Due to the size, this document has been divided into 6 parts, all of which are available on the ENCePP EU PAS register.
Part I = this pdf; Parts II - VI = under "Other documents "

PASS Information

Title	Comparative Assessment of VTE and Other Risks among Patients with Rheumatoid Arthritis treated with Baricitinib versus Tumour Necrosis Factor Inhibitors: A Multi-database Observational Cohort Study (I4V-MC-B023)
Version identifier of the final study report	1.0
Date of last version of the final study report	Not applicable
EU PAS register number	EUPAS 32271
Active substance	L04AA37 Baricitinib
Medicinal product(s):	Baricitinib: Olumiant®
Product reference:	EU/1/16/1170
Procedure number:	EMEA/H/C/004085/II/0031
Marketing authorisation holder(s)	Eli Lilly and Company
Joint PASS	No
Research question and objectives	This study aimed to evaluate the safety of patients with rheumatoid arthritis (RA) treated with baricitinib compared to patients treated with tumour necrosis factor inhibitors (TNFi). This aim was achieved using post marketing data from multiple sources and through the following objectives, that were addressed by a meta-analysis of analytic results across individual data sources: **Primary Objective:** • To compare the risk of VTE among patients with RA treated with baricitinib with the risk among similar patients treated with TNFi. **Secondary Objectives:** • To compare the risk of MACE among patients with RA treated with baricitinib with the risk among similar patients treated with TNFi. • To compare the risk of incident serious infection among patients with RA treated with baricitinib with the risk among similar patients treated with TNFi. • To describe the risk of tuberculosis (TB) requiring hospitalization among patients with RA treated with baricitinib.
Country(-ies) of study	United States, France, Germany, Japan, Sweden, United Kingdom
Author	GPS Epidemiology Eli Lilly and Company Lilly Corporate Centre Indianapolis, IN 46285 United States
Signature of principal investigator	Signature on file/see approval date below

Approval Date: 30-Jun-2022 GMT

Marketing Authorisation Holder

Marketing authorisation holder (MAH)	Eli Lilly and Company
MAH contact person	GPS Epidemiology Eli Lilly and Company Lilly Corporate Canter Indianapolis, IN 46285 United States

Table of Contents

List of Tables

List of Figures

List of Tables from Individual Data Sources

Baseline Demographics VTE Cohorts, Matched	**US Data Sources** Table 2_Aetna Table 2_HIRD Table 2_Cor_US Table 2_HUM Table 2_MTSCN Table 2_MDR Table 2_Optum Table 2_PP Table 2_PS20
	Europe and Japan Data Sources Table 2_ARTIS Table 2_BKK Table 2_Cor_JP Table 2_JMDC Table 2_SNDS
Baseline Demographics MACE Cohorts, Matched	**US Data Sources** Table 3_Aetna Table 3_HIRD Table 3_Cor_US Table 3_HUM Table 3_MTSCN Table 3_MDR Table 3_Optum Table 3_PP Table 3_PS20
	Europe and Japan Data Sources Table 3_ARTIS Table 3_BKK Table 3_Cor_JP Table 3_JMDC Table 3_SNDS
Clinical Characteristics Primary VTE Cohorts, Matched	**US Data Sources** Table 7_Aetna Table 7_HIRD Table 7_Cor_US Table 7_HUM Table 7_MTSCN Table 7_MDR Table 7_Optum Table 7_PP Table 7_PS20

	Europe and Japan Data Sources Table 7_ARTIS Table 7_BKK Table 7_Cor_JP Table 7_JMDC Table 7_SNDS
Clinical Characteristics MACE Cohorts, Matched	**US Data Sources** Table 8_Aetna Table 8_HIRD Table 8_Cor_US Table 8_HUM Table 8_MTSCN Table 8_MDR Table 8_Optum Table 8_PP Table 8_PS20
	Europe and Japan Data Sources Table 8_ARTIS Table 8_BKK Table 8_Cor_JP Table 8_JMDC Table 8_SNDS
Baseline Healthcare Resource Utilisation Primary VTE Cohorts, Matched	**US Data Sources** Table 12B_Aetna Table 12_HIRD Table 12B_HUM Table 12B_MTSCN Table 12_MDR Table 12B_Optum Table 12_PP Table 12B_PS20
	Europe and Japan Data Sources Table 12_ARTIS Table 12_BKK Table 12B_JMDC Table 12_SNDS
Baseline Healthcare Resource Utilization MACE Cohorts, Matched	**US Data Sources** Table 13B_Aetna Table 13_HIRD Table 13B_HUM Table 13B_MTSCN Table 13_MDR Table 13B_Optum Table 13_PP Table 13B_PS20
	Europe and Japan Data Sources Table 13_ARTIS Table 13_BKK

	Table 13B_JMDC Table 13_SNDS

Clinical Characteristics of RA Patients with VTE, Primary Definition	**US data sources** Table 40_PS20
	Europe and Japan data sources Table 40_ARTIS Table 40_BKK Table 40_SNDS
Pattern of RA Medication Use in Patients with VTE, Primary Definition	**US data sources** Table 41_PS20
	Europe and Japan data sources Table 41_BKK Table 41_SNDS
Time to First VTE Event (Days), Primary Definition	**US data sources** Table 42_PS20
	Europe and Japan data sources Table 42_ARTIS Table 42_BKK Table 42_SNDS
Incidence rate of VTE, Primary Definition	**US data sources** Table 45_Aetna Table 45_HIRD Table 45_COR_US Table 45_HUM Table 45_MTSCN Table 45_MDR Table 45_Optum Table 45_PP Table 45_PS20
	Europe and Japan data sources Table 45_ARTIS Table 45_BKK Table 45_Cor_JP Table 45_JMDC Table 45_SNDS
Comparative Risk of Incident VTE, Primary Definition	**US data sources** Table 48_MTSCN Table 48_MDR Table 48_PS20
	Europe and Japan data sources Table 48_ARTIS Table 48_BKK Table 48_SNDS Table 48_PS20
Clinical Characteristics of RA Patients with MACE	**US data sources** Table 51_PS20
	Europe and Japan data sources Table 51_ARTIS Table 51_BKK Table 51_SNDS

Pattern of RA Medication Use in Patients with MACE	**US data sources** Table 52_PS20
	Europe and Japan data sources Table 52_BKK Table 52_SNDS
Time to First MACE (Days)	**US data sources** Table 53_PS20
	Europe and Japan data sources Table 53_ARTIS Table 53_BKK Table 53_SNDS
Incidence Rate of MACE	**US data sources** Table 54_Aetna Table 54_HIRD Table 54_COR_US Table 54_HUM Table 54_MTSCN Table 54_MDR Table 54_Optum Table 54_PP Table 54_PS20
	Europe and Japan data sources Table 54_ARTIS Table 54_BKK Table 54_Cor_JP Table 54_JMDC Table 54_SNDS
Comparative Risk of MACE	**US data sources** Table 55_COR_US Table 55_Optum Table 55_PS20
	Europe and Japan data sources Table 55_ARTIS Table 55_BKK Table 55_SNDS
Incidence Rate of First Serious Infection	**US data sources** Table 59_Aetna Table 59_HIRD Table 59_COR_US Table 59_HUM Table 59_MTSCN Table 59_MDR Table 59_PP Table 59_PS20

	Europe and Japan data sources Table 59_ARTIS Table 59_BKK Table 59_Cor_JP Table 59_JMDC Table 59_SNDS Table 59_PS20
Comparative Risk of First Serious Infection Event	**US data sources** Table 61_Aetna Table 61_HIRD Table 61_HUM Table 61_Optum Table 61_PP Table 61_PS20
	Europe and Japan data sources Table 61_ARTIS Table 61_SA ARTIS Table 61_BKK Table 61_Cor_JP Table 61_SNDS Table 61_PP

1. Abstract

Title: Comparative Assessment of VTE and Other Risks among Patients with Rheumatoid Arthritis treated with Baricitinib versus Tumour Necrosis Factor Inhibitors: A Multi-database Observational Cohort Study (I4V-MC-B023)

Keywords: baricitinib, safety, VTE, meta-analysis, observational study

Rationale and background: The safety profile of baricitinib, an oral reversible JAK 1/2 inhibitor, is based on integrated clinical data from over 14,000 PY of exposure. Despite the extensive exposure to baricitinib from the integrated clinical program, information is limited from placebo or active comparator periods, which is especially limiting when considering less common safety events. In the 24-week placebo-controlled period, there was a numerical imbalance of VTE between baricitinib- and placebo-treated patients. The available information was not sufficient to support a definitive assessment of the risk of VTE associated with baricitinib treatment. The limited amount of placebo-controlled data also affected the ability to evaluate the risk of other safety outcomes such as MACE and serious infections. Post-marketing safety studies conducted within real-world populations are thus needed to better characterize the safety profile of baricitinib.

This cohort study evaluated the safety of baricitinib relative to TNFi treatment in a large number of patients with RA across multiple real-world data sources from Europe, US, and Japan.

Research question and objectives: The main goal of this study was to evaluate the safety of patients with RA treated with baricitinib compared to patients treated with TNFi. This aim was achieved through the following specific objectives:

Primary objective: To compare the risk of VTE among patients with RA treated with baricitinib with the risk among similar patients treated with TNFi.

Secondary objectives:

- to compare the risk of MACE among patients with RA treated with baricitinib with the risk among similar patients treated with TNFi
- to compare the risk of incident serious infection among patients with RA treated with baricitinib with the risk among similar patients treated with TNFi,
- to describe the risk of TB requiring hospitalization among patients with RA treated with baricitinib.

Study design: This was a multi-database cohort study including patients with RA across multiple real-world data sources from US, Europe, and Japan. The main study results are summarized as an aggregate incidence rate ratio based on meta-analysis of results from individual data sources. Overall incidence rate differences are also estimated. In addition to the meta-analyses, results from individual data sources are presented.

Setting and data sources: This cohort study assessed data from 16 sources in 6 countries (France, Germany, Japan, Sweden, United Kingdom, and US). All data included in this study were based on information previously collected for other purposes. Data sources included:

- RA disease and bDMARD registries: ARTIS, CorEvitas Japan and US RA registries
- Insurance claims records from commercial and national health insurance systems

  Europe and Japan: Betriebskrankenkasse, Cegedim THIN France, Clinical Practice Research Database, JMDC, Système National des Données de Santé,

  US: Aetna, Humana, Anthem HIRD, MarketScan, Military Health System Data Repository, Optum's de-identified Clinformatics® Data Mart Database, HealthVerity Private Source 20, IQVIA PharMetrics® Plus.

Subjects and study size: The study population consisted of adult patients diagnosed with RA who were incident users of baricitinib or the specific TNFi that qualified entry into the cohort during the study period. No maximum was set for the number of patients included in the study and data sources were selected for inclusion based on expected sizes of the baricitinib cohort. The study aimed to identify 6000 PY of baricitinib exposure to allow detection of a treatment effect as small as 1.8-fold greater among baricitinib- compared to TNFi-treated patients. The total number of patients included in analyses varied by data source and was ultimately determined by the proportion retained after propensity score matching.

Variables: **Exposure** – All available patients meeting the eligibility criteria and with exposure to baricitinib during the study period were included in the baricitinib cohort. For the TNFi cohort, exposure status was classified based on use of any specific, eligible TNFi medication.

Outcomes – VTE is the primary outcome of interest and reflects the occurrence of either PE, DVT, or other venous thrombosis. Secondary outcomes were defined as follows: MACE was a composite outcome based on the occurrence of either MI or stroke; serious infections was a composite endpoint of any serious infection, including bacterial, opportunistic, or viral infection requiring hospitalization. TB was limited to hospitalized TB.

Potential Confounders – potentially confounding variables were balanced across comparison groups in analyses for respective outcomes using propensity score matching, based on available information. Quantitative bias analysis was conducted to assess the impact of specific potential confounders (smoking, obesity, and disease severity) on analysis for VTE and MACE.

Statistical Analysis: All eligible patients were propensity score matched to address the potential imbalance of risk factors across treatment cohorts. The main study result for each outcome was estimated from a meta-analysis of IRR calculated from modified Poisson regression in individual data sources. Survival analysis of time-to-event data was investigated using a Cox proportional hazards regression model to generate HR. Frequently, HR could not be estimated due to zero events in the reference TNFi group. Incidence rate differences (IRD) were estimated by subtracting the rate in the TNFi treatment cohort from that in the baricitinib cohort.

Results: A total of 14 data sources (11 insurance claims databases and 3 registries) contributed to the meta-analysis. There were 7606 new baricitinib users (5879.2 PY baricitinib) from a total

of 9013 eligible, who were propensity score-matched and compared to similar patients treated with TNFi. A total of 97 patients with VTE were identified across both treatment cohorts, of whom 56 were treated with baricitinib. For VTE, the aggregate IRR from meta-analysis ($IRR_{meta-analysis}$) was 1.51 (95% CI 1.10, 2.08). The IR was greater among patients treated with baricitinib than with TNFi, with a difference of 0.26 (95% -0.04, 0.57) per 100 PY. There were a total of 93 patients with MACE, 54 of whom were treated with baricitinib. For MACE, the $IRR_{meta-analysis}$ was 1.54 (95% CI 0.93, 2.54). The IR was greater among patients treated with baricitinib than with TNFi, with a difference of 0.22 (95% CI -0.07, 0.52) per 100 PY. There were also 318 patients with serious infections, with 175 among patients treated with baricitinib. For serious infection, the $IRR_{meta-analysis}$ was 1.36 (95% CI 0.86, 2.13). The IR was greater among patients treated with baricitinib than with TNFi, with a difference of 0.57 (95% CI -0.07, 1.21) per 100 PY. There were no hospitalized TB events among baricitinib-treated patients included in this study.

Discussion: This multi-database multi-country study is the largest study of VTE, MACE, and serious infection among patients treated with baricitinib in a real-world setting. This study included case validation of VTE identified in claims data sources and implemented several design and analysis strategies to control for potential confounding including the use of an active comparator, new user study design, implementation of inclusion criteria in US claims-based data sources to approximate the required indication that baricitinib be used after TNFi, propensity score matching, and further adjustment for variables that were imbalanced after matching. Although these strategies were used to address potential confounding, this study was not able to evaluate the existence of differences in the prevalence of BMI, smoking, or disease severity between treatment groups. Since these are important risk factors for the outcomes evaluated, we cannot exclude the possibility that results may be biased. However, based on quantitative evaluation of the magnitude of bias that could have occurred, it is unlikely that the study results were meaningfully impacted by the lack of information about these factors.

Despite the size and scope of this study, almost half of data sources reported zero VTE or MACE in either the baricitinib or TNFi cohorts. In addition, the 2 largest data sources, ARTIS and SNDS, were from Europe and together contributed 71% of the total person-time and 76% or 80% of all VTE or MACE evaluated. Combined, all 9 US data sources contributed 13.9% of baricitinib exposure, highlighting the limited information available from this region.

In summary, the findings from this large multi-country study provide additional information about the safety of baricitinib. There is strong evidence of an association between baricitinib and risk of VTE compared to TNFi based on the significant overall meta-analysis result. This agrees with other studies that have found an imbalance of VTE during placebo-controlled periods of randomized studies and the ORAL surveillance study results for tofacitinib, another JAKi. For MACE, the meta-analysis result suggests a modest increase in risk associated with baricitinib compared to TNFi treatment and for serious infection, the overall meta-analysis result suggests a small increase in risk.

Conclusion: Lilly concludes that these results do not alter the current benefit-risk assessment for baricitinib. Lilly strengthened the warnings and precautions in the company core data sheet based

on early analyses from Study B023 and concludes that no additional changes are warranted to the company core data sheet or the risk management plan. Findings from this study reported in comparison with TNFi should be considered in context with the study limitations and evidence from other studies evaluating the safety of baricitinib and other JAK inhibitors. Healthcare professionals should consider information from this and other studies in aggregate and take appropriate precautions in patients with cardiovascular risk factors, risk factors for DVT/PE including immobilization, and in those with risk factors that may increase risk of serious infection.

Marketing Authorisation Holder(s): Eli Lilly and Company

Names and affiliations of principal investigator:

Pharmacoepidemiologist
Global Patient Safety
Eli Lilly and Company
Indianapolis, IN USA

2. List of abbreviations

Term	Definition
AE	adverse event: Any untoward medical occurrence in a patient or clinical investigation subject administered a pharmaceutical product that does not necessarily have a causal relationship with this treatment. An adverse event can therefore be any unfavourable and unintended sign (including an abnormal laboratory finding), symptom, or disease temporally associated with the use of a medicinal (investigational) product, whether or not related to the medicinal (investigational) product.
AEP	Aetion Evidence Platform
ARR	apparent relative risk, refers to the IRR estimates calculated by the study
ARTIS	Anti-Rheumatic Therapy in Sweden
ATC	Anatomical Therapeutic Chemical classification
bDMARD	biologic disease-modifying anti-rheumatic drug
BKK	Betriebskrankenkasse
BMA	British Medical Association
BMI	body mass index
b/tsDMARD	biologic / targeted synthetic disease-modifying antirheumatic drug
cDMARD	conventional disease-modifying anti-rheumatic drug
CCAM	classification commune des actes médicaux
CDAI	Clinical Disease Activity Index
CIRAS	Claims-Based Index for RA Severity
CI	confidence interval
CM	Clinical Modifications
CIP	Code Identifiant de Présentation
CPT	Current Procedural Terminology
CORPUS	Cohorte d'Observation Rhumatologique des Pratiques et des USages
CPRD	Clinical Practice Research Datalink
CV	cardiovascular
DAS28	Disease Activity Score 28

DDD	defined daily dose
DoD	US Department of Defense
DMARD	disease-modifying anti-rheumatic drug
DPC	diagnosis procedure combination
DVT	deep vein thrombosis
ED	emergency department
EDI	Electronic data interchange
EDW	Aetna Enterprise Data Warehouse
EMR	electronic medical record
ESPOIR	Etude et Suivi des Polyarthrites Indifférenciées Récentes
EU PAS	European Union electronic Register of Post-Authorisation Studies
FDA	US Food and Drug Administration
GKV	Gesetzliche Krankenversicherung
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
ID	identifier
INSERM	Institut national de la santé et de la recherche médicale (*Translation: French National Institute of Health and Medical Research)*
IRD	incidence rate difference
IR	incidence rate
IRR	incidence rate ratio
ISPE	International Society for Pharmacoepidemiology
ISPOR	International Society for Pharmacoeconomics and Outcomes Research
IV	intravenous
JAK	Janus kinase

JMDC	Japanese Medical Data Center
Lilly	Eli Lilly and Company
LTD	Long Term Disease
MACE	major adverse cardiovascular events
MDR	Military Health System Data Repository
MHS	Military Health System
MI	myocardial infarction
MTX	methotrexate
NDC	National Drug Code
NPI	National Provider Identifier
ORD	Optum Research Database
OUS	outside of the United States
PC0	prevalence of the potential confounding factor among the 'unexposed' (TNFi) cohort
PC1	prevalence of the potential confounding factor in the 'exposed' (baricitinib) cohort
PCS	Procedure Coding System
PE	pulmonary embolism
PHI	protected health information
PPV	positive predictive value
PS	propensity score
PY	person-years
QBA	quantitative bias analysis
QC	Quality Control
RA	rheumatoid arthritis
RD	rate difference
rEHR	reconstituted electronic health records
RR	rate ratio
RR_{CD}	Relative risk as measure of the association between the potential confounding factor and the outcome

SAP	statistical analysis plan
SHI	statutory health insurance
SNDS	Système National des Données de Santé
TB	tuberculosis
THIN	The Health Improvement Network
TNF	tumour necrosis factor
TNFi	tumour necrosis factor inhibitor
tsDMARD	targeted synthetic disease-modifying anti-rheumatic drug
UCD	Unité commune de dispensation
USPI	United States Package Insert
VTE	venous thromboembolism
WHO	World Health Organisation

3. Investigators

Principal investigator:

Pharmacoepidemiologist
Global Patient Safety
Eli Lilly and Company
Indianapolis, IN USA

Coordinating investigators:

Contact information for investigators from each of the below analytic partners is available upon request.

Aetion, Inc
Real World Evidence Solutions
5 Penn Plaza, 7th Floor
New York, NY 10001

ARTIS
Karolinska Institutet, Stockholm, Sweden
Clinical Epidemiology Division
Department of Medicine Solna
Karolinska University Hospital, T2:01
SE-171 76 Stockholm
Sweden

Bordeaux PharmacoEpi (BPE)
Université Bordeaux – CHU Bordeaux – Adera
Bâtiment Le Tondu – Case 41
146, rue Léo Saignat
33076 Bordeaux cedex, France

CorEvitas, LLC
Drug Safety and Pharmacovigilance
1440 Main Street, Suite 310
Waltham, MA 02451

HealthCore
Safety & Epidemiology
HealthCore, Inc
123 Justison Street, Suite 200
Wilmington, DE USA

IQVIA
Real World Evidence Solutions
Epidemiology and Drug Safety
New York, New York

4. Other responsible parties

See list in Section 3.

5. Milestones

Milestone	Planned Date	Actual Date	Comments
Registration in the EU PAS register	--	17 December 2020	
Final report of study results	30 June 2022	See Page 1	

6. Rationale and background

RA is a common systemic autoimmune inflammatory disease characterized by synovial inflammation that causes pain, swelling, stiffness, and leads to progressive destruction and deformity of small and large joints. In addition to the hallmark musculoskeletal features of the disease, patients with RA experience an increased risk of clinically important non-musculoskeletal comorbidities including:

- malignancy (Simon et al. 2015)
- infection (Doran et al. 2002)
- VTE (Kim et al. 2013; Lee and Pope 2014; Ogdie et al. 2018), as well as the individual events of DVT and PE (Choi et al. 2013)
- CV disease (Picerno et al. 2015), which includes MACE of myocardial infarction, stroke, and CV death (Aviña Zubieta et al. 2008; Aviña Zubieta et al. 2012), and
- overall early mortality (Mutru et al. 1985; Sihvonen et al., 2004).

Furthermore, patients with RA also experience a direct detrimental effect on everyday activities that results in impaired physical function, social participation, and health-related quality of life (Smolen et al. 2016).

Current treatment of RA prioritizes timely initiation and modification of DMARD therapy to bring patients to a target of sustained low disease activity or remission (Singh et al. 2016; Smolen et al. 2017). Achievement of these targets improves short- and long-term patient health outcomes, including prevention of progressive irreversible structural joint damage (Maini et al. 2004; Smolen et al. 2017). Patients typically begin treatment with oral cDMARDs such as methotrexate and modify treatment, as needed and tolerated, to achieve these targets.

Treatment modification often involves targeted DMARD therapy, which includes injectable bDMARDs such as TNFi and oral tsDMARDs, such as JAK inhibitors. With enhanced focus on tight disease control and increased availability of novel targeted therapies, the prognosis of patients with RA has greatly improved in recent years.

Baricitinib and the need for post marketing safety studies

The safety profile of baricitinib, a JAK inhibitor, is based on integrated clinical data from over 14,000 PY of exposure but includes limited information in relation to placebo or active comparator especially when considering less common events (Taylor et al. 2022). The 52-week duration of the controlled periods of the baricitinib clinical program also limits the ability to compare long-term risk across treatment groups. Further, while safety information about medications is collected during clinical development and pivotal trials, the studies are designed to evaluate overall safety and more common safety events, rather than rare outcomes such as VTE. In the 24-week placebo-controlled period, there was a numerical imbalance of VTE reported between baricitinib- and placebo-treated patients. The available information was not sufficient to support a definitive assessment of the risk of VTE associated with baricitinib treatment. The limited amount of placebo-controlled data also affected the ability to evaluate the risk of other safety outcomes such as MACE and TB. Each of these outcomes was observed

among baricitinib-treated patients during placebo- or active comparator-controlled periods, but these events are also observed more frequently in the general RA population relative to age-matched general population controls (Gabriel 2008; van den Hoek et al. 2017). In this scenario, low numbers can complicate the assessment of any potential causal relationship. Patients in clinical trials may also not be representative of the overall population who may be treated with a medication in the real world. Therefore, postmarketing safety studies conducted within real-world populations are needed to better characterize and establish the safety profile of baricitinib.

This cohort study evaluated the safety of baricitinib relative to the standard of care therapy (ie, treatment with TNFi) in a large number of patients with RA, across multiple real-world data sources from US, Europe, and Japan. Because outcomes such as VTE and MACE are not common, exposed patients from multiple sources were included to investigate potential associations with baricitinib treatment. Additionally, this approach allowed for potential replication of results across different populations and healthcare systems, which provided additional characterization about the safety of baricitinib. Results from individual data sources are presented as well as the main study results, which combined results from all sources using meta-analysis.

7. Research question and objectives

This study aimed to evaluate the safety of patients with RA treated with baricitinib in comparison to those treated with TNFi. This aim was achieved using post marketing data from multiple US and OUS sources and through the following objectives, addressed by a meta-analysis of analytic results across individual data sources.

Primary objective:

This study compared the risk of VTE among patients with RA treated with baricitinib to the risk among similar patients treated with TNFi.

Secondary objectives:

- To compare the risk of MACE among patients with RA treated with baricitinib with the risk among similar patients treated with TNFi.
- To compare the risk of incident serious infection among patients with RA treated with baricitinib with the risk among similar patients treated with TNFi.
- To describe the risk of TB requiring hospitalization among patients with RA treated with baricitinib. Due to the small number of events expected overall, incidence rates of hospitalized TB will be estimated, and no comparison will be made between treatment groups. A comparative analysis will be done if a sufficient number of TB events accrue to support at least 80% statistical power to detect a relative difference as small as 3.0 between treatment cohorts, if such a difference truly exists.

Due to the small number of expected events, and consequently limited statistical power in individual data sources, investigating the potential association of each outcome with baricitinib in individual data sources should be considered an exploratory objective.

8. Amendments and updates

There were no substantial amendments made to protocol after start of data collection. Full description of non-substantial amendments made to the protocol following initial protocol approval are found in Section 5 of protocol (see Annex 19).

9. Research methods

9.1. Study design

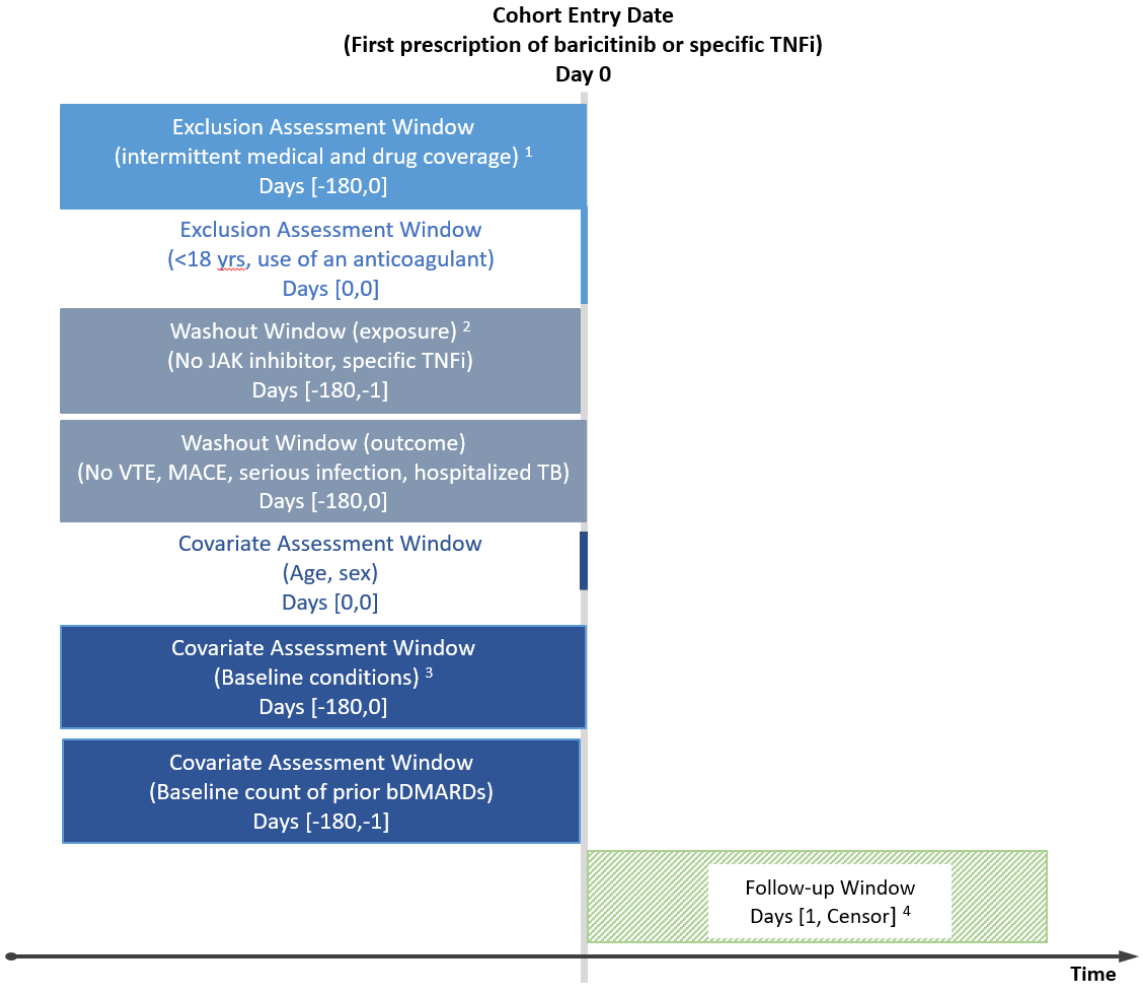
This multi-database cohort study relied on data from prospective cohorts in independent bDMARD and disease registries and retrospective cohorts in administrative claims databases and national healthcare systems to evaluate the safety of baricitinib relative to the standard of care therapy (ie, TNFi). Figure 9.1 provides a graphical depiction of the study design, detailing assessment of eligibility, covariates at baseline, index date, and follow-up. This design applied individually to data from each source.

Details of the contributing data sources are provided in Section 9.5. These data sources included information on patient demographics, RA diagnosis, records of filled prescriptions or administration of RA treatment, comorbidities, hospitalizations, and use of medications other than those for treatment of RA. All data sources were pre-existing independent collections of information and were not established or modified for the purposes of this study. A single protocol and overarching SAP was developed and shared with each data partner.

The primary outcome was VTE, and secondary outcomes included MACE (composite of MI and stroke), serious infection, and TB (Section 9.4.3). Incidence rates were estimated for each outcome in each data source (Section 9.9), with presentation of IRD, and IRR for comparative analysis. Survival analysis was also conducted and HR estimated, as possible. Although analyses from individual data sources had limited statistical power to estimate the association between baricitinib and the study outcomes, point estimates and confidence intervals were available for comparison regarding trend and direction. Meta-analysis was conducted for the main study result to obtain a single, combined estimate of the association between treatment and each primary or secondary study outcome.

There were several elements of the multi-database cohort study design and analyses that contributed to the scientific robustness of the study including the

- diversity of the data sources and populations represented provided a good foundation for ensuring the generalizability of any findings
- ability to consider replication across different populations, healthcare systems, and data sources. While infrequently applied, such replication to assess consistency of findings represents standard practice within epidemiological research when results may inform substantial decisions (Peng et al. 2006), and
- validation of the case definition for the primary VTE outcome in 2 settings.



Abbreviations: bDMARDs = biologic disease-modifying anti-rheumatic drugs; JAK = Janus kinase; MACE = major adverse cardiovascular events; TB = tuberculosis; TNFi = tumour necrosis factor inhibitor; tsDMARD = targeted synthetic disease modifying anti-rheumatic drug; VTE = venous thromboembolism

1 Up to 45-day gaps in medical or pharmacy enrolment will be allowed.

2 After cohort entry, patients in the TNFi treatment group may not be exposed to any specific TNFi previously used during the washout window.

3 Baseline conditions evaluated vary by outcome and are described in Table 4 of the protocol.

4 Earliest occurrence of the outcome of interest (ie, incident VTE, MACE, serious infection, or hospitalized TB, depending on the analysis), discontinuation of study medication +30 days or switch to a medication in another exposure cohort, initiation of a concomitant bDMARD or tsDMARD, disenrollment from the database or registry, or, where available, death, or the end of the study period.

Figure 9.1. Study design schematic for an administrative claims data source.

9.2. Setting

9.2.1. Overview of data sources

As this is a multi-database cohort study utilizing pre-existing real-world data collected independently of the study, the setting is best considered in the context of the data sources (summarized in Table 9.1). Defining characteristics for each data source include

- **Geography**: US or OUS, details which influence health care practices and date for study period given timing of market availability of baricitinib
- **Type of data:**
  - Prospective registry with primary data collection (eg, CorEvitas US and Japan RA registries, ARTIS)
  - Administrative claims data, from a commercial (eg, Anthem HealthCore Integrated Research Database [HIRD]), or national insurer (eg, SNDS, the French national healthcare database), or from an aggregated source with information from multiple insurers (eg, Marketscan)

Table 9.1. Overview of Study Setting, by Data Source

Data Sources	Country	Type of Data	Index Period (ie, Patient Eligibility Period)[a]
US Data			
Aetna (Healthagen)	United States	Claims, single insurer	Jun 2018 – Apr 2020
Anthem (HIRD)		Claims, single insurer,	Jun 2018 – Feb 2021
CorEvitas[b] US RA Registry		Disease registry	Jun 2018 – Dec 2018
Humana		Claims, single insurer	Jun 2018 – Jun 2020
Marketscan		Claims, multiple insurers	Jun 2018 – Jun 2021
Military Health System Data Repository (MDR)		Claims, single insurer	Jun 2018 – Aug 2020
Optum's de-identified Clinformatics® Data Mart Database		Claims, single insurer	Jun 2018 – Dec 2020
Pharmetrics® Plus		Claims, multiple insurers	Jun 2018 – Dec 2020
Private Source 20 (PS20) (HealthVerity)		Claims, multiple insurers	Jun 2018 – Sept 2020
Europe and Japan Data			
ARTIS	Sweden	Disease registry linked to national data	Feb 2017 – Dec 2020[d]
BKK	Germany	Claims, regional insurer	Feb 2017 – Dec 2018
Cegedim THIN France[c]	France	Ambulatory EMR	Sept 2017 – Dec 2020
CorEvitas[b] Japan RA Registry	Japan	Disease registry	Jul 2018 – Dec 2020
CPRD (Gold and Aurum)	United Kingdom	Primary care based electronic medical record	Feb 2017 – Jul 2021
JMDC	Japan	Claims, single insurer	Jul 2017 – Aug 2020
SNDS	France	Claims, universal national insurer	Sept 2017 – Dec 2019

Abbreviations: ARTIS = Anti-rheumatic therapy in Sweden; BKK = Betriebskrankenkasse; CPRD = Clinical Practice Research Database; EMR = electronic medical record; HIRD = HealthCore Integrated Research Database; OUS= outside of the United States; SNDS = Système National des Données de Santé; THIN = The Health Improvement Network; US= United States.

[a] The period for identifying eligible patients (ie, index period) commenced at the time of market launch for baricitinib, which differed by region, and continued until the most recent data available at the time of extraction. Data prior to index informed eligibility and baseline information for analyses. In ARTIS, eligible patients were identified between Feb 2017 and Dec 2019 with follow-up continuing through Dec 2020. CorEvitas Japan RA registry did not have first patient enrollment until 1 year after market launch.

[b] CorEvitas was previously known as Corrona RA Registry

[c] Patients in the French Cegedim THIN data are also in the SNDS population, which includes all French residents.

9.2.2. Study period and follow-up

The period for identifying eligible patients (ie, index period) commenced at the time of market launch for baricitinib, which differed by region, and continued until the most recent data availability at the time of data extraction. The specific dates for the index period vary by geography and individual data source (Table 9.1). Details of each contributing data source are in Section 9.5.

The study used a new user design to identify patients who newly initiated baricitinib or a specific TNFi (Lund et al. 2015). The study design is depicted in Figure 9.1, which also shows the periods for assessment of study eligibility, cohort entry, covariate assessment, and follow-up in administrative claims data sources. These periods within the study design are defined for each data source. In summary,

- **Cohort entry date**, or index date, was defined as the date of the first dispensing or prescription fill (claims data) or as the first recorded prescription of either baricitinib or the specific TNFi included in the comparator group (registry data).
- Regardless of treatment group, the earliest cohort entry date possible was the date of market availability of baricitinib for that data source.
- **Follow-up** started on the day after cohort entry. In claims data, all patients were required to maintain continuous enrolment, defined as continuous medical and prescription drug coverage with a gap not longer than 45 days.
- The **covariate assessment window** (“baseline”) for claims data relied on the 6-month enrolment period prior to the index dispensing, including the day of dispensing except when assessing numbers of prior DMARDs. For data sources with primary data collection (registries), medical history and physician diagnosis of RA on the day of prescribing was used. For example, with CorEvitas, medical history information available on cohort entry reflected medical history “at any time in the past”, as this reflects how data is collected upon enrolment into the registry.

- **Censoring** (end of follow-up) criteria were the same for all analyses, except for the first criterion where the specific incident event differs, as each outcome was analysed separately. Censoring occurred at the earliest of any of the following:
    - occurrence of an incident event
    - discontinuation of study medication plus 30 days, including due to a switch to a different TNFi for patients in the TNFi cohort
    - switch to a medication in the other exposure cohort
    - initiation of a concomitant bDMARD, including a TNFi, or tsDMARD
    - disenrollment from the database or registry
    - death (where available), or
    - end of study period.

9.3. Subjects

The study population consisted of adult patients diagnosed with RA who, during the study period, were incident users of baricitinib or the specific TNFi that qualified a user to enter the cohort. Only new users of baricitinib or a specific TNFi were included in the study. New users were defined as patients without prior use of baricitinib or the specific TNFi during the 6-month period immediately prior to the index date or the equivalent, based on medical records or data collection forms. All patients in the main analyses were also new users of JAK inhibitors. Additional detail in Section 9.4.1 describes how exposure to the index medications was specified.

All patients meeting the eligibility criteria below were included in the main analyses and an attrition table was generated. Because the methods used to identify RA diagnoses and medications to treat RA differ depending on the type of data, the inclusion criteria are specified for each type of data source.

Inclusion Criteria: Registry or other source relying on primary data collection

1. Adult patients diagnosed with RA who have newly initiated treatment with baricitinib or a TNFi; the date of treatment initiation is the cohort entry date (also known as the index date). New users of TNFi should not have previous exposure to baricitinib.

Inclusion Criteria: Administrative claims data

1. Patients have an RA diagnosis code from a physician encounter and initiated baricitinib or a TNFi*; the date of treatment initiation is the cohort entry date (also known as the index date)
2. Patients are aged ≥18 years on the cohort entry date,
3. Patients have continuous medical and prescription drug coverage for at least 6 months before cohort entry.

* Initiation of a TNFi is defined as dispensing of 1 of the following TNFi without prior dispensing of that same TNFi during the baseline period: adalimumab, certolizumab pegol, etanercept, golimumab, or infliximab.

Inclusion Criteria: US administrative claims data

In addition to criteria 1-3 for administrative claims data, in US claims, patients with RA who newly initiate treatment with a specific TNFi required prior treatment with at least 1 TNFi. This aligns the comparison groups in US data with the indicated population for baricitinib, which is patients with moderate to severe RA who have had an inadequate response to 1 or more TNF antagonist therapies (Olumiant package insert, 2020).

Exclusion criteria are specific to analyses and are included in Section 9.9.2.2.

9.4. Variables

9.4.1. Exposure

In claims data sources, exposure to baricitinib or other medications indicated for the treatment of RA were ascertained based on a recognized classification scheme such as the ATC/DDD system (WHO 2018), or the National Drug Code or Generic Product Identifier for outpatient pharmacy dispensing and based on Health Care Common Procedure Coding Systems for injections or infusions that occurred in a healthcare setting.

All available patients meeting the eligibility criteria and with exposure to baricitinib during the period of the study were included in the baricitinib cohort. For the TNFi cohort, exposure status was classified based on use of any specific, eligible TNFi medication (Table 9.2). Two treatment groups were created for these analyses:

- **Baricitinib cohort:** Patients newly initiating baricitinib,
- **TNFi cohort (referent group):** Patients newly initiating a specific TNFi.

Table 9.2. Eligible Treatments for RA Patients

Conventional DMARD (as concomitant medication)		TNFi Cohort		Baricitinib Cohort
Cyclosporine	Methotrexate	Adalimumab	Infliximab	Baricitinib
Gold sodium thiomalate	Mycophenolate mofetil	Certolizumab pegol		
Hydroxychloroquine	Penicillamine	Etanercept		
Leflunomide	Sulfasalazine	Golimumab		

Patients were assigned to treatment groups based on several considerations:

- Because patients treated with baricitinib represent a smaller proportion of patients with RA than do patients treated with TNFi, inclusion in the baricitinib cohort was prioritized when possible, to maximize the group size. Given the relatively larger size of the TNFi population, this did not impact the availability of patients for the comparison group. This was required to maximize the number of patients in the baricitinib group and increased

the probability of having sufficient events in each groups to allow comparative analyses. However, this approach may select baricitinib patients who initiate TNFi followed by initiation of baricitinib during the study period which tend to increase differences between groups with respect to disease progression and severity. To the extent that these differences are risk factors for the outcomes under analysis and patients treated with baricitinib have more severe disease, this prioritization may tend to bias results away from the null towards increased relative risk (see Section 9.6 for detail).

- Patients were permitted to contribute person-time and events to only a single treatment group in an analysis, *either* the TNFi cohort or the baricitinib cohort. Study results therefore apply to 'persons' rather than to 'treatment episodes'.
- As described in Section 9.2, in US administrative claims data only, assignment to the TNFi cohort required patients to have treatment with at least one other TNFi prior to the index TNFi. Since US patients treated with baricitinib must have received prior treatment with a TNFi, this permits the baricitinib and TNFi treatment groups to be more similar to each other, including with respect to baseline disease severity and activity.
- A patient treated with a biosimilar was defined as initiating a TNFi only if the patient did not have prior exposure to the originator drug or another biosimilar to the originator.
- Discontinuation or bridging of a gap in exposure was based on the time elapsed since the last exposure. Discontinuation (one of the censoring criteria outlined in Section 9.2.2) was defined as a gap in exposure exceeding the recommended dosing interval for the specific medication + 30 days. In contrast, if exposure was interrupted by a gap less than this interval + 30 days, the exposure was bridged, and exposure was defined as continuous. Protocol Table 2 and Figure 2 (Annex 19) provide detailed information on the specific gap thresholds used for each medication.
- Concomitant treatment with cDMARDs was not restricted (Table 9.2). Concomitant use was defined based on dispensing of a cDMARD between consecutive dispensings of the same b/tsDMARD and was described by treatment group. In some claims data (ie, Aetna, Cegedim THIN France, Humana, JMDC, MarketScan, Optum, and PS20), concomitant use was defined based on having at least two dispensings of the cDMARD during treatment with baricitinib or the index TNFi. Concomitant use of bDMARDs, including TNFi, or tsDMARDS was not permitted and initiation of such concomitant use was a criterion for censoring.

9.4.2. Outcomes

9.4.2.1. Venous thromboembolic events

Venous thromboembolism is the primary outcome of interest. It is an outcome that reflects the occurrence of either PE, DVT, or other venous thrombosis. In claims data, the case definition was based on ICD-10 discharge or diagnosis codes (Annex 19) occurring in ED, hospital, or outpatient settings, with an additional requirement for low molecular-weight heparin or oral anticoagulant dispensing within 31 days after diagnosis. The requirement for evidence of dispensing of an anticoagulant was based on the diagnostic setting and, for inpatient diagnoses, the position of the code (primary/secondary). This definition is summarized below and is

consistent with work conducted in a large validation study of VTE in 4 US integrated health care delivery systems (Fang et al. 2017). The event date was the date of the first recorded qualifying diagnosis of VTE during follow-up.

VTE case definition in claims data

VTE = occurrence of PE or DVT or other venous thrombosis. Each component is defined by:

- **PE**: ICD-10 diagnosis codes in the primary position for inpatient or ED setting. Inpatient PE diagnoses in the secondary position require anticoagulation within 31 days of the event, (ie, from the date of discharge). No outpatient codes for PE will be included.
- **DVT**
    - **Of the lower extremity:** ICD-10 diagnosis codes in the primary position for inpatient or ED setting. Inpatient diagnoses in the secondary position and outpatient diagnoses require anticoagulation within 31 days of the event (ie, from the date of discharge).
    - **Of the upper extremity:** ICD-10 diagnosis codes in the primary position for inpatient or ED settings and outpatient diagnoses require anticoagulation within 31 days of the event (ie, from the date of discharge). No inpatient diagnoses in the secondary position will be included.
        - **Other venous thrombosis:** ICD-10 diagnosis codes in the primary or secondary position for inpatient or ED settings and outpatient diagnoses require anticoagulation within 31 days of the event.

These case definitions were based on information from US healthcare settings and were modified to account for regional differences in healthcare or in use of diagnosis codes among geographies. Modifications were noted in SAP addenda for the individual data source.

To ensure that the VTE events identified in claims data accurately reflect true events, the case definition described above was validated using clinical information from medical chart review in the US and clinical review of patient claims histories in French patients (see summary in Section 9.4.2.1.1).

In CorEvitas data, VTE was identified based on physician diagnosis and adjudication of endpoints within the registry procedures. In ARTIS, VTE was defined based on a validation study of a case definition for incident VTE in Swedish patients with RA in the Swedish National Patient Register (Molander et al. 2022).

9.4.2.1.1. Validation of VTE case definition

Validation of the main VTE definition for claims-based data sources was based on two strategies, one using a US data source and one using a data source with information on French patients.

9.4.2.1.1.1. Validation in US claims data

In US data, a separate medical chart abstraction study was conducted to evaluate the use of administrative claims data to identify VTE among patients with RA (Study I4V-MC-B029; Annex 17). This validation study utilized health care claims data (medical and pharmacy) from the ORD, a closed claims data source, to identify patients with a presumptive VTE based on the case definition described in Section 9.4.2.1. Patient clinical data was abstracted from targeted medical charts to confirm the VTE diagnosis using pre-specified case confirmation rules based on clinical guidance. This study was reviewed and approved by an institutional review board and privacy board (WCG IRB #1284955). Commercially insured patients and those enrolled in a Medicare insurance plan with medical and Part D coverage with a claim for a RA therapy of interest between 01 May 2016 through 30 November 2020 were identified from ORD.

In this VTE validation study (Study B029), a total of 155 RA cases from the ORD database met the study inclusion criteria based on meeting at least one category of the claims-based primary VTE algorithm and procurement and abstraction of at least one medical chart. The VTE identified based on the primary case definition described above was validated in medical charts for 117 of the 155 (PPV = 75.5%; 95% CI 68.7%, 82.3%) identified cases.

The protocol described 2 alternate case definitions for VTE designed to bracket the main definition in terms of sensitivity and specificity (Study B023 protocol Section 8.3.2.1, Annex 19). Validation was planned only for the primary case definition. Due to the availability of additional resources, a modest number of cases meeting the less stringent case definition 2 were also evaluated. Of 76 patients with presumptive VTE identified based on case definition 2, 40 cases were confirmed after clinical review of medical charts, resulting in a lower PPV of 52.6% (95% CI 41.4%, 63.9%). As a result of this assessment, use of the less stringent case definition 2 for sensitivity analyses was not pursued further in the main study.

9.4.2.1.1.2. Validation in French data

In French patients, an alternative approach was taken to validate the VTE case definition (Annex 18). General Data Protection Regulation does not permit re-identification of subjects in the SNDS and thus prevents linkage of individual claim records with patient medical records without consent from the patient. Instead, adjudication of reconstituted EHR (rEHR) was used to validate the main VTE case definition (Thurin et al. 2021). Briefly, medical history available in the SNDS (eg, drug dispensing, procedure codes including surgery and imaging, hospital discharge diagnoses, laboratory tests) was used to generate anonymized rEHRs. Medical experts then blindly adjudicated the rEHR of the selected patients and expert conclusions were used to estimate the PPV.

The initially planned VTE case definition for French claims data had a low PPV (61%). The case definition used for VTE in French claims was refined based on expert clinical consultation. After evaluation of an initially planned algorithm revealed an unsuitable PPV, and adjustment of the algorithm based on expert clinical feedback, the new version of the algorithm had a PPV of 92% (95% CI 86%, 98%) for VTE identification in the SNDS. This updated version of the VTE case definition was used in the final analysis.

9.4.2.2. Major adverse cardiovascular events

MACE, a secondary outcome, was a composite outcome based on the occurrence of either MI or stroke. In claims data, the occurrence of MACE was captured based on ICD-10 diagnosis codes. As described for VTE, the case definition for MACE was based on information from US healthcare setting and may have been modified within a data source, as appropriate for the geography. In registry data, or other sources where outcomes have been recorded based on clinical evaluation, MACE was defined based on clinical information such as physician diagnosis and adjudication of endpoints within the registry procedures.

MACE case definition in claims

MACE = occurrence of either MI or stroke

- **MI**: a primary discharge code of acute MI (any I21 ICD-10 code) for an inpatient visit (PPV ≥93%; Fralick et al. 2018). Fatal MI was also included for analyses in the CorEvitas (US and Japan) registry data. No other data sources included MI due to death.
- **Stroke**: a primary discharge code of acute ischemic or haemorrhagic stroke (any I60, I61, or I63 ICD-10 code) for an inpatient visit (PPV ≥82% for ischemic and ≥87% for haemorrhagic stroke; McCormick et al. 2015).

9.4.2.3. Serious infections

Serious infections, another secondary outcome, was a composite endpoint of any serious infection, including bacterial, opportunistic, or viral infection requiring hospitalization. As described in Section 9.9.2.2, patients with evidence of serious infection in the 6 months prior to cohort entry were excluded from analysis. Therefore, this analysis is specific to incident serious infection. In claims data, this outcome was defined based on ICD-10 diagnosis codes. Infections were identified based on inpatient codes in the primary position. In the HIRD data, inpatients stays were required to be at least 3 days. The PPV was 90.2% when expert review of medical records was used as the gold standard (Schneeweiss et al. 2007).

In primary data sources, serious infection was based on clinical judgment and when available (CorEvitas), on adjudicated events. In the CorEvitas registries, serious infection isdefined as an infection requiring hospitalization and/or treatment with IV antibiotics. .

9.4.2.4. Tuberculosis requiring hospitalization

Hospitalized TB, another secondary outcome, was defined based a primary hospitalization discharge diagnosis code (any ICD-10 code A15-A19) in claims data and based on clinical data from physicians in registries. In claims, the PPV for active TB was between 9% and 54% when expert review of medical records was used as the gold standard (Winthrop et al. 2011). Due to the limited number of events and the descriptive nature of this secondary outcome, all further information about hospitalized TB is in Section 10.1.6.

9.4.3. Potential confounding factors

Patient baseline characteristics were described from information available prior to cohort entry. This corresponded to information from the 6 months prior to initiation of index medication

(claims data, EMR, patient registries) and/or information collected at enrolment in a registry (patient registries). Characteristics included demographics, history of RA treatment, medical history, and healthcare resource utilisation for patients in each individual data source (unmatched but eligible and matched cohorts).

Imbalances of patient characteristics across comparison groups can lead to biased results if those characteristics are also risk factors for the outcome under analysis. A comprehensive list of covariates considered in the propensity score for each outcome is included in the protocol (Table 4 in Section 8.3.3, Annex 19).

Some information is typically not available in claims data, such as race, education, BMI, alcohol and tobacco use, measures of disease activity, or disease duration. Proxies for some of these covariates, such as the CIRAS index (Ting et al. 2008) may sometimes be used to adjust statistical models to account for potential confounding factors that would not otherwise be available. Three important potential confounders that are not well-measured in claims data, and the methods employed to attempt to adjust for them, are described further.

BMI

BMI is an important potential confounder. In data sources where BMI is not directly available (claims data), ICD-10 diagnosis codes corresponding with BMI categories were used (Z68 codes for categories of BMI <20, ≥20-29, ≥30-39, ≥40; and E66 for 'obesity'). The accuracy of these Z68 codes increases with increasing category of BMI (Lau et al. 2015).

The information that is available was included in the propensity score model, but incomplete coverage of smoking and BMI (eg, missing values, information not collected, or imprecise or incomplete measures of the variable) would leave the potential for residual confounding. For this reason, information about the distribution of these variables was abstracted from medical charts for a sample group of patients included in US claims data and from a large cohort of French patients with RA. Additional information about how this information was collected is available in 2 separate protocols for US (Optum) and OUS (French) data (Annex 17and Annex 18, respectively). The abstracted information was used for quantitative bias analysis, to provide context for understanding the possible impact to the study results of the observed smoking and BMI distributions in the baricitinib and TNFi-treated patients (see Section 9.9.4.2). Patients were also compared with respect to the duration of exposure, which defined follow-up.

Smoking

Smoking is an important potential confounder. Smoking status was not directly available in claims data sources. However, diagnosis codes for indirect indicators of smoking were queried (ICD-10: F17, O99.33, T65.21, Z87.891; CPT: 99406, 99407, G0436, G0437, G9016, S9453, S4995, G9276, G9458, 1034F, 4004F, 4001F). This includes claims for eg, tobacco use disorder, records of counselling visits for smoking, and anti-smoking prescription medication use. One study reported that the PPV of this smoking algorithm based on medical claims was 100% (95% CI 79.4%, 100%), but with sensitivity 26.2% (95% CI 15.2%, 37.3%) (Desai et al. 2016).

RA severity

Disease severity is an important potential confounder. Data on clinical status indicators among RA patients (eg, number of flares, physician global rating, functional and ambulatory status, presence of swollen joints, etc) are not collected in claims databases. However, claims data do contain information on factors that correlate with RA disease severity, such as physician visits and medication fills. Although the CIRAS index (Ting et al. 2008) was used to control for disease severity in statistical analyses, the index is not a strong proxy and may not be able to fully control for any existing imbalance in severity between comparison groups. RA disease severity was also addressed partially by adjusting for number of rheumatologist visits and number of unique DMARD medications ordered during the baseline period.

9.5. Data sources

This study was conducted using information from patients in multiple data sources across several geographic regions, including the US, EU, and Japan. Data used for this study was primarily from administrative claims databases, but also from electronic medical records and RA patient registries. The diversity of data sources and populations was designed to provide a good foundation for ensuring the generalizability of findings, since this study analyses data from a large proportion of the total available baricitinib exposure. All data used in this study should be considered secondary data as none was collected for the purpose of the B023 study.

Registries typically identify patients based on a specific medical condition such as a disease or an exposure such as a medication. Patient registries collect data in a uniform manner, often at regular intervals within predefined visit windows. This study includes the CorEvitas registry data (US and Japan) and the ARTIS data, which includes information from the Swedish SRQ registry.

Administrative claims databases collect information on healthcare services for billing purposes for health plan enrolees, whether those plans are commercial US plans or national plans, as in the French SNDS data. Claims records include patient-level information about dispensing of medications, diagnoses, performed procedures, and demographics. Administrative claims data sources are not generated with research intent, they are designed to maximize reimbursement and reflect coding practices.

An overview of the data sources is outlined in Table 9.1. This section provides detailed information regarding each data source.

9.5.1. US data sources

9.5.1.1. Aetna (Healthagen)

The Aetna EDW includes medical and pharmacy health plan membership, medical and pharmacy health plan eligibility, medical claims (ie, diagnosis and procedure claim codes), outpatient pharmacy claims, outpatient lab test results, and data derived from Aetna's care management processes for Aetna's non-administrative services Commercial members and Medicare Advantage, with service dates from 3 years plus the current year. The population of focus for this project will be Aetna fully insured and Medicare Advantage patients. These

segments of the Aetna population purchase insurance directly so longitudinal medical and prescription data are available for these patients.

Aetna serves people in all 50 states and multiple US territories.

9.5.1.2. Anthem (HealthCore Integrated Research Database [HIRD])

HealthCore, Inc. is a wholly owned subsidiary of Anthem, Inc., which is one of the largest health benefits companies in the US in terms of medical membership. Anthem is an independent licensee of the Blue Cross and Blue Shield Association. HealthCore is the health services research entity for Anthem that integrates the public health, pharmacoepidemiologic, health outcomes, and pharmacoeconomic concerns of these companies and their clients to conduct outcomes analyses.

The HIRD is a large health care database maintained by HealthCore for use in health outcomes and pharmacoepidemiologic research. The HIRD is a broad, clinically rich, and geographically diverse spectrum of longitudinal medical and pharmacy claims data with near real-time access from health plan members from across the US. Member enrolment, inpatient and outpatient medical care (professional and facility claims), outpatient prescription drug use, certain outpatient laboratory test results, and healthcare utilisation are tracked for health plan members in the database dating back to January 2006.

In the HIRD, diagnoses and procedures are identified by the following types of codes for both outpatient visits and inpatient stays: ICD-9-CM (Ninth Revision, Clinical Modification), ICD-10-CM (Tenth Revision, Clinical Modification, since October 2015), CPT, and HCPCS (Healthcare Common Procedure Coding System). Drug claims are recorded using National Drug Codes (NDCs), which can then be translated to broader categories of coding, such as Generic Product Identifier codes. Information on drug exposure is obtained from claims for pharmacy dispensations and includes data on prescribing physician specialty, drug dispensed, quantity and date dispensed, drug strength, and days' supply. Data on underlying medical indication are not available. 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, national vital statistics records, cancer and vaccine registries (state-by-state), disease and device registries, member and provider surveys, and point of care clinical data.

Claims data in the HIRD are updated monthly, with an approximate 3-month time lag for greater than 85% capture of paid medical claims. The lag time for pharmacy data is shorter, with approximately 98% paid within 30 days.

9.5.1.3. CorEvitas – United States

CorEvitas is a US prospective registry of patients with RA which collects data from patients and physicians at the time of routine clinic visit. Clinically relevant information is collected on disease activity, including joint counts, visual analogue scales for physicians and patients, laboratory values and diagnostic tests, and radiographic outcomes. Patient demographic and lifestyle habits, including smoking, alcohol consumption, BMI, employment, insurance type and status, are routinely recorded. Medication start and stop dates and reasons for the start or switch are recorded by the treating physician. Targeted follow-up forms are implemented for outcomes

of interest including VTE, MACE, and serious infection. As of March 2019, CorEvitas had enrolled over 51,600 patients with RA from 769 different rheumatologists in 42 US states. Data from physician-patient encounters collected via targeted event follow-up forms are confirmed by medical record review in order to confirm, validate, and adjudicate reports from physician-patient encounters (Kremer 2016).

9.5.1.4. Humana

The Humana Research Database (Louisville, KY) contains enrolment information linked to medical and pharmacy claims data for Medicare and commercially insured members across the US. Humana is the second largest private Medicare insurer in the US, with more than 4.5 million Medicare Advantage members enrolled as of 31 December 2020. The Database includes administrative claims data for approximately 30 million individuals enrolled from 2007 to the present.

Enrolment data includes coverage start and end dates, date of birth, sex, geographic region, race/ethnicity, and insurance line of business. The medical claims data includes information related to facility and provider claims such as service date, diagnosis code(s), procedure code(s), and place of treatment. Diagnoses and procedures in the Database are coded according to the ICD-9-CM schema until October 1, 2015 when transition to ICD-10-CM occurred. Pharmacy claims data includes data elements such as prescription fill date, NDC, quantity dispensed, and days' supply. A unique member identifier is assigned to each individual and remains constant regardless of any gap in plan enrolment or transition between lines of business (eg, transition from an employer-sponsored Humana commercial plan to a Humana Medicare Advantage plan).

9.5.1.5. Marketscan (IBM Watson)

The IBM Watson Marketscan database contains healthcare claims-based information for enrolees in large employer-sponsored health insurance plans across the US. Commercial claims databases contain information on inpatient claims, outpatient claims, outpatient pharmacy claims, and plan enrolment. Information available for enrolees includes sex, age, starting and ending dates of enrolment, and diagnoses and performed procedures. During the period of this study, diagnosis information was indicated using ICD-10-CM codes and, in the inpatient data, procedures performed were indicated using ICD-10-CM codes and the American Medical Association's CPT code sets.

9.5.1.6. Military Health System Data Repository (MDR)

The Military Health System is a comprehensive medical network with the DoD provides health care to all US military personnel, their dependents, and retirees. MHS operates the largest cradle-to-grave health care system in the US, with over 10 million patients actively receiving care on an annual basis. Patients enrolled in the MHS receive benefits through the TRICARE nationwide managed care program, which combines health care from DoD facilities with those from the private sector. Patients are not required to use military medical facilities, and many use their MHS coverage to obtain care in civilian facilities. Thirteen percent of the MHS population are active duty military, meaning that most enrolees are non-military. Males represent 51% of the population and the age distribution of patients in the MHS is similar to the US population.

Comprehensive medical care is recorded in the MDR, which is a continually updated longitudinal EMR database, capturing and integrating all health care events for the entire DoD network since 2000. The DoD's EMR system, the Armed Forces Health Longitudinal Technology Application, is linked to the Composite Health Care System, which allows access to laboratory, pathology, and radiology orders and results for a subset of patients receiving care in military treatment facilities (15-20%). Currently, the MDR contains data on more than 10 million active beneficiaries receiving care at more than 65 hospitals and over 500 military clinics, as well as at private hospitals and clinics throughout the country. Examples of data recorded in such encounters include demographic data (eg, age, sex, race), provider information (eg, provider ID, specialty, and facility), diagnostic codes (ICD-10-CM and ICD-9-CM codes for outpatient encounters; Diagnosis-Related Group classifications for inpatient encounters), CPT codes for procedures, and, among visits within military treatment facilities (15% to 20% of the encounters), additional vitals and lifestyle information (eg, BMI, smoking behaviours, alcohol use, and chemistry laboratory results). Mortality information is obtained via a master death file that compiles, processes, and validates all death records from inpatient hospitalization discharge dispositions, ambulatory and outpatient encounter records, a casualty death feed of combat-related deaths among active duty service members, survivor self-report, and an established, recurring Social Security Death Index feed from the Social Security Administration. Death data are updated monthly and directly linked to all eligible beneficiaries.

All prescribing and dispensing details, including mail order and inpatient medications, are electronically coded in the MDR. Details include the prescribed drug name and NDC, the treating facility, department(s) rendering care, the treating provider(s), start and end dates of the prescription order, amount dispensed, date dispensed, route of medication, units, number of refills, and remaining refills. The DoD manages a drug formulary that has been well established for decades. In general, DoD providers follow the recommendations of the DoD formulary. However, the approval process to prescribe a non-formulary drug is historically expedient, allowing providers to prescribe what they consider to be the optimal treatment regimen. In general, a DoD provider can prescribe most, if not all, available medications in a therapeutic class to a DoD beneficiary. In the civilian network, in which patients utilize TRICARE insurance in the pharmacy setting (ie, retail and mail order), there are even fewer restrictions on prescription drugs as prescribed by the provider. Notably, it is not uncommon for drugs recently approved by the FDA to be prescribed in the DoD pharmacy system much faster than in other integrated delivery networks and commercial insurance plans.

9.5.1.7. Optum Clinformatics Data Mart

This is an administrative health claims database with information from medical and pharmacy benefits coverage for members of a large national managed care company affiliated with Optum. The population is geographically diverse, spanning all 50 states and including approximately 12 to 13 million annual covered lives with commercial health plan data and Medicare Advantage members. For each participant, the data contain demographic information, health plan enrolment status, inpatient and outpatient medical encounters (including diagnosis and procedure codes from encounters), and drugs filled on an outpatient basis, including national drug code, quantity

dispensed, and days' supply. The underlying information from the study database is geographically diverse and reasonably representative of the US population.

9.5.1.8. IQVIA PharMetrics® Plus

The PharMetrics Plus database is the largest claims database of integrated medical claims in the US. As of February 2022, the aggregated PharMetrics Plus database is comprised of adjudicated claims for approximately 210 million unique enrolees across the US with over 54 million enrolees having at least 3 years continuous enrolment. Data are available from 2006 onwards. PharMetrics Plus has diverse representation of geography, employers, payers, providers and therapy areas, with the majority of 3-digit ZIP codes in the US covered and reported. Patients in the PharMetrics Plus database are similar to the national, commercially insured population in terms of age and sex for individuals age 65 and under. The data are also longitudinal, with more than 61 million patients with 3 or more years of continuous enrolment.

The PharMetrics Plus database contains information including (but not limited to): inpatient and outpatient diagnoses (ICD-10-CM and ICD-9-CM), inpatient and outpatient procedures (CPT, 4th edition), HCPCS, ICD-9-CM, and ICD-10-CM), retail and mail order prescription records (inpatient dispensings are not available), inpatient stays (eg, admission type and source, discharge status), provider details (eg, specialty), dates of service, demographic variables (eg, age, sex, and geographic region), and start and stop dates of health-plan enrolment. Mortality and cause of death information are not currently available in the database. Data contributions are subjected to a series of quality checks to ensure a standardised format and to minimize error rates.

9.5.1.9. HealthVerity Private Source 20 (PS20)

Source 20 represents the largest source of closed payer medical claims available from health insurance companies. Covering more than 150 unique payers and more than 120 million total patients, Source 20 medical claims offer a comprehensive view of a patient's entire medical journey including both inpatient and outpatient diagnoses, procedures and physician-administered medications over the enrolment/eligibility period. Unlike other closed claims datasets, Source 20 captures enrolment/eligibility information as well as treating physician identity (NPI), patient/physician geography, and detailed physician visit information including the cost of a paid claim. Because closed payer data captures all insurance-related patient transactions, the availability of both common and rare diseases, co-morbidities, outcome tracking, and in-office medication administration is excellent. Furthermore, coverage of Commercial, Medicare and Medicaid is available at levels far exceeding competitive data offerings.

Source 20 is an ideal resource when comprehensive treatment information is required over defined patient treatment durations. The data has been used successfully in pharmaceutical research supporting rare disease, surgical outcomes, patient adherence to physician-administered medications, and physician profiling. Combining Source 20 medical and pharmacy data in HealthVerity Marketplace will provide a detailed view of a patient's interactions with the healthcare system during the time that they are enrolled with the payer. The diversity of payers in

the sample also increases the chances of following a patient's journey across multiple payers and potentially from one insurance type to another.

Medical claims: Data relating to an encounter with a medical professional for which reimbursement is requested from the patient's insurer. Medical claims are typically submitted electronically using EDI 837/835 transactions.

Most Recent Transaction: 90 days old

Adjudication: Collected (post-adjudicated claims)

9.5.2. Europe and Japan data sources

9.5.2.1. ARTIS

The Swedish Biologics Register (ARTIS) is a registry of patients with rheumatologic diseases maintained by the Swedish Society for Rheumatology. Data in the register are used as a clinical decision-making tool for rheumatologists, who, together with the patient, enter the data. Despite the name including the term biologics, the register has collected information for patients receiving treatment with baricitinib. Data in the register were linked to various other registers through a unique personal identifier assigned to all Swedish residents. The National Patient Register provides information on diagnosis codes of outpatient visits in Swedish primary care since 2001. The Prescribed Drug Register is a nationwide public register with complete coverage of dispensing of prescribed drugs in Sweden, including quantity, dose, and date. The Total Population Register provides data on residency for all subjects who have ever resided in Sweden since 1961, and the Causes of Death Register is a national register that provides information on dates and causes of death for all deceased residents since 1961. ARTIS provides a high coverage and accuracy on bDMARD and tsDMARD exposures, including baricitinib, for patients with RA (Wadström et al. 2015).

9.5.2.2. Betriebskrankenkasse (BKK)

Germany's multi-payer healthcare system consists of a combination of health data from Sick fund providers (statutory health insurance [SHI]), that is, "Gesetzliche Krankenversicherung (GKV)," and private health insurance. Health insurance is mandatory for all citizens and permanent residents since 2009 and approximately 90% of the German population are members of SHI, entitled to comprehensive benefits, including inpatient and outpatient care, physician services, and prescriptions drugs. All prescription drugs are included unless explicitly excluded or pending evaluation.

Starting in 2004, the BKK contains healthcare claims data on more than 5 million patients (6-8% of the German GKV population) who are representative, in terms of age and sex, to the larger GKV population. The database includes information on patient demographics and inpatient and outpatient care, as well as outpatient medical prescriptions. Available demographic information includes sex, age, insurance status, time insured, and region of residence. Inpatient and outpatient diagnoses are coded via ICD-10, German Modification (ICD-10-GM) codes and procedures are available. For outpatient diagnoses, only the quarter of the diagnosis is reported (ie, the actual date is unavailable). Outpatient physician specialty information and inpatient

medical department of care can be accessed. Medical prescriptions from retail pharmacies are coded using the ATC hierarchy and the date of the prescription dispensing is available. Information on patient anthropometric data (eg, height, weight, body mass index [BMI]) is not available nor are laboratory test results (eg, absolute lymphocyte count, blood lipid levels) or clinical measurements (eg, blood pressure).

9.5.2.3. Cegedim THIN France

The THIN® France database is a longitudinal observational database established in 1994. The THIN® France database contains the anonymized electronic patient records of 2,000 GPs and 1000 specialists. These practitioners meet standard criteria regarding the quality of data entry: they have been selected to be representative of the global practitioner cohort in terms of gender, age and geographic locations.

The records consist of demographic (age, sex) and medical data including physical examination data (height, weight, blood pressure), diagnoses, drug prescriptions, laboratory test and medical procedures prescriptions and results, as well as reimbursement data for medical and paramedical procedures, including hospital outpatient consultations, specialist referrals, and hospitalizations, and drug delivery.

Each patient has a unique identification number associated with a THIN® France panel physician. All diagnoses are coded according to the International Classification of Diseases, tenth revision, Clinical Modification (ICD-10-CM). Drug prescriptions comprised information on trade names, formulations, and active substances, and are encoded with Anatomic Therapeutic Chemical (ATC) classification.

THIN is a Cegedim proprietary database of ambulatory electronic medical records; as such THIN patients are also part of the SNDS (which encompasses all patients in France and is described in Section 9.5.2.7). However available variables in THIN are much more extensive because they are based on medical records at physicians' level while SNDS variables are based on claim data.

As such, results from this data source are described individually, but were not included in the meta-analysis (as the patients were already contributing to the meta-analysis through SNDS).

9.5.2.4. CorEvitas – Japan

The CorEvitas Japan registry is patterned after the US CorEvitas RA registry (Section 9.5.1.3) but collects data from Japanese patients and physicians during routine clinic visits. Data are compatible across RA registries. As of September 2021, CorEvitas Japan had enrolled over 2199 patients with RA from 47 sites.

9.5.2.5. CPRD (GOLD and Aurum)

CPRD is a UK-based primary care electronic medical records (EMR) database of anonymized patient medical records representative of the UK general population in terms of age, sex, and ethnicity (Herrett et al. 2015). Both CPRD GOLD and CPRD Aurum EMR datasets were analysed for this study. Together, these datasets included 596 unique UK general practices and 15.9 million unique patients (Ghosh 2019), representing approximately 23.8% of the UK

population (WorldOMeters 2019). The CPRD includes information on demographics (including year of birth and sex), prescriptions and diagnoses. Dates of prescriptions and receipt of diagnoses are also available. As CPRD is a general practice EMR database, specialist hospital-generated prescription data have the potential to be missing. However, guidance from the British Medical Association (BMA) Medical Ethics Department advises specialists to inform patients' general practitioner of the results of the investigations, the treatment provided, and other information necessary for the continuing care of the patient (BMA 2009). Therefore, a number of specialist prescriptions are expected to be entered into CPRD upon receipt of this report by the general practitioner.

9.5.2.6. JMDC

This database, using data collected from medical institutions in Japan, consists of claims (for hospitalization and outpatient treatment), diagnosis procedure combination (DPC) assessment forms, and clinical laboratory test values. The oldest data in this database that can be accessed relate to treatment in April 2014. At the end of October 2019, the number of medical institutions was 218, consisting of 131 DPC-eligible hospitals and 87 DPC-ineligible hospitals. This database includes not only data from DPC-eligible hospitals but also data from some DPC-ineligible hospitals (Nagai et al. 2020).

9.5.2.7. Système National des Données de Santé (SNDS)

The French nationwide healthcare database, SNDS, covers more than 99% of the French population, near 67 million inhabitants, lifelong irrespective of socioeconomic status, even if a subject relocates, changes occupation, or retires. Using a unique pseudonymized identifier, the SNDS merges all reimbursed outpatient claims from all French healthcare insurance schemes with hospital-discharge summaries from public and private hospitals, and the national death registry (Bezin et al. 2017; Tuppin et al. 2017). The SNDS captures general characteristics (eg, gender, year of birth, area of residence); registration for Long Term Disease (LTD) qualifying for full health insurance coverage; outpatient encounter details (eg, medical and paramedical visits, medical and imaging procedures, laboratory tests, drug and medical device dispensing); inpatient details (eg, discharge diagnosis codes, medical procedures and paraclinical examinations, dates of hospital stay, drugs invoiced in addition to the stay). For each expenditure, dates, associated costs, and prescriber and caregiver information are provided.

Hospital discharge diagnosis and LTD are coded according to the International Classification of Diseases, 10th revision (ICD-10). Quality of this coding is ensured by regular internal and external audits (Gilleron et al. 2018). Reimbursed drugs are identified according to the Anatomic Therapeutic Chemical (ATC) classification. Drug brand, dosage and packaging are recorded using respectively presentation identifier codes (*Code Identifiant de Présentation*, CIP) in outpatient settings, and common dispensing units (*Unité commune de dispensation,* UCD) in hospital settings. Inpatient and outpatient procedures are coded according to the *classification commune des actes médicaux* (CCAM). Though, neither medical indication nor result are recorded, the level of details of the recorded information, enable an accurate characterization of patient healthcare journeys (Thurin et al. 2021).

9.6. Bias

In clinical trials, randomization creates treatment groups with similar baseline probability of experiencing the outcome(s) of interest. For an observational study, in the absence of randomization, rheumatologists and their patients select treatments based on characteristics that are also risk factors for the outcome of interest (eg, age, lifestyle factors, comorbidities, existence of refractory disease, etc. This can create confounding and lead to biased results if the prevalence of the risk factor differs between the groups being compared. In Study B023, confounding was addressed through the study design (ie, new user active comparator design [Lund et al. 2015]) and the analytic approach (ie, propensity score matching), which were used to create balance between groups with respect to risk factors for the outcomes. This approach works well for imbalances in risk factors that are available in the data. Confounding by unmeasured factors, such as BMI, smoking, and RA disease activity, may remain.

Information bias is another potential source of systematic error for claims data. Important limitations exist related to lack of detailed clinical detail, availability of variables for adjustment or matching, and the uncertain diagnostic validity of outcomes identified through billing codes. Because prescription records simply record dispensing, these data do not confirm patient exposure to medication. This can be addressed by ensuring that exposures are defined based on more than one prescription record. Limitations regarding the outcomes can be addressed through linkage to clinical information to validate the algorithms used to identify events and clarify the ability of the selected case definition to find true cases. Such linkage can also help address unmeasured confounding by providing insight into the importance of a potential confounder, absent from the main data, and then allowing adjustment of the overall analysis to correct for the unmeasured cofounder.

9.7. Study size

Naïve sample size and statistical power estimates were obtained from standard formulas for time-to-event studies that are based on a minimum number of events being observed. To ensure 80% power to detect a difference between treatment cohorts in the case of a true hazard ratio of 1.8, at least 90 patients with events are required cumulatively, over both baricitinib and TNFi cohorts, based on a 1:1 ratio of baricitinib users to TNFi users and a 1-sided Type I error rate of 0.025. Based on a background incidence rate of 05 to 0.9 per 100 PY for VTE, it was anticipated that at least the required number of patients with events will be observed in 6000 PY of exposure to baricitinib.

The final study size and statistical power depend on a number of factors, including the number of eligible patients in each data source, the follow-up time available in each data source, and the number of events observed during the follow-up.

9.8. Data transformation

Data were managed according to the standard procedures required by Eli Lilly and Company (Lilly) and the study partners with access to the health care data. Analytic data files from this study were originally intended to be shared with a regulatory agency. Therefore, prior to

initiating analyses, some data were transformed to safeguard patient privacy and ensure the anonymity of records per HIPAA. This applied to select US data and JMDC only.

Lilly analysed the CPRD data. Analysis of all other individual data sources was conducted by their respective organization, or a third-party platform as depicted in Table 9.3.

One such third party platform was Aetion, Inc. who executed analyses using the Aetion Evidence Platform. The platform is a data-handling technology that allows for the analysis of large patient datasets by indexing patient data into a form that can be queried by an internal patient variable language. When calendar dates were provided by an individual data vendor, there were no transformations to the data. In the circumstance where calendar dates are not provided within a data source and instead events have indicators relating the number of days a healthcare encounter occurred since index (either pre or post index date), the data were minimally transformed at the point of connection to the Aetion Evidence Platform in a way that the original format and temporal relationships in the source data were preserved. At the point of data connection to the platform some discard rules were applied depending on the data source. Specifically, patients were excluded if patient identifier (ID) was missing and patient events were excluded if there were no dates associated with them or if the start date of the event was preceded by the end date of the event (eg, discharge date precedes admission date for an inpatient event). Aetion IDs were assigned to source patient IDs and a crosswalk file is kept as a protected file available upon request to authorized parties. Individual-level patient data were analysed within the Aetion Evidence Platform, and aggregated results of the analyses will be exported from the platform.

Data partners who conducted analyses outside of the Aetion platform maintained appropriate data storage and performed appropriate quality control checks for all programming, consistent with their internal procedures.

IQVIA implemented a transformation to all date variables within the PharMetrics Plus data. This consisted of shifting all dates for each patient by a random number of days between -7 and +7. The same shift was applied to all dates for a given patient. Dates were not shift when the first or last claim date fell on the first or last date of the study period, respectively.

Table 9.3. Data Partner Conducting Patient-Level Analysis

Data Partner Conducting Analysis	US Data	Europe & Japan Data
Aetion, Inc	Aetna, Humana, Marketscan, Optum, HealthVerity PS20	Cegedim THIN (France), JMDC (Japan)
ARTIS (Karolinska University)	-	ARTIS (Sweden)
Bordeaux PharmacoEpidemiology Group (Université de Bordeaux)	-	SNDS (France)
CorEvitas, Inc.	CorEvitas US	CorEvitas (Japan)
HealthCore	Anthem's HIRD	-
IQVIA	MDR, PharMetrics Plus	BKK (Germany)
Lilly	-	CPRD (UK)

Abbreviations: CPRD = Clinical Practice Research Datalink; HIRD = HealthCore Integrated Research Database; IBM = International Business Machines; JMDC = JMDC, Inc's claims database; MDR = Military Health System Data Repository; NA = not applicable; OUS = outside the United States; RA = rheumatoid arthritis; SNDS = Système National des Données de Santé; UK = United Kingdom; US = United States.

9.9. Statistical methods

The statistical analysis plan for this study is detailed in the protocol (Annex 19), as well as a stand-alone SAP, both of which were shared with each contributing data partner.

As indicated in the SAP, changes in the comparative analysis strategy could be made based on the total number of patients and events in one or both comparison groups. Where appropriate, these changes are noted in Section 9.9.6.

Analyses were conducted separately for each outcome. The analysis population for each outcome includes all adult patients present in the data who meet the eligibility criteria. For all comparative analyses, baricitinib was the treatment of interest and the TNFi cohort was the reference group.

9.9.1. Main summary measures

9.9.1.1. Baseline characteristics

Within each data source, patient characteristics at baseline were summarized by exposure group (namely, baricitinib versus TNFi cohorts) in both unmatched and matched cohorts. For descriptive tables, cells with less than or equal to 5 or 10 counts (depending on data source) were redacted and reported as '≤10' to maintain patient confidentiality. Specifically, summaries were provided for the following categories of baseline characteristics, using measures appropriate for the variable type (counts, percentages, mean, median, range):

- patient demographics (age, sex)
- clinical histories (comorbidities, c/bDMARD use, other prescription medication use)
- healthcare resource utilisation (eg, number of outpatient and inpatient visits, emergency department visits),
- prevalence of outcomes of interest (VTE, MACE, serious infections, TB).

Tables presenting descriptive information for the unmatched cohorts focused on the characteristics of patients eligible for the main VTE analysis. The number of patients included in each outcome analysis varied since exclusion criteria were applied to help ensure that patients had similar baseline risk of experiencing the event at the time they enter the cohort. For example, patients with a prior history of the event under evaluation were excluded from the cohort (see Section 9.9.2.2 for detail), since those with a prior history are typically at increased risk of a recurrence relative to those without prior histories.

Standardised differences for each descriptive characteristic were also calculated in the unmatched and matched cohorts to assess the balance between groups before and after propensity score matching. Among matched patients, and in data sources where dose information

is available, baseline characteristics are also described by dose of baricitinib if there are at least 30 patients exposed to baricitinib 4 mg.

The covariates from the baseline period were considered for inclusion in the propensity score model.

9.9.1.2. Characteristics of patients under follow-up

Several characteristics of follow-up are described for data sources, when executed, including

- duration of follow-up for unmatched cohorts for analysis of VTE
- duration of follow-up for propensity score matched cohorts by outcome,
- descriptions of baseline characteristics for unmatched (VTE) and propensity score matched (all outcomes) patients stratified by duration of exposure in both cohorts: <6 months, 6 to <12 months, 12 months to <24 months, and ≥24 months.

9.9.2. Main statistical methods

9.9.2.1. Propensity score matching

Comparative analyses were implemented after propensity-score matching. Propensity scores aim to address imbalances 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.

Propensity scores were separately estimated for each outcome (Rosenbaum and Rubin 1983; Rubin 1997). They were estimated for each patient using a priori, multivariable logistic regression models. The treatment was specified as the dependent variable, and the model included all variables that are known risk factors for the outcomes of interest that are also associated with the exposures. Thus, the 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). Covariates considered for inclusion in the propensity score model for each outcome are provided in Table 4 of the protocol, available in Annex 19).

Propensity score matching was used to help ensure that confounding factors are evenly distributed across the exposure groups being compared. Propensity score matching was implemented using nearest-neighbour matching. The SAP detailed matching with a variable ratio of up to 1:3 based on a matching calliper of between 0.01 and up to 0.02 on the propensity score scale (Rassen et al. 2012; Austin 2011). However, as allowed per protocol to alter the application of propensity score matching, this strategy was changed to a fixed 1:1 matching. The rationale for this change is outlined in Section 9.9.6.

Patients with RA from regions outside of the US may be treated with baricitinib directly after cDMARD treatment, or they may switch to baricitinib therapy after a biologic such as a TNFi. Baricitinib-treated patients in non-US data will therefore consist of a mixture of TNFi-naïve and TNFi-experienced patients, with potentially varying proportions in different data sources. Other

relevant potential confounding factors will be included as they are for the propensity score models in US data.

Before initiating any outcome analyses, the ability of the propensity-score matching to balance the distribution of baseline confounders and reduce channelling bias was evaluated. Standardised differences were used to assess differences between the cohorts across measured baseline covariates before and after propensity score matching. Standardised differences less or equal to 0.10 indicated balance.

Exposure-specific propensity score distributions were plotted to inspect the suitability of the comparison group (Walker et al. 2013). The propensity score balance achieved within each database was inspected by displaying all patient characteristics by treatment status and by examining the absolute standardised differences (Austin 2011; Franklin et al. 2014). A post-matching C-statistic was also computed in some data sources as a summary metric for confounder balance. C-statistics close to 0.5 represent good overall balance (Franklin et al. 2013).

9.9.2.2. Exclusion criteria

The overall study inclusion criteria were detailed in Section 9.3. Because exclusion criteria were relevant to specific analyses (ie, applied separately for each outcome analysis), the exclusion criteria for comparative analyses are described here:

1. Use of JAK inhibitor during the 6 months prior to cohort entry
2. Concomitant use of more than 1 advanced therapy to treat RA (ie, dispensing of any combination of 2 or more bDMARDS and/or tsDMARDS at the time of cohort entry)
3. Patients with evidence of the outcome of interest (VTE, MACE, serious infection or hospitalized TB) in the 6 months prior to cohort entry were excluded from analysis of that respective outcome (but were considered for analysis for other outcomes)

For VTE and MACE analyses, the following additional exclusion criterion was applied

4. Use of an anticoagulant at the time of cohort entry

9.9.2.3. Comparative outcome analysis

Comparative analyses were performed separately on each of the outcomes: VTE, MACE, and serious infections and for within each data source separately. These analyses included only patients initiating their index treatment (either baricitinib or TNFi) and were based on aggregate baricitinib doses (2 mg and 4 mg), where dose information was available.

A variety of descriptive statistics were generated for the outcome of interest for the baricitinib and the TNFi cohorts, in all patients, unmatched eligible patients, and propensity score-matched patients included in the comparative analyses:

- Pattern of RA medication use, including drug switching and dosing, after cohort entry, by treatment cohort. The use of methotrexate was also described over the study period by treatment

- Distribution of survival time until first event of outcome of interest
- Crude IR and 95% confidence interval for the first ('incident') occurrence of the outcome of interest and stratified by baseline concomitant methotrexate use for the VTE and MACE analyses,
- Analysis of serious infections additionally summarized the distribution of the number of serious infections per patient and the incidence of serious infections prior to use of baricitinib and after commencing baricitinib.

The protocol described use of Cox proportional hazards regression as the primary statistical model for comparing risk of each outcome between the baricitinib and TNFi cohorts. However, due to low numbers of events and presence of zero events in numerous data sources, this strategy was changed to Poisson regression models. Rationale for which is outlined in Section 9.9.6. All models included the exposure cohort and any variables that remained unbalanced after propensity score matching.

Some medications commonly used in combination with b/tsDMARDs may themselves act to influence risk of an outcome. MTX, for example, has been recognized to increase risk of VTE (Methotrexate package insert, 2016) while potentially exhibiting cardioprotective effects against MACE (Marks and Edwards, 2012). For this reason, use of concomitant MTX treatment was evaluated by treatment group and inclusion of a time-dependent covariate was considered in regression models for both VTE and MACE. Similarly, glucocorticoid use raises risk of serious infection and may also serve as a proxy for disease severity. For this reason, glucocorticoid use was also included in an alternate model as a time-dependent covariate. In both cases, when possible based on counts, the time-dependent covariates were included in the final models if inclusion lead to a ≥10% change in the estimated hazard ratio.

9.9.3. Missing values

The protocol outlined options for imputation of missing values for variables used to adjust for confounding, or for generating propensity scores (see Section 8.6.1.1 of protocol for detail, Annex 19). However, imputation was not needed in analysis thus no imputation was conducted.

9.9.4. Sensitivity analyses

The protocol pre-defined numerous sensitivity analyses that could have been performed to examine the impact of study results to various factors and assumptions (see Section 8.7.9 of protocol, Annex 19). Due to the limited exposures and events in large majority of the data sources, most of the sensitivity analyses were not executed. The analyses that were conducted are outlined below, with the protocol including detail for those analyses that would have been executed if possible.

The exclusion criteria applied in sensitivity analyses, included:

1. Use of a JAK inhibitor during the 6 months prior to cohort entry.
2. Use of an anticoagulant at the time of cohort entry.
3. Concomitant use of more than 1 advanced therapy to treat RA (ie, dispensing of any combination of 2 or more bDMARDs and/or tsDMARDs at the time of cohort entry).

4. Patients with evidence of the outcome of interest (VTE, MACE, or serious infection, respectively) in the 6 months prior to cohort entry.

In US claims data, a sensitivity analysis was planned that would restrict inclusion to those with prior treatment with ≥1 TNFi if more than 5% of baricitinib-treated patients appeared to be TNFi-naïve. Due to the limited number of events in US data sources, this sensitivity analysis would not have been informative and was therefore not executed.

9.9.4.1. Descriptive analysis by dose for VTE

Where feasible based on having sufficient events and data by dose of baricitinib (ie, European and Japan data), a descriptive analysis was executed to estimate incidence rate of VTE by dose of baricitinib. Since the patients described in these analyses are not propensity-score matched to each other or to the comparator TNFi group, these incidence rates should not be compared with each other.

9.9.4.2. Quantitative bias analysis

This analysis aimed to estimate the direction, magnitude, and uncertainty in the study results that can arise due to systematic errors in the measurement or ascertainment of exposures (Lash et al. 2014). Specifically, the potential impact on study results of covariates for BMI and smoking status that are not available in claims data were assessed through this method. These analyses focused specifically on the potential impact to evaluating risk of VTE associated with RA therapy (baricitinib vs TNFi). The impact of disease severity to estimating risk of VTE was included subsequently as was a quantitative bias analysis focused on the potential impact to evaluating risk of MACE associated with baricitinib vs. TNFi.

This analysis attempts to explain how the strength of an unmeasured confounder, or risk factor, and differences in prevalence between the baricitinib and TNFi cohorts may have affected the observed or ARR. Using an array approach described by Schneeweiss (2006), various scenarios were modelled to understand whether (a) the strength of an unmeasured potential confounding factor (ie, smoking, BMI, or disease severity), and (b) an imbalance in the prevalence of each factor between treatment groups might have affected the IRRs calculated for VTE and MACE. The results are reported as an adjusted IRR and a quantitative measure of bias. Separate analyses were conducted for each outcome and for the observed, or ARRs generated in US and French data. The prevalence of smoking and obesity in each region was obtained, respectively, by medical record abstraction from patients with RA identified in the Optum Market Clarity data (US) and from the ESPOIR and CORPUS RA cohorts (Annex 17 and Annex 18).

The array approach is helpful for exploring the effect of residual confounding over a wide range of parameter estimates (ie, risk factor strengths and comparison group differences. The apparent and potentially confounded ARR can be expressed as a fully adjusted or 'true' RR by multiplying by the estimated bias (ie, $RR_{adjusted}$ = ARR x $Bias_M$), which is an expression of the imbalance of a binary confounding factor between 'exposed' or baricitinib (P_{C1}) and 'unexposed' or TNFi (P_{C0}) groups. Using the equation below and varying the values of RR_{CD} (ie, the magnitude of the association between the potential confounding factor and the outcome) and

those of PC1, but keeping ARR and P_{C0} fixed, the resulting RR reflects the "true" or adjusted RR after addressing unmeasured confounding .

$$.. = \frac{ARR}{\left[\frac{P_{C1}(RR_{CD}-1)+1}{P_{C0}(RR_{CD}-1)+1}\right]}$$

- **P_{C1}** – prevalence of the potential confounder (BMI or smoking) in the baricitinib cohort; varied in the sensitivity analyses
- **P_{C0}** – prevalence of the potential confounder (BMI or smoking) in the TNFi cohort; considered a fixed quantity
- **RR_{CD}** relative risk between confounder and outcome; varied in the sensitivity analysis
- **ARR** – Fixed quantity, the observed point estimate

Results can also be displayed graphically using 3-dimensional plots (see Figures X-Z in the results) which describe the relationship between (a) the adjusted IRR as a function of P_{C1}, holding RR_{CD} constant, and (b) the adjusted IRR as a function of RR_{CD}, holding P_{C1} constant. It is then possible to visualize to what extent it is likely that the strength of smoking or obesity on the likelihood of VTE (or MACE) and the imbalance of that factor between patients treated with baricitinib or TNFi, would plausibly produce the 'true' RR observed.

9.9.5. Meta-analysis

The results of comparative analyses for each outcome from each individual data source were combined through a meta-analysis. This combining of results via meta-analysis allowed for a more precise estimate of the effect size related to baricitinib exposure than was available from analyses of individual data sources.

The protocol detailed use of Cox regression and the DerSimonian-Laird method (see Section 8.7.8 of protocol, Annex 19) as the primary method for meta-analysis, with Poisson regression reserved as an alternative analysis strategy reserved for when zero events were observed in any treatment group. Due to the limited and often zero event counts across a large majority of data sources, the Poisson regression was selected as the preferred approach for meta-analysis. Thus, instead of generating an estimate on hazard ratio, a modified Poisson regression analysis was used to generate a pooled IRR for each outcome in patients treated with baricitinib versus TNFi. The rationale for this shift in analytic strategy is further described in Section 9.9.6.2.

The presence of heterogeneity in the treatment effect was assessed using the standard Cochran χ^2 test, and the magnitude of the heterogeneity was evaluated using the *I-squared* statistic (Higgins et al. 2003).

9.9.6. Amendments to the statistical analysis plan

9.9.6.1. Matching ratio

Per the study protocol and SAP, propensity score matching using a 1:3 ratio of baricitinib- to TNFi-treated patients was to be implemented to maximise the study's statistical power. However, the protocol anticipated feasibility challenges with 1:3 matching because of small and varying sample size ratios in each data source and allowed for "other applications of the

propensity score, such as changing the calliper permitted for matches, allowing matching with replacement, implementing variable ratio matching, or allowing full matching" (Section 8.7.4 in Annex 19), to be considered. Owing to the limited sample size and varying sample size ratios across the data sources – where standard 1:3 matching was possible in some data sources, but would have resulted in the removal of a substantial portion of the baricitinib-treated patients in other data – a fixed 1:1 propensity score-matching ratio was implemented across all data sources for this analysis. This fixed 1:1 ratio matching was feasible to implement in each data source and has the desirable effects of

- allowing for use of the same estimand from each data source for inclusion in the meta-analysis
- including the greatest proportion of the eligible baricitinib-treated patients for matching from each data source, which maintains broad generalizability of results, and
- weighting each data source appropriately in the meta-analysis (instead of depending in part on the matching process used in each data source).

9.9.6.2. Modified Poisson regression

Per the study protocol, HRs from Cox proportional hazards models from each data source were originally planned as the primary data for the meta-analysis. Poisson regression analysis was presented as an option for use secondary to the Cox regression models in the event of zero events in a treatment group. Upon review of event counts from several data sources, zero events in 1 or both treatment groups were identified for several. Because a Cox model would have excluded approximately half the data, the primary analytic approach was revised to rely on Poisson regression instead, as an option described in protocol (Section 8.7.8 in Annex 19). Poisson regression allows for inclusion of all exposed person-time, in the presence of sparse or zero events in one or even both treatment groups. A published report (Spittal et al. 2015) showed that the Poisson regression model estimated pooled IRR without bias, with good coverage, and generally had the lowest mean square error. These results held in a variety of situations, for instance, when

- only a small number of studies were included in each meta-analysis
- high baseline variability or high heterogeneity existed, and
- many zeros were present in the data (from none to approximately 80%).

Both random effects and fixed effect Poisson regression models were implemented. Only results from the fixed effect model are reported since random effect parameter estimates approached zero and provided estimates similar to the fixed effect parameters.

9.9.6.3. Incidence rate difference

In addition to the IRR calculated from Poisson regression, an updated Mantel-Haenszel meta-analysis of incidence rate differences (and 95% CI) was included as a supplemental analysis. The RD analysis includes available exposure time from all data sources and allows for estimation of individual study treatment differences even with zero events. Rate differences provide additional context that supplements the relative risk by providing information about the public health burden of an exposure. This is especially useful in the case of rare events.

9.10. Quality control

The research team documented the progress and scientific and quality review of all study activities and deliverables (for example, protocol, data management, data analysis, reports, manuscripts, etc.). This documentation recorded the major study tasks related to a specific study activity performed by data partners to develop and execute the requirements of the protocol. Quality assurance measures performed for each study activity during the conduct of the study were also recorded. This ensures 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 followed acceptable programming standards as agreed by Lilly and the data partners. Typically, the programming standards are a set of documents describing data extraction methods and provide a guideline for basic, frequently used terms and definitions and respective coding information to maintain operational consistency. Data validation occurred throughout the data management and analysis process. Data quality checks should typically include, but would not be limited to, independent programming checks by an individual who is not a main programmer for the study, internal consistency of the dataset, and checks to ensure that protocol criteria were met. If validation checks are not satisfied, then an examination of the problem is performed on the dataset(s) until the problem resolved. All data validation, quality checks, and resolution of issues identified and resolved would be documented.

As described in Section 9.8, the analysis of the 15 contributing data sources was conducted by 6 analytic partners. The quality processes followed by each of the 6 analytic partners is described below.

9.10.1. Aetion

Analysis for 6 individual data sources was conducted within the Aetion Evidence Platform (Table 9.3).

Data validation process at Aetion

When data is received at Aetion from a vendor, the Data Operations team creates a templated data ingestion ticket and tags all responsible team members. Details of the ingestion ticket include the location where the data was uploaded, the location where the data should be transferred, and the date and time of the upload. All data is retained and stored by client, dataset, and revision.

The schema of the data received is compared to the expected schema per vendor documentation for new data sets and to previously received data for data updates to identify missing variables or new variables. Event density distributions are created for new data sets and data updates in order to explore event data over time and identify possible gaps or missing files. For some frequently received data, lists of files received are compared between data updates in order to identify files that have been added or unexpectedly dropped.

The conversion of the data into the Aetion data model is thoroughly validated as part of quality control processes. Starting with the source data received from the data vendor or customer, the Data Engineering team applies transformations according to logic defined in the database connector specification document. The validation is performed in steps, first on a sample of the data, then on the entire data. The validation is performed on every patient and event attribute and includes a comparison and checks of the converted data through a double-programming approach. The Data Engineering team produces results via scala code, which reflect the data as they would be available in the platform after transformation. The Data Science team independently implements the logic defined in the database connector specification using SQL and compares the results to the spark-based ones, to ensure that the connector was coded correctly. Any discrepancies are investigated and resolved. The data are not deployed to the Platform until the validation and QC process is completed.

A rule-based validation framework is applied in order to provide automation, flexibility, transparency and scalability to the validation process. Instead of individual queries being used to perform validation checks, functions and rules are used to generate an automated validation report that compares the results generated by the Data Science and Data Engineering teams and flags any discrepancies.

All code is version-controlled and stored in Github by dataset, client, and revision. All raw data, code, transformed data, metadata, and validation reports are stored in a client-specific bucket and organized by dataset and revision in a structured folder system. The code and data used for a particular data connection or data update can be accessed at any time by users of the Data Integration and Engineering teams with appropriate permissions in s3.

Scientific quality control processes at Aetion

All measures created, cohorts developed, statistical analyses implemented, and tables populated undergo rigorous quality control review. Creation of each study variable, patient cohorts, and analysis is performed by an individual, and then quality checked by a second individual to ensure accurate and reproducible implementation under the supervision of a senior scientific leader.

Quality control methods include:

- Checks for the validity and logical content of measure definition components
- Checks for missing values and variables
- Tabulation of values to identify potential inconsistencies and errors,
- Examination of the distribution of values for each variable, including potential outliers.

Aetion employs the following quality control processes on results:

- Results output is populated into tables. One individual enters the information, while a second individual quality checks those table entries for accuracy.
- The senior scientific study leader then reviews all tables for consistency, accuracy, and clarity.

Aetion prides itself on the delivery of transparent, auditable, and comprehensive reports. The completion of such reports is facilitated by key features of AEP. AEP's audit trail maintains study version summaries, including variable, cohort, and analysis selections. Any data transformations and methodological decisions made in an analysis on AEP are fully and automatically documented. All archived documentation is stored in complete, readable form, and conforms to standards set by organizations such as ISPE and ISPOR.

9.10.2. ARTIS

The ARTIS register linkage data is obtained from Socialstyrelsen (The National Board of Health and Welfare) who perform independent quality control checks on the delivered data. ARTIS team members perform quality control checks according to a pre-specific manual on this data when received. Programming checks for the specific project are performed by a member of the ARTIS team who is not the primary analyst. A quality control checklist is filled out by this independent team member, which ensures that the programming is correct, definitions are in alignment with the SAP/protocol, results/reports are correct and consistent, and all relevant research documentation is completed and accurate. The results delivered are reviewed by at least 2 independent ARTIS team members, to ensure that the results are also consistent with other reports delivered by ARTIS.

The ARTIS PI is responsible for ensuring that the quality system is followed. The ARTIS team, together with internal experts, are responsible for the continuous development of quality system, regularly follow-up and internal audit-like assessments. The project manager is responsible for implementing quality system adapted to each project's conditions. All ARTIS employees are responsible for complying with the quality system and communicating shortcomings and suggesting improvements through the project specific deviation log.

9.10.3. BPE - Bordeaux PharmacoEpi (French SNDS)

The Bordeaux PharmacoEpi, INSERM CIC1401, has implemented a quality management system for all its activities and is certified ISO 9001:v2015 for its activities in pharmacoepidemiology research. An independent double programming was performed for main criteria analyses, and the results compared for validation. All statistical logs are kept and can be provided.

9.10.4. CorEvitas (US and Japan)

All data gathering and analyses will be overseen by 2 pharmacoepidemiologists experienced in the field of registry-based research. Programming for this project will be conducted by a primary statistician and second, validation statistician. In addition, the full project will be overseen and reviewed by a senior statistician and all AE information and counts are validated by a senior pharmacovigilance expert. Programming associated with all (primary and sensitivity analysis) cohort creation, key variable definition, and primary analysis results generation will be double-coded. All code will be reviewed by the validation statistician as well as the senior statistician.

9.10.5. Anthem (HIRD)

The research team will document the progress and scientific and quality review of all study activities and deliverables (eg, protocol, data management, data analysis, reports, manuscripts, etc.) in a QC Log. The QC 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 QC Log documents the quality assurance measures performed for each study activity during the conduct of the study. This is necessary to document and track edits and revisions to project documents and ensure that the most up to date versions of relevant documents are readily identifiable.

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 QC Log.

9.10.6. IQVIA (PharMetrics Plus, MDR, and BKK)

Analysis of 3 individual data sources was conducted by IQVIA (Table 9.3).

Datasets and analytic programs were stored according to IQVIA procedures with access restricted to study personnel. IQVIA confidentiality agreements were signed by all employees and included data protection and strict prohibitions on reidentification attempts.

All aspects of the study were conducted within the framework of the IQVIA Quality Management System. A QC plan for the study was developed and executed, which included QC on SAP addendum, programming, data management and analysis, and study results. Furthermore:

- The study QC plan established ownership for the execution of the individual QC steps
- The Principal in Charge of the study ensured that individuals responsible for the execution of specific QC steps had the knowledge, capability, and experience necessary to perform the assigned tasks, and
- The result of the execution of the individual steps of the QC plan was documented, and included the required corrective actions, if any. The execution of any required corrective action was also documented.

The QC plan was subjected to a final review and approval for sufficiency and completeness from the Principal in Charge of the study.

10. Results

10.1. Overall summary

10.1.1. Descriptive data

Detailed descriptions of the cohorts that were analysed from each data source that contributed to the meta-analysis include:

US Data

- Aetna (Healthagen) (Section 10.2.1.1)
- Anthem (HIRD) (Section 10.2.1.2)
- CorEvitas US (Section 10.2.1.3)
- Humana (Section 10.2.1.4)
- MarketScan (Section 10.2.1.5)
- MDR (Section 10.2.1.6)
- Optum (Section 10.2.1.7)
- PharMetrics Plus (Section 10.2.1.8)
- PS20 (Section 10.2.1.9)

Europe and Japan Data

- ARTIS (Section 10.2.2.1)
- BKK (Section 10.2.2.2)
- Cegedim (Section 10.2.2.3)
- CorEvitas Japan (Section 10.2.2.4)
- JMDC (Section 10.2.2.5)
- SNDS (Section 10.2.2.6)

10.1.2. Patient summary

In total, 9,013 patients with RA treated with baricitinib from 14 data sources across the US, Japan, France, Germany, and Sweden, met preliminary eligibility criteria and 7,606 were retained in comparative analyses of VTE after propensity score matching to TNFi patients. The contribution of respective data sources to the overall study are described in Table 10.1 for VTE. Because eligibility criteria are specific to analyses of each outcome, the number of patients and exposures will vary by outcome. Only contributions for VTE analyses are presented since the relative contribution for other outcomes did not differ meaningfully. Because analyses were designed and conducted separately in each data source, neither patients nor person-time should be combined other than to assess the size of this study.

Table 10.1. Contribution of patients and person-time to VTE analyses, by data source

Data Source[a]	Country	Eligible (Unmatched) Patients		Patients in Analyses[b]			
		Baricitinib (n)	TNFi (n)	Baricitinib (n)	Baricitinib Exposure PY (% total)	TNFi (n)	TNFi Exposure (PY)
US Data							
Aetna/Healthagen	US	69	289	37	12.8 (<1)	37	22.2
Anthem (HIRD)	US	255	1,304	123	69.4 (1)	123	99
CorEvitas US	US	118	1,897	112	76.2 (1)	112	84.6
HealthVerity PS20	US	933	3,953	748	235.5 (4)	748	377.5
Humana	US	89	154	49	19.8 (<1)	49	20.6
Marketscan	US	257	1,599	185	84.4 (1)	185	77.6
MDR	US	188	1,686	114	61 (1)	114	70
Optum	US	348	1,441	284	118.0 (2)	284	163.1
PharMetrics Plus	US	473	6,576	261	141 (2)	261	159
Europe and Japan Data							
ARTIS	Sweden	1,737	6,230	1,685	2,313.6 (39)	1,685	2,608.3
BKK	Germany	851	3,332	765	539 (9)	765	544
CoreEvitas JP	Japan	210	354	171	199.5 (3)	171	247.8
JMDC	Japan	243	1,721	213	154.0 (3)	213	115.0
SNDS	France	3,242	10,202	2,859	1,855 (32)	2,859	1,923
TOTAL		**9,013**	**40,738**	**7,606**	**5,879.2**	**7,606**	**6,511.7**

Abbreviations: ARTIS = Anti-Rheumatic Therapy in Sweden; BKK = Betriebskrankenkasse; CPRD = Clinical Practice Research Datalink; HIRD = HealthCore Integrated Research Database; JMDC = Japanese Medical Data Center, Inc.'s claims database; JP = Japan; MACE = major adverse cardiovascular event; PY = person-years; SNDS = Système National des Données de Santé; THIN = The Health Improvement Network; TNF = tumour necrosis factor.

[a] All data sources included in the meta-analysis are summarised here, except: Cegedim THIN France, whose patients are a subset of the SNDS data, and CPRD, which has insufficient exposure in the eligible pre-matched patients to permit meaningful analyses.

[b] Available information on sample size (n) and baricitinib exposure (PY) among eligible propensity-score-matched patients analysed to assess risk of VTE.

Detailed baseline characteristics are presented for each data source in Section 10.2. This includes descriptions for both the baricitinib and TNFi cohorts. Selected baseline characteristics for data sources contributing ≥100 total PY baricitinib exposure are summarized in Table 10.2. This summary is specific to the matched baricitinib cohort from the VTE analysis. The average age of patients treated with baricitinib included in analyses ranged from 51.55 (JMDC) to 60.9 years (CorEvitas Japan). Patients from the ARTIS and SNDS data contributed the greatest proportion of baricitinib exposure to the VTE meta-analysis, with 39% (2313.6 PY) and 32% (1855 PY) of the total. The longest median baricitinib exposure (in days) occurred in ARTIS and CorEvitas

Japan patients with 454 and 385 days, respectively. Selected baseline characteristics (comorbidities, comedication use, and disease severity) should be interpreted cautiously as:

(a) treatment cohorts in each data source have been modified by propensity-score-matching,

(b) variables were not defined consistently across all data sources, depending on the data source and regional coding scheme.

The distribution of the selected risk factors presented is not intended to be the same across data sources, but rather within each data source when comparing the baricitinib to the TNFi cohort. This was achieved through analytic methods applied within the respective data source (ie, Propensity score matching to TNFi cohort and subsequent adjustment in the regression model).

Table 10.2. Selected Baseline Characteristics of Baricitinib-Treated Patients in Propensity Score-Matched VTE Cohorts, by Data Source.

Patient Characteristic	US Data			Europe and Japan Data				
	Health Verity PS20	Optum	PharMetrics Plus	ARTIS	BKK	CorEvitas Japan	JMDC	SNDS
N, baricitinib	748	284	261	1,685	765	171	213	2,859
Exposure								
Total (PY)	235.5	118	141	2,313.60	539	199.5	154	1,855
mean (SD), days	114.56 (111.77)	151.68 (139.04)	197.5 (168)	502 (345.7)	256.5 (211)	426.1 (253.2)	263.93 (209.95)	236.8 (195.2)
median, days	68.5	96	141	454	194	385	216	173
min, max, days	0.0, 718.0	2.0, 707.0	1.0, 814.0	1, 1310	2.0, 962.0	12.0, 1071.0	2.0, 934.0	0.0, 831.0
Demographic								
Mean age, years (SD)	54.85 (11.26)	58.69 (11.94)	53.5 (11)	59 (13.6)	56.5 (13)	60.9 (13.6)	51.55 (10.20)	58.4 (13.2)
Sex, female (%)	643 (86.0%)	244 (85.9%)	218 (84%)	1382 (82%)	572 (75%)	145 (84.8%)	169 (79.3%)	2268 (79.3%)
RA-related								
RA Severity (SD)	NA	NA	NA	DAS28: 4.6 (1.27)	NA	CDAI: 23.4 (13.0)	NA	NA
RA Severity: CIRAS index (SD)	4.18 (1.23)	4.21 (1.28)	4.4 (1)	NA	7.1 (2)	NA	6.59 (1.30)	6.5 (1.4)
cDMARDs, baseline (%)	352 (47.1%)	177 (62.3%)	163 (62%)	935 (55%)	441 (58%)	105 (61.4%)	180 (84.5%)	1945 (68.0%)
cDMARDs, concomitant with bDMARD, baseline	158 (21.1%)	50 (17.6%)	69 (26%)	NA	100 (13%)	NA	78 (36.6%)	910 (31.8%)
Glucocorticoids, baseline (%)	349 (46.7%)	172 (60.6%)	211 (81%)	1073 (64%)	549 (72%)	40 (23.4%)	131 (61.5%)	2026 (70.9%)
Comorbid conditions								
Diabetes (%)	145 (19.4%)	57 (20.1%)	35 (13%)	141 (8%)	124 (16%)	19 (11.1%)	6 (2.8%)	283 (9.9%)

Patient Characteristic	US Data			Europe and Japan Data				
	Health Verity PS20	Optum	PharMetrics Plus	ARTIS	BKK	CorEvitas Japan	JMDC	SNDS
Dyslipidaemia (%)	269 (36.0%)	99 (34.9%)	72 (28%)	<5	203 (27%)	22 (12.9%)	16 (7.5%)	NA
Hypertension (%)	321 (42.9%)	133 (46.8%)	89 (34%)	12 (1%)	349 (46%)	48 (28.1%)	0 (0.0%)	NA
Anti-hypertensive medication (%)	265 (35.4%)	155 (54.6%)	121 (46%)	718 (43%)	326 (43%)	40 (23.4%)	49 (23.0%)	973 (34.0%)
Obesity (%)	212 (28.3%)	70 (24.6%)	52 (20%)	NA	125 (16%)	14 (8.2%)	0	NA
Smoking (%)[b]	111 (14.8%)	46 (16.2%)	19 (7%)	803 (48%)[c]	93 (12%)	18 (10.5%)[c]	9 (4.2%)	NA

Abbreviations: ARTIS = Anti-Rheumatic Therapy in Sweden; bDMARDs = biologic disease-modifying antirheumatic drugs; BKK = Betriebskrankenkasse; CDAI = clinical disease activity index; cDMARDs = conventional synthetic disease-modifying antirheumatic drugs; CIRAS = claims-based index of rheumatoid arthritis severity; DAS28 = Disease Activity Score 28; HIRD = HealthCore Integrated Research Database; JMDC = JMDC, Inc's claims database; N = number in specific category; max = maximum; min = minimum; PS20 = Private Source 20; PY = person-years exposure; RA = rheumatoid arthritis; SD = standard deviation; SNDS = Système National des Données de Santé; US = United States

[a] Disease severity is based on the CIRAS index in US insurance claims data: Aetna, HIRD, Humana, MarketScan, Optum; CDAI in CorEvitas US and Japan, BKK, SNDS; and DAS28 for ARTIS data.

[b] Smoking status is not directly available in claims data sources and is based on ICD-10 codes related to smoking cessation and other measures such as tobacco use disorder, counselling visits for smoking, and antismoking prescription medications. In registry-based data (ARTIS, CorEvitas US and Japan), smoking status is collected directly from patients. In JMDC, smoking is defined based on information recorded in the variable "Annual health checkup – Smoking habit".

[c] Smoking is defined as follows ARTIS: current or former smoker, and CorEvitas Japan: current smoking.

10.1.3. Overall results VTE

A total of 97 patients with VTE were identified in the analytic cohorts, 56 of whom were treated with baricitinib. The IR of VTE in the baricitinib and TNFi cohorts from each data source are presented in Table 10.3. For VTE the IRR was significantly elevated overall (IRR=1.51; 95% CI 1.10, 2.08). The IR was greater among patients treated with baricitinib than with TNFi, the difference was 0.26 (95% CI -0.04, 0.57) per 100 PY. This IRD did not reach statistical significance. Risk of VTE was elevated, but not statistically significant, in the two largest data sources (ARTIS and SNDS), which contributed approximately two-thirds of exposure to the meat-analysis (IRR_{ARTIS} = 1.85; 95% CI 0.95, 3.60 and IRR_{SNDS} = 1.59; 95% CI 0.79, 3.21). With 9% of total baricitinib exposure, the IRR from German BKK data was 0.50, but the CI was wide (95% CI 0.13, 2.02).

Table 10.3. Comparison of VTE between Baricitinib- and TNFi-Treated Patients in Multiple Databases

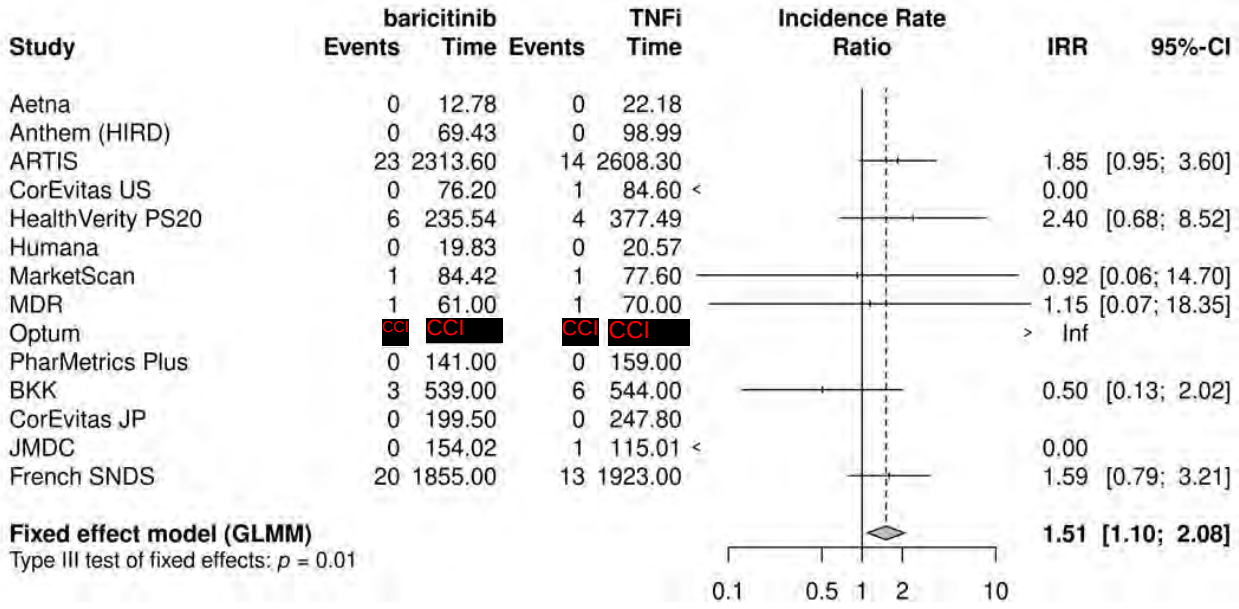
Data Source[a]	Baricitinib **n** [Events, PY]	TNFi **n** [Events, PY]	$IR_{baricitinib}$, IR_{TNFi} **IRD (95% CI)** per 100 PY	HR **(95% CI)**[a]	IRR **(95% CI)**
			All Available Data		
Overall	**-**	**-**	**IRD = 0.26 (-0.04, 0.57)**	**-**	**1.51 (1.10, 2.08)**
			US Data		
Aetna /Healthagen	**37** [0, 12.8]	**37** [0, 22.2]	$IR_{baricitinib}$ = 0 (0, 28.88) IR_{TNFi} = 0 (0, 16.64) IRD = 0 (-11.51, 11.51)	-	-
Anthem (HIRD)	**123** [0, 69.4]	**123** [0, 99.0]	$IR_{baricitinib}$ = 0 (0, 5.31) IR_{TNFi} = 0 (0, 3.73) IRD = 0 (-2.40, 2.40)	-	-
CorEvitas US	**112** [0, 76.2]	**112** [1, 84.6]	$IR_{baricitinib}$ = 0 (0, 4.8) IR_{TNFi} = 1.2 (0, 6.6) IRD = -1.18 (-4.49, 2.12)	-	0
HealthVerity PS20	**748** [6, 235.5]	**748** [4, 377.5]	$IR_{baricitinib}$ = 2.55 (0.94, 5.54) IR_{TNFi} = 1.06 (0.29, 2.71) IRD = 1.49 (-0.77, 3.75)	2.15 (0.60, 7.71)	2.40 (0.68, 8.52)
Humana	**49** [0, 19.8]	**49** [0, 20.6]	$IR_{baricitinib}$ = 0 (0, 18.60) IR_{TNFi} = 0 (0, 17.94) IRD = 0 (-9.14, 9.14)	-	-
MarketScan	**185** [1, 84.4]	**185** [1, 77.6]	$IR_{baricitinib}$ = 1.19 (0.03, 6.60) IR_{TNFi} = 1.29 (0.03, 7.18) IRD = -0.10 (-3.51, 3.31)	0.79 (0.05, 12.62)	0.92 (0.06, 14.70)
MDR	**114** [1, 61]	**114** [1, 70]	$IR_{baricitinib}$ = 1.6 (0.2, 11.7) IR_{TNFi} =1.40 (0.2, 10.1) IRD = 0.21 (-4.02, 4.44)	1.23 (0.1, 20.7)	1.15 (0.07, 18.35)
Optum	**284** [CCI, CCI]	**284** [CCI, CCI]	$IR_{baricitinib}$ = 1.69 (0.21, 6.12) IR_{TNFi} = 0 (0, 2.26) IRD = 1.69 (-1.02, 4.41)	-	-
PharMetrics Plus	**261** [0, 141]	**261** [0, 159]	$IR_{baricitinib}$ = 0 (0, 2.62) IR_{TNFi} = 0 (0, 2.32) IRD = 0 (-1.30, 1.30)	-	-

Data Source[a]	**Baricitinib** **n** [Events, PY]	**TNFi** **n** [Events, PY]	**$IR_{baricitinib}$, IR_{TNFi}** **IRD (95% CI)** per 100 PY	**HR (95% CI)[a]**	**IRR (95% CI)**
All Available Data					
Overall	**-**	**-**	**IRD = 0.26 (-0.04, 0.57)**	**-**	**1.51 (1.10, 2.08)**
Europe and Japan Data					
ARTIS	**1685** [23, 2313.6]	**1685** [14, 2608.3]	$IR_{baricitinib}$ = 0.99 (066, 1.50) IR_{TNFi} = 0.54 (0.32, 0.91) IRD = 0.46 (-0.03, 0.95)	1.83 (0.95, 3.55)	1.85 (0.95, 3.60)
BKK	**765** [3, 539]	**765** [6, 544]	$IR_{baricitinib}$ = 0.6 (0.1, 1.6) IR_{TNFi} = 1.1 (0.4, 2.4) IRD = -0.55 (-1.63, 0.53)	0.49 (0.1, 2.0)	0.50 (0.13, 2.02)
CorEvitas JP	**171** [0, 199.5]	**171** [0, 247.8]	$IR_{baricitinib}$ = 0 (0, 1.8) IR_{TNFi} = 0 (0, 1.5) IRD = 0 (-0.89, 0.89)	-	-
JMDC	**213** [0, 154.0]	**213** [1, 115.0]	$IR_{baricitinib}$ = 0 (0, 2.40) IR_{TNFi} = 0.87 (0.02, 4.84) IRD = -0.87 (-3.11, 1.37)	-	0
SNDS (France)	**2859** [20, 1855]	**2859** [13, 1923]	$IR_{baricitinib}$ = 1.1 (0.7, 1.7) IR_{TNFi} = 0.7 (0.4, 1.2) IRD = 0.40 (-0.19, 1.00)	1.57 (0.78, 3.18)	1.59 (0.79, 3.21)

Abbreviations: ARTIS = Anti-Rheumatic Therapy in Sweden; BKK = Betriebskrankenkasse; CI = confidence interval; $IR_{baricitinib}$ = incidence rate for the baricitinib-treated cohort; IR_{TNFi} = incidence rate for the TNFi-treated cohort; IRD = incidence rate difference; HIRD = HealthCore Integrated Research Database; HR = hazard ratio; IR = incidence rate; IRD = incidence rate difference; IRR = incidence rate ratio; JMDC = Japanese Medical Data Center, Inc.'s claims database; JP = Japan; MDR = Military Data Repository; PY = person-years; SNDS = Système National des Données de Santé; TNFi = tumour necrosis factor inhibitor; VTE = venous thromboembolism.

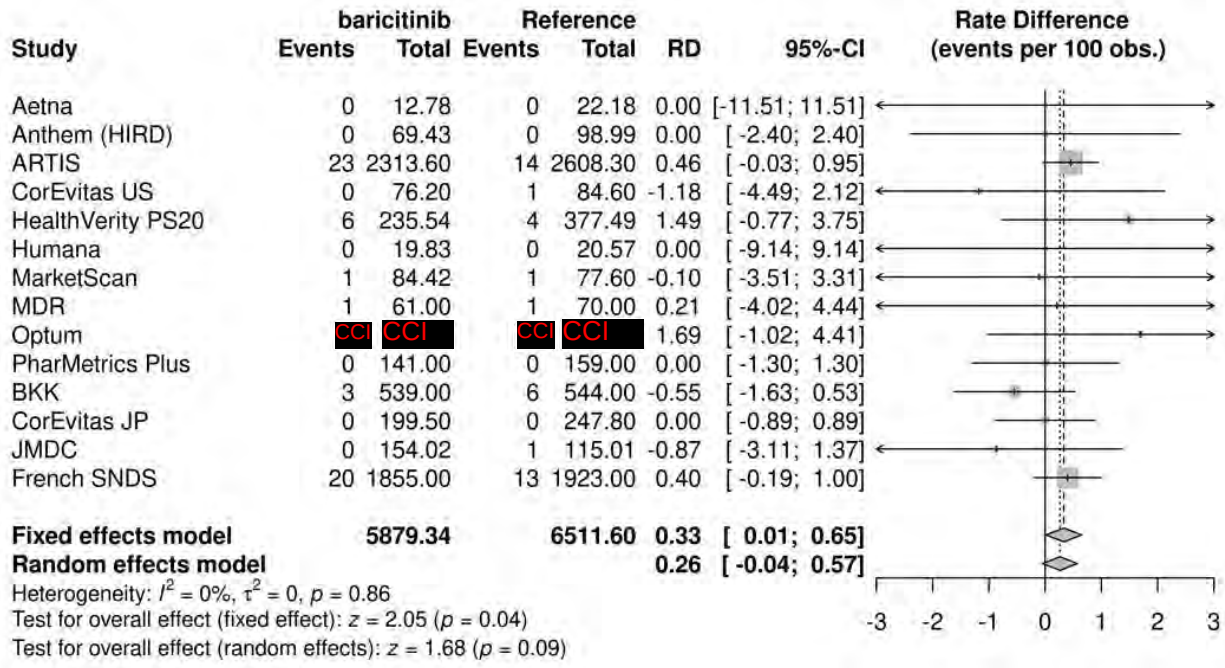
Note: Any differences between confidence intervals for the individual IRD reported here and in Section 10.3 are due to different methods used for the calculation.

[a] Base or minimally adjusted HR.



Abbreviations: BKK = Brietriebskrankenkasse; CI = confidence interval; GLMM = generalised linear mixed model; HIRD = HealthCore Integrated Research Database; IRR = incidence rate ratio; JMDC = JMDC, Inc's claims database; JP = Japan; MDR = Military Data Repository; SNDS = Système National des Données de Santé; TNFi = tumour necrosis factor inhibitor; VTE = venous thromboembolism.
Source: lillyce\prd\ly3009104\i4v_mc_b023\final\output\shared\irr_vte_14ds.png

Figure 10.1. Meta-analysis on incidence rate ratios for VTE comparing baricitinib and TNFi.



Abbreviations: ARTIS = Anti-Rheumatic Therapy in Sweden; BKK = Betriebskrankenkasse; CI = confidence interval; HIRD = HealthCore Integrated Research Database;

= incidence rate difference; JMDC = JMDC, Inc's claims database; JP = Japan; MDR = Military Data Repository; obs. = observations; PS20 = HealthVerity Private Source 20; RD = incidence rate difference; SNDS = Système National des Données de Santé; TNFi = tumour necrosis factor inhibitor; VTE = venous thromboembolism.
Note: Any differences between confidence intervals for the individual incidence rate differences reported here and in Section 10.3 are due to different methods used for the calculation.
Source: lillyce\prd\ly3009104\i4v_mc_b023\final\output\shared\final_ird_vte3.png

Figure 10.2. Meta-analysis on incidence rate differences in VTE between baricitinib and TNFi.

10.1.4. Overall results MACE

There were a total of 93 patients with MACE, 54 of whom were treated with baricitinib. The IR of MACE in the baricitinib and TNFi cohorts from each data source are presented in Table 10.4. For MACE the overall IRR was greater for baricitinib compared to TNFi but did not attain statistical significance (IRR=1.54; 95% CI 0.93, 2.54). The IR was greater among patients treated with baricitinib than with TNFi, the difference was 0.22 (95% CI -0.07, 0.52) per 100 PY. This IRD was not statistically significant since the CI includes 0. A statistically significantly increased risk of MACE was found in SNDS (IRR$_{SNDS}$ =2.33; 95% CI$_{SNDS}$ 1.15, 4.7) and not in the largest data source (IRR$_{ARTIS}$ = 0.94; 95% CI$_{ARTIS}$ 0.5, 1.96).

Table 10.4. Comparison of MACE between Baricitinib- and TNFi-Treated Patients in Multiple Databases

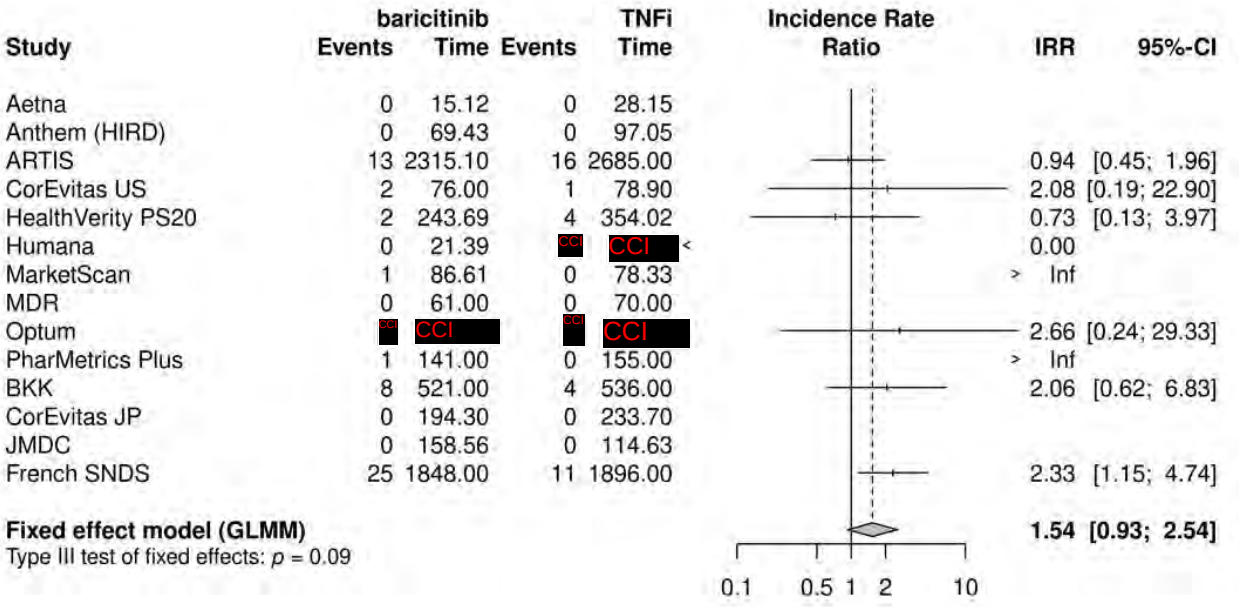
Data Source[a]	Baricitinib n [Events, PY]	TNFi n [Events, PY]	$IR_{baricitinib}$, IR_{TNFi}, IRD (95% CI) per 100 PY	HR (95% CI)[a]	IRR (95% CI)
Overall	-	-	**IRD = 0.22 (-0.07, 0.52)**	-	**1.54 (0.93, 2.54)**
					US Data
Aetna /Healthagen	**43** [0, 15.1]	**43** [0, 28.2]	$IR_{baricitinib}$ = 0 (0, 24.4) IR_{TNFi} = 0 (0, 13.10) IRD = 0 (-9.69, 9.69)	-	-
CorEvitas US	**114** [2, 76.0]	**114** [1, 78.9]	$IR_{baricitinib}$ = 2.6 (0.3, 9.5) IR_{TNFi} = 1.3 (0, 7.1) IRD = 1.36 (-3.00, 5.73)	1.60 (0.11, 23.61)	2.08 (0.19, 22.90)
Anthem (HIRD)	**123** [0, 69.4]	**123** [0, 97.1]	$IR_{baricitinib}$ = 0 (0, 5.31) IR_{TNFi} = 0 (0, 3.80) IRD = 0 (-2.42, 2.42)	-	-
HealthVerity PS20	**743** [2, 243.7]	**743** [4, 354.0]	$IR_{baricitinib}$ = 0.82 (0.10, 2.97) IR_{TNFi} = 1.13 (0.31, 2.89) IRD = -0.31 (-1.89, 1.27)	0.73 (0.13, 3.99)	0.73 (0.13, 3.97)
Humana	**51** [0, 21.4]	**51** [[redacted], CCI]	$IR_{baricitinib}$ = 0 (0, 17.25) IR_{TNFi} = 7.85 (0.95, 28.36) IRD = -7.85 (-20.56, 4.86)	-	0
MarketScan	**192** [1, 86.6]	**192** [0, 78.3]	$IR_{baricitinib}$ = 1.15 (0.03, 6.43) IR_{TNFi} = 0 (0, 4.71) IRD = 1.15 (-2.07, 4.38)	-	-
MDR	**114** [0, 61]	**114** [0, 70]	$IR_{baricitinib}$ = 0 (0, 6.05) IR_{TNFi} = 0 (0, 5.27) IRD = 0 (-2.96, 2.96)	-	-
Optum	**287** [[redacted], CCI]	**287** [[redacted], CCI]	$IR_{baricitinib}$ = 1.65 (0.20, 5.95) IR_{TNFi} = 0.62 (0.02, 3.45) IRD = 1.03 (-1.54, 3.59)	2.63 (0.24, 29.38)	2.66 (0.24, 29.33)

Data Source[a]	**Baricitinib** **n** [Events, PY]	**TNFi** **n** [Events, PY]	**$IR_{baricitinib}$, IR_{TNFi}, IRD (95% CI)** per 100 PY	**HR (95% CI)[a]**	**IRR (95% CI)**
			All Available Data		
Overall	**-**	**-**	**IRD = 0.22 (-0.07, 0.52)**	**-**	**1.54 (0.93, 2.54)**
PharMetrics Plus	**262** [1, 141]	**262** [0, 155]	$IR_{baricitinib}$ = 0.7 (0.1, 5.0) IR_{TNFi} = 0 (0, 2.38) IRD = 0.71 (-1.19, 2.61)	-	-
Europe and Japan Data					
ARTIS	**1681** [13, 2315.1]	**1681** [16, 2685.0]	$IR_{baricitinib}$ = 0.56 (0.33, 0.97) IR_{TNFi} = 0.60 (0.37, 0.97) IRD = -0.03 (-0.46, 0.39)	0.92(0.45, 1.90)	0.94 (0.45, 1.96)
BKK	**757** [8, 521]	**757** [4, 536]	$IR_{baricitinib}$ = 1.5 (0.7, 3.0) IR_{TNFi} = 0.7 (0.2, 1.9) IRD = 0.79 (-0.49, 2.07)	2.09 (0.6, 6.9)	2.06 (0.62, 6.83)
CorEvitas JP	**168** [0, 194.3]	**168** [0, 233.7]	$IR_{baricitinib}$ = 0 (0, 1.9) IR_{TNFi} = 0 (0, 1.6) IRD = 0 (-0.92, 0.92)	-	-
JMDC	**224** [0, 158.6]	**224** [0, 114.6]	$IR_{baricitinib}$ = 0 (0, 2.33) IR_{TNFi} = 0 (0, 3.22) IRD = 0 (-1.48, 1.48)	-	-
SNDS (France)	**2864** [25, 1848]	**2864** [11, 1896]	$IR_{baricitinib}$ = 1.4 (0.9, 2.0) IR_{TNFi} = 0.6 (0.3, 1.0) IRD = 0.77 (0.14, 1.40)	2.33 (1.14, 4.77)	2.33 (1.15, 4.74)

Abbreviations: ARTIS = Anti-Rheumatic Therapy in Sweden; BKK = Betriebskrankenkasse; CI = confidence interval; HIRD = HealthCore Integrated Research Database; HR = hazard ratio; IR = incidence rate; $IR_{baricitinib}$ = incidence rate for the baricitinib-treated cohort; IR_{TNFi} = incidence rate for the TNFi-treated cohort;IRD = incidence rate difference; IRR = incidence rate ratio; JMDC = JMDC, Inc's claims database; JP = Japan; MACE = major adverse cardiovascular event; MDR = Military Data Repository; PY = person years; SNDS = Système National des Données de Santé; TNFi = tumour necrosis factor inhibitor; US = United States.

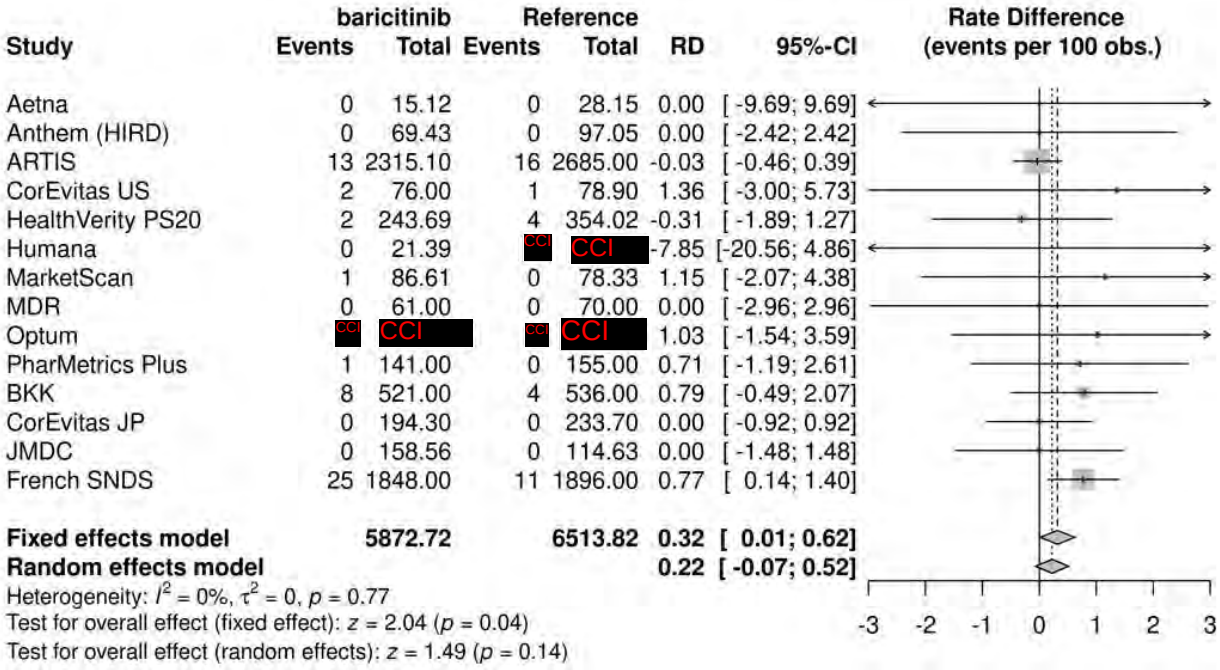
Note: Differences between confidence intervals for the individual incidence rate differences reported here and in Section 10.3 are due to different methods used for the calculation.

[a] Base or minimally adjusted HR.



Abbreviations: ARTIS = Anti-Rheumatic Therapy in Sweden; BKK = Betriebskrankenkasse; CI = confidence interval; GLMM = generalised linear mixed model; HIRD = HealthCore Integrated Research Database; IRR = incidence rate ratio; JMDC = Japan Medical Data Center, Inc.'s claims database; JP = Japan; MACE = major adverse cardiovascular event; MDR = Military Data Repository; PS20 = HealthVerity Private Source 20; SNDS = Système National des Données de Santé; TNFi = tumour necrosis factor inhibitor; US = United States.
Source: lillyce\prd\ly3009104\i4v_mc_b023\final\output\shared\irr_mace_14ds.png

Figure 10.3. Meta-analysis on incidence rate ratios for MACE comparing baricitinib and TNFi.



Abbreviations: ARTIS = Anti-Rheumatic Therapy in Sweden; BKK = Betriebskrankenkasse; CI = confidence interval; HIRD = HealthCore Integrated Research Database; JMDC = Japan Medical Data Center, Inc.'s claims database; JP = Japan; MACE = major adverse cardiovascular event; MDR = Military Data Repository; PS20 = HealthVerity Private Source 20; RD = incidence rate difference; SNDS = Système National des Données de Santé; TNFi = tumour necrosis factor inhibitor; US = United States.

Note: Any differences between confidence intervals for the individual incidence rate differences reported here and in Section 10.3 are due to different methods used for the calculation.

Source: lillyce\prd\ly3009104\i4v_mc_b023\final\output\shared\final_ird_mace3.png

Figure 10.4. Meta-analysis on incidence rate differences in MACE comparing baricitinib and TNFi.

10.1.5. Overall results serious infection

There were 318 serious infections identified, with 175 among patients treated with baricitinib. For serious infection overall IRR was numerically greater for baricitinib compared to TNFi but did not attain statistical significance (IRR=1.36; 95% CI 0.86, 2.13). The IR was greater among patients treated with baricitinib than with TNFi, the difference was 0.57 (95% CI -0.07, 1.21) per 100 PY. This IRD did not reach statistical significance. A statistically significantly increased risk of serious infection was found in the largest data source (IRR_{ARTIS} = 1.65; 95% CI 1.20, 2.26) and not in the second largest (IRR_{SNDS} = 1.04; 95% CI_{SNDS} 0.65, 1.65).