patient records to the appropriate party (e.g., regulators or the Marketing Authorisation Holder) as they would in normal practice as required by applicable laws, regulations, and practices.

12. Plans for Disseminating and Communicating Study Results

Final reports will be submitted to regulatory agencies. The study will also be registered in the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) Registry. Study results may be disseminated via presentation at scientific conferences and/or publication in a peer-reviewed journal.

13. References

Allaire S, Wolfe F, Niu J, LaValley MP, Zhang B, Reisine S. Current risk factors for work disability associated with rheumatoid arthritis: recent data from a US national cohort. *Arthritis Rheum.* 2009;61(3):321-328.

Accortt NA, Lesperance T, Liu M, Rebello S, Trivedi M, Li Y, Curtis JR. Impact of sustained remission on the risk of serious infection in patients with rheumatoid arthritis. *Arthritis Care Res (Hoboken)*. 2018;70(5):679-684.

Askling J, Berglind N, Franzen S, Frisell T, Garwood C, Greenberg JD, Ho M, Holmqvist M, Horne L, Inoue E, Michaud K, Nyberg F, Pappas DA, Reed G, Tanaka E, Tran TN, Verstappen SM, Yamanaka H, Wesby-van Swaay E, Symmons D. How comparable are rates of malignancies in patients with rheumatoid arthritis across the world? A comparison of cancer rates, and means to optimise their comparability, in five RA registries. *Ann Rheum Dis*. 2016;75(10):1789-1796.

Austin PC, Mamdani MM. A comparison of propensity score methods: a case-study estimating the effectiveness of post-AMI statin use. *Stat Med*. 2006;25(12):2084-2106.

Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behav Res*. 2011;46(3):399-424.

Chang K, Yang SM, Kim SH, Han KH, Park SJ, Shin JI1. Smoking and rheumatoid *Arthritis Int J Mol Sci*. 2014;15(12): 22279-22295.

[CDC] Centers for Disease Control and Prevention. Rheumatoid arthritis (RA). 2015. Available at: http://www.cdc.gov/arthritis/basics/rheumatoid.htm. Accessed June 7, 2017.

Choy EH, Panayi GS. Cytokine pathways and joint inflammation in rheumatoid arthritis. *N Engl J Med.* 2001;344(12):907-916.

Chung WS, Peng CL, Lin CL, Chang YJ, Chen YF, Chiang JY, Sung FC, Kao CH. Rheumatoid arthritis increases the risk of deep vein thrombosis and pulmonary thromboembolism: a nationwide cohort study. *Ann Rheum Dis*. 2014;73(10):1774-1780.

[Cimzia SmPC]. Cimzia Summary of Product Characteristics 2017. Available at: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Product_Information/human/001037/WC500069763.pdf. Accessed August 04, 2017.

Colmegna I, Ohata BR, Menard HA. Current understanding of rheumatoid arthritis therapy. *Clin Pharmacol Ther.* 2012;91(4):607-620.

Crowson CS, Matteson EL, Davis JM III, Gabriel SE. Contribution of obesity to the rise in incidence of rheumatoid arthritis. *Arthritis Care Res (Hoboken)*. 2013;65(1):71-77.

D'Agostino RB Jr, D'Agostino RB Sr. Estimating treatment effects using observational data. *JAMA*. 2007;297(3):314-316.

Delaney JAC, Seeger JD. Sensitivity analysis. In: Velentgas P, Dreyer NA, Nourjah P, Smith SR, Torchia MM, editors. Developing a protocol for observational comparative effectiveness research: a user's guide. Rockville, MD: Agency for Healthcare Research and Quality (US); 2013. Chap 11:p 145-160.

Desai RJ, Solomon DH, Shadick N, Iannaccone C, Kim SC. Identification of smoking using Medicare data- A validation study of claims-based algorithms. Pharmacoepidemiology and drug safety. 2016;25(4):472-475.

Dougados M, Soubrier M, Antunez A, Balint P, Balsa A, Buch MH, Casado G, Detert J, El-Zorkany B, Emery P, Hajjaj-Hassouni N, Harigai M, Luo SF, Kurucz R, Maciel G, Mola EM, Montecucco CM, McInnes I, Radner H, Smolen JS, Song YW, Vonkeman HE, Winthrop K, Kay J. Prevalence of comorbidities in rheumatoid arthritis and evaluation of their monitoring: results of an international, cross-sectional study (COMORA). *Ann Rheum Dis*. 2014;73(1):62-68.

Humira [SmPC]. Humira Summary of Product Characteristics 2017. Available at: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Product_Information/human/000481/WC500050870.pdf. Accessed August 04, 2017.

Lakatos E. Sample sizes based on the log-rank statistic in complex clinical trials. *Biometrics*. 1988;44(1):229-241.

Lash TL, Fox MP, MacLehose RF, Maldonado G, McCandless LC, Greenland S. Good practices for quantitative bias analysis. *Int J Epidemiol*. 2014;43(6):1969-1985.

Lund JL, Richardson DB, Sturmer T. The active comparator, new user study design in pharmacoepidemiology: historical foundations and contemporary application. *Curr Epidemiol Rep.* 2015;2(4):221-228.

Mack CD, Glynn RJ, Brookhart MA, Carpenter WR, Mayer AM, Sandler RS, Sturmer T. Calendar time-specific propensity scores and comparative effectiveness research for stage III colon cancer chemotherapy. *Pharmacoepidemiol Drug Saf*. 2013;22(8):810-818.

Marks JL, Edwards CJ. Protective effect of methotrexate in patients with rheumatoid arthritis and cardiovascular comorbidity. *Ther Adv Musculoskelet Dis*. 2012;4(3):149-157.

Maro JC, Menzin T, Hornbuckle K, Giles JT, Kavanaugh A, Dorner T, Martin D, Huang T, Salinas CA. SAT0140 Risk of venous thromboembolism in rheumatoid arthritis patients treated with biologic and non-biologic dmards. *Ann Rheum Dis.* 2018;77:932.

Matta F, Singala R, Yaekoub AY, Najjar R, Stein PD. Risk of venous thromboembolism with rheumatoid arthritis. *Thromb Haemost.* 2009;101(1):134-138.

Methotrexate [SmPC]. Methotrexate Summary of Product Characteristics 2016. Available at: http://www.medicines.org.uk/emc/print-document?documentId=12034. Accessed August 04, 2017.

Peng RD, Dominici F, Zeger SL. Reproducible epidemiologic research. *Am J Epidemiol*. 2006;163(9):783-789.

Ramiro S, Sepriano A, Chatzidionysiou K, Nam JL, Smolen JS, van der Heijde D, Dougados M, van Vollenhoven R, Bijlsma JW, Burmester GR, Scholte-Voshaar M, Falzon L, Landewé RBM. Safety of synthetic and biological DMARDs: a systematic literature review informing the 2016 update of the EULAR recommendations for management of rheumatoid arthritis. *Ann Rheum Dis*. 2017;76(6):1101-1136.

Rosenbaum PR, Rubin, DB. The central role of the propensity score in observational studies for causal effects. *Biometrika*. 1983;70(1):41-55.

Royston P. Multiple imputation of missing values. *Stata J*. 2004;4(3):227-241.

Royston P, White I. Multiple imputation by chained equations (MICE): implemented in stata. *J Stat Soft*. 2011;45(4):1-20.

Rubin DB. The design versus the analysis of observational studies for causal effects: parallels with the design of randomized trials. *Stat Med*. 2007;26(1):20-36.

Schneeweiss S. Developments in post-marketing comparative effectiveness research. *Clin Pharmacol Ther*. 2007;82(2):143-156.

Schneeweiss S, Gagne JJ, Glynn RJ, Ruhl M, Rassen JA. Assessing the comparative effectiveness of newly marketed medications: methodological challenges and implications for drug development. *Clin Pharmacol Ther*. 2011;90(6):777-790.

Seeger JD, Walker AM, Williams PL, Saperia GM, Sacks FM. A propensity score-matched cohort study of the effect of statins, mainly fluvastatin, on the occurrence of acute myocardial infarction. *Am J Cardiol*. 2003;92(12):1447-1451.

Singh JA, Saag KG, Bridges SL Jr, Aki EA, Bannuru RR, Sullivan MC, Vaysbrot E, McNaughton C, Osani M, Shmerling RH, Curtis JR, Furst DE, Parks D, Kavanaugh A, O'Dell J, King C, Leong A, Matteson EL, Schousboe JT, Drevlow B, Ginsberg S, Grober J, St Clair EW, Tindall E, Miller AS, McAlindon T. 2015 American College of Rheumatology Guideline for the Treatment of Rheumatoid Arthritis. *Arthritis Care Res (Hoboken)*. 2016;68(1):1-25.

Smolen J, Steiner G. Therapeutic strategies for rheumatoid arthritis. *Nat Rev Drug Discov*. 2003;2(6):473-488.

Solomon DH, Curtis JR, Saag KG, Lii J, Chen L, Harrold LR, Herrinton LJ, Graham DJ, Kowal MK, Kuriay B, Liu L, Griffin MR, Lewis JK, Rassen JA. Cardiovascular risk in rheumatoid arthritis: comparing TNF-α blockade with nonbiologic DMARDs. *Am J Med*. 2013;126(8):730.e9-730.e17.

Spreeuwenberg MD, Bartak A, Croon MA, Hagenaars JA, Busschback JJ, Andrea H, Twisk J, Stijnen T. The multiple propensity score as control for bias in the comparison of more than two treatment arms: an introduction from a case study in mental health. *Med Care*. 2010;48(2):166-174.

Tamariz L, Harkins T, Nair V. A systematic review of validated methods for identifying venous thromboembolism using administrative and claims data. *Pharmacoepidemiol Drug Saf*. 2012;21 (Suppl 1):154-162.

Thode HC. Power and sample size requirements for tests of differences between two Poisson rates. *Statistician (J R Stat Soc)*. 1997;46(2):227-230.

van Buuren S, Boshuizen HC, Knook DL. Multiple imputation of missing blood pressure covariates in survival analysis. *Stat Med*. 1999;18(6):681-694.

Wasserman AM. Diagnosis and management of rheumatoid arthritis. *Am Fam Physician.* 2011;84(11):1245-1252.

Annex 1. List of Stand-Alone Documents

Not applicable.

Annex 2. ENCePP Checklist for Study Protocols

Study title:
Study I4V-MC-B004: A Retrospective Cohort Study to Assess the Long-Term Safety of Baricitinib Compared with Other Therapies Used in the Treatment of Adults with Moderate-to-Severe Rheumatoid Arthritis in the Course of Routine Clinical Care
Study reference number:

Section 1: Milestones	**Yes**	**No**	**N/A**	**Page Number(s)**
1.1 Does the protocol specify timelines for				
1.1.1 Start of data collection[1]	☒	☐	☐	17
1.1.2 End of data collection[2]	☒	☐	☐	17
1.1.3 Study progress report(s)	☐	☐	☒	
1.1.4 Interim progress report(s)	☐	☒	☐	
1.1.5 Registration in the EU PAS register	☒	☐	☐	1, 17
1.1.6 Final report of study results.	☒	☐	☐	17

Comments:

Section 2: Research question	**Yes**	**No**	**N/A**	**Page Number(s)**
2.1 Does the formulation of the research question and objectives clearly explain:				
2.1.1 Why the study is conducted? (e.g. to address an				

[1] Date from which information on the first study is first recorded in the study dataset or, in the case of secondary use of data, the date from which data extraction starts.
[2] Date from which the analytical dataset is completely available.

Section 2: Research question	**Yes**	**No**	**N/A**	**Page Number(s)**
important public health concern, a risk identified in the risk management plan, an emerging safety issue)	☒	☐	☐	20
2.1.2 The objective(s) of the study?	☒	☐	☐	20
2.1.3 The target population? (i.e. population or subgroup to whom the study results are intended to be generalised)	☒	☐	☐	21
2.1.4 Which formal hypothesis(-es) is (are) to be tested?	☒	☐	☐	20
2.1.5 If applicable, that there is no *a priori* hypothesis?	☐	☐	☒	

Comments:

Section 3: Study design	**Yes**	**No**	**N/A**	**Page Number(s)**
3.1 Is the study design described? (e.g. cohort, case-control, randomised controlled trial, new or alternative design)	☒	☐	☐	24
3.2 Does the protocol specify the primary and secondary (if applicable) endpoint(s) to be investigated?	☒	☐	☐	30-32
3.3 Does the protocol describe the measure(s) of effect? (e.g. relative risk, odds ratio, deaths per 1000 person-years, absolute risk, excess risk, incidence rate ratio, hazard ratio, number needed to harm (NNH) per year)	☒	☐	☐	34-44

Comments:

Section 4: Source and study populations	**Yes**	**No**	**N/A**	**Page Number(s)**

Section 4: Source and study populations	**Yes**	**No**	**N/A**	**Page Number(s)**
4.1 Is the source population described?	☒	☐	☐	21
4.2 Is the planned study population defined in terms of:				
4.2.1 Study time period?	☒	☐	☐	21
4.2.2 Age and sex?	☒	☐	☐	21
4.2.3 Country of origin?	☒	☐	☐	21
4.2.4 Disease/indication?	☒	☐	☐	21
4.2.5 Co-morbidity?	☒	☐	☐	19
4.2.6 Seasonality?	☐	☐	☒	
4.3 Does the protocol define how the study population will be sampled from the source population? (e.g. event or inclusion/exclusion criteria)	☒	☐	☐	21

Comments:

Section 5: Exposure definition and measurement	**Yes**	**No**	**N/A**	**Page Number(s)**
5.1 Does the protocol describe how exposure is defined and measured? (e.g. operational details for defining and categorising exposure)	☒	☐	☐	24-30
5.2 Does the protocol discuss the validity of exposure measurement? (e.g. precision, accuracy, prospective ascertainment, exposure information recorded before the outcome occurred, use of validation sub-study)	☒	☐	☐	24-30
5.3 Is exposure classified according to time windows? (e.g. current user, former user, non-use)	☒	☐	☐	24-30
5.4 Is exposure classified based on biological mechanism of action and taking into account the pharmacokinetics and pharmacodynamics of the drug?	☒	☐	☐	24-30

Section 5: Exposure definition and measurement	**Yes**	**No**	**N/A**	**Page Number(s)**
5.5 Does the protocol specify whether a dose-dependent or duration-dependent response is measured?	☒	☐	☐	43

Comments:

Section 6: Endpoint definition and measurement	**Yes**	**No**	**N/A**	**Page Number(s)**
6.1 Does the protocol describe how the endpoints are defined and measured?	☒	☐	☐	30-32
6.2 Does the protocol discuss the validity of endpoint measurement? (e.g. precision, accuracy, sensitivity, specificity, positive predictive value, prospective or retrospective ascertainment, use of validation sub-study)	☒	☐	☐	30-32

Comments:

Section 7: Confounders and effect modifiers	**Yes**	**No**	**N/A**	**Page Number(s)**
7.1 Does the protocol address known confounders? (e.g. collection of data on known confounders, methods of controlling for known confounders)	☒	☐	☐	44
7.2 Does the protocol address known effect modifiers? (e.g. collection of data on known effect modifiers, anticipated direction of effect)	☐	☒	☐	

Comments:

Section 8: Data sources	**Yes**	**No**	**N/A**	**Page Number(s)**
8.1 Does the protocol describe the data source(s) used in the study for the ascertainment of:				
8.1.1 Exposure? (e.g. pharmacy dispensing, general practice prescribing, claims data, self-report, face-to-face interview, etc.)	☒	☐	☐	33
8.1.2 Endpoints? (e.g. clinical records, laboratory markers or values, claims data, self-report, patient interview including scales and questionnaires, vital statistics, etc.)	☒	☐	☐	33
8.1.3 Covariates?	☒	☐	☐	32, 33
8.2 Does the protocol describe the information available from the data source(s) on:				
8.2.1 Exposure? (e.g. date of dispensing, drug quantity, dose, number of days of supply prescription, daily dosage, prescriber)	☒	☐	☐	33
8.2.2 Endpoints? (e.g. date of occurrence, multiple event, severity measures related to event)	☒	☐	☐	33
8.2.3 Covariates? (e.g. age, sex, clinical and drug use history, co-morbidity, co-medications, life style, etc.)	☒	☐	☐	32
8.3 Is a coding system described for:				
8.3.1 Diseases? (e.g. International Classification of Diseases (ICD)-10)	☐	☐	☒	
8.3.2 Endpoints? (e.g. Medical Dictionary for Regulatory Activities (MedDRA) for adverse events)	☐	☐	☒	
8.3.3 Exposure? (e.g. WHO Drug Dictionary, Anatomical Therapeutic Chemical (ATC)Classification System)	☒	☐	☐	24
8.4 Is the linkage method between data sources described? (e.g. based on a unique identifier or other)	☒	☐	☐	33

Comments:

Section 9: Study size and power	**Yes**	**No**	**N/A**	**Page Number(s)**
9.1 Is sample size and/or statistical power calculated?	☒	☐	☐	33

Comments:

Section 10: Analysis plan	**Yes**	**No**	**N/A**	**Page Number(s)**
10.1 Does the plan include measurement of excess risks?	☐	☒	☐	
10.2 Is the choice of statistical techniques described?	☒	☐	☐	35-47
10.3 Are descriptive analyses included?	☒	☐	☐	35-47
10.4 Are stratified analyses included?	☒	☐	☐	35-47
10.5 Does the plan describe methods for adjusting for confounding?	☒	☐	☐	35-47
10.6 Does the plan describe methods addressing effect modification?	☒	☐	☐	35-47

Comments:

Section 11: Data management and quality control	**Yes**	**No**	**N/A**	**Page Number(s)**
11.1 Is information provided on the management of missing data?	☒	☐	☐	35
11.2 Does the protocol provide information on data storage? (e.g. software and IT environment, database maintenance and anti-fraud protection, archiving)	☐	☐	☒	
11.3 Are methods of quality assurance described?	☒	☐	☐	47

Section 11: Data management and quality control	**Yes**	**No**	**N/A**	**Page Number(s)**
11.4 Does the protocol describe possible quality issues related to the data source(s)?	☒	☐	☐	47
11.5 Is there a system in place for independent review of study results?	☒	☐	☐	

Comments:

Lilly has a process to ensure independent review of study results produced by this protocol.

Section 12: Limitations	**Yes**	**No**	**N/A**	**Page Number(s)**
12.1 Does the protocol discuss:				
12.1.1 Selection biases?	☒	☐	☐	35-47
12.1.2 Information biases? (e.g. anticipated direction and magnitude of such biases, validation sub-study, use of validation and external data, analytical methods)	☒	☐	☐	35-47
12.2 Does the protocol discuss study feasibility? (e.g. sample size, anticipated exposure, duration of follow-up in a cohort study, patient recruitment)	☒	☐	☐	21
12.3 Does the protocol address other limitations?	☒	☐	☐	48-51

Comments:

Section 13: Ethical issues	**Yes**	**No**	**N/A**	**Page Number(s)**
13.1 Have requirements of Ethics Committee/Institutional Review Board approval been described?	☒	☐	☐	52
13.2 Has any outcome of an ethical review procedure been addressed?	☐	☐	☒	

Section 13: Ethical issues	**Yes**	**No**	**N/A**	**Page Number(s)**
13.3 Have data protection requirements been described?	☒	☐	☐	34, 35

Comments:

Section 14: Amendments and deviations	**Yes**	**No**	**N/A**	**Page Number(s)**
14.1 Does the protocol include a section to document future amendments and deviations?	☒	☐	☐	16

Comments:

Section 15: Plans for communication of study results	**Yes**	**No**	**N/A**	**Page Number(s)**
15.1 Are plans described for communicating study results (e.g. to regulatory authorities)?	☒	☐	☐	54
15.2 Are plans described for disseminating study results externally, including publication?	☒	☐	☐	54

Comments:

Name of the main author of the protocol: _____

Date: / /

Signature: _____

Annex 3. Additional Information

Not applicable.