NON-INTERVENTIONAL (NI) STUDY PROTOCOL

Study information

Title	A network meta-analysis of real-world studies comparing tofacitinib with other advanced therapies in the treatment of moderate-to-severe ulcerative colitis
Protocol number	A3921447
Protocol version identifier	2.0
Date	09 December 2024
EU Post Authorization Study (PAS) register number	EUPAS108141
Active substance	L04AA29 - Tofacitinib citrate
Medicinal product	Xeljanz (tofacitinib)
Research question and objectives	*Research questions* • What is the real-world effectiveness of tofacitinib, compared to alternative advanced therapies, for the treatment of moderate-to-severe UC? • How does the safety profile of tofacitinib compare with these alternative advanced therapies? *Primary objectives* • To estimate the difference in the likelihood of achieving a clinically meaningful response, in terms of effectiveness outcomes, between patients treated with tofacitinib compared to other advanced therapies. • To estimate the relative risk of serious adverse events (AEs) between patients treated with tofacitinib versus other advanced therapies. *Secondary objective* • To estimate the incidence rate (IR) of various AEs, and of mortality, on each therapy.
Author	

1. TABLE OF CONTENTS

2. LIST OF ABBREVIATIONS

Abbreviation	Definition
AE	Adverse event
ASUC	Acute severe ulcerative colitis
CrI	Credible interval
DIC	Deviance information criterion
EC	Ethics committee
EMA	European Medicines Agency
ENCePP	European Network of Centres for Pharmacoepidemiology and Pharmacovigilance
EQ-5D	EuroQol- 5 Dimension questionnaire
GPP	Good Pharmacoepidemiology Practices
HZV	Herpes zoster virus
IBDQ	Inflammatory Bowel Disease Questionnaire
ICMJE	International Committee of Medical Journal Editors
IRB	Institutional review board
IQR	Interquartile range
ISPE	International Society for Pharmacoepidemiology
JAK	Janus Kinase
LOE	Loss of efficacy
MACE	Major adverse cardiovascular events
MCMC	Markov chain Monte Carlo
MH	Mucosal healing
NMSC	Non-melanoma skin cancer

OR	Odds ratio
PD	Pharmacodynamics
PGA	Physician Global Assessment
PK	Pharmacokinetics
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analysis
PROMIS	Patient-Reported Outcomes Measurement Information System
SAP	Statistical analysis plan
SD	Standard deviation
SF-36	36-Item Short Form Survey
SLR	Systematic literature review
TNF	Tissue necrosis factor
UC	Ulcerative colitis
VTE	Venous thromboembolism

3. RESPONSIBLE PARTIES

Principal Investigator(s) of the Protocol

Name, Degree(s)	Job Title	Affiliation	Address
[redacted]			
	Group Lead		

4. ABSTRACT

Stand Alone document, see ANNEX 1.

5. AMENDMENTS AND UPDATES

Version Identifier	Date	Amendment Type (substantial or administrative)	Protocol Section(s) Changed	Summary of Amendment(s)	Reason
1.0	09 December 2024	Substantial	6. Milestones	Update milestones to reflect updated end of data collection and study report due date	Later end of data collection date than originally anticipated due to contracting finalization, data analysis, and quality control requirements for the PASS study

6. MILESTONES

Milestone	Planned Date
Start of data collection	17 January 2024
End of data collection	30 April 2024
Registration in the EU PAS register	02 January 2024
Final study report	25 March 2025

7. RATIONALE AND BACKGROUND

Ulcerative colitis (UC) is a chronic idiopathic inflammatory bowel disease of the colon that causes continuous mucosal inflammation starting in the rectum and extending to the more proximal colon, with variable extents. Typical symptoms include bloody diarrhea, abdominal pain, urgency, and tenesmus.[1] Moderate-to-severe UC is clinically defined as 4 to 6 bowel movements daily with moderate to severe rectal bleeding in the absence of constitutional signs or symptoms. Endoscopically, it is defined as marked mucosal erythema, absent vascularization, friability, granularity, spontaneous bleeding, and ulcerations.[2] Patients with moderate-to-severe UC experience remarkable disease burden with frequent flares and hospitalizations, which are associated with a significant economic burden.[3]

The primary therapeutic goal in UC is to induce and maintain long-term disease remission.[4] However, there is no single treatment pathway for patients. Several advanced therapies have become available for the induction and maintenance of remission in moderate-severe UC, including tissue necrosis factor (TNF)-α inhibitors (i.e., infliximab, adalimumab, and golimumab), interleukin inhibitor (i.e., ustekinumab), integrin receptor inhibitor (i.e., vedolizumab), and small-molecule Janus Kinase (JAK) inhibitor (i.e., tofacitinib).[2] Sphingosine 1-phosphate receptor modulator ozanimod, JAK inhibitors filgotinib and upadacitinib are recently approved drug for the treatment of moderate-to-severe UC.[5-7]

Tofacitinib (Xeljanz®) is an oral JAK inhibitor for the treatment of moderate-to-severely active UC. Evidence from clinical trial indicates that treatment with tofacitinib was more efficacious compared to placebo for induction of remission and mucosal healing (MH). Also, the maintenance therapy with tofacitinib was more effective than placebo in sustaining remission and MH.[8] A meta-analysis of real-world studies demonstrated the effectiveness of tofacitinib in a highly refractory population of patients with moderate-to-severe UC. Tofacitinib was also shown to have an acceptable safety profile.[9]

Several network meta-analyses (NMA) have been published comparing efficacy and safety of biologics and small molecules for the treatment of moderate-to-severe UC.[10-12] However, all these NMAs were conducted using data from the randomized trials. Previous meta-analyses examining the efficacy and safety of vedolizumab[13] based on randomized trials and of ustekinumab[14,15] based on real-world evidence have been published. Despite the compelling evidence on efficacy and safety of tofacitinib from clinical and real-world studies, there is lack of evidence about the comparative effectiveness and safety of tofacitinib with other therapies approved for the treatment of moderate-to-severely active UC from real-world studies.

The purpose of the study is to assess the feasibility and conduct a NMA to compare the real-world effectiveness and safety of tofacitinib with other advanced therapies in the treatment of moderate-to-severe UC. This noninterventional study is designated as a PASS and is conducted voluntarily by Pfizer.

8. RESEARCH QUESTION AND OBJECTIVES

8.1. Research Question

Research questions to be addressed by this study are as follows:

1. What is the real-world effectiveness of tofacitinib, compared to alternative advanced therapies, for the treatment of moderate-to-severe UC?

2. How does the safety profile of tofacitinib compare to these alternative advanced therapies?

8.2. Study Objectives

8.2.1. Primary Objectives

The primary objectives for this study are:

1. To estimate the difference in the likelihood of achieving a clinically meaningful response, in terms of effectiveness outcomes, between patients treated with tofacitinib compared to other advanced therapies.
2. To estimate the relative risk of serious adverse events (AEs) between patients treated with tofacitinib versus other advanced therapies.

8.2.2. Secondary Objectives

The secondary objectives for this study are:

1. To estimate the incidence rate (IR) of various AEs, and of mortality, on each therapy.

9. RESEARCH METHODS

9.1. Study Design

The study is designed as a NMA with the primary objective to compare the effectiveness of tofacitinib with other advanced therapies in real-world studies for the treatment of patients with moderate-to-severe UC. The secondary objective of the study is to compare the safety outcomes as IR assessed through a meta-analysis of tofacitinib and other advanced therapies in real-world studies of patients with moderate-to-severe UC. These analyses will be performed on data collected from studies published in literature in the form of a systematic literature review (SLR) and no patient enrollment will be done.

9.2. Setting

A SLR was conducted as per the standards published by the Cochrane Collaboration[16] and in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines[17] to identify the real-world studies reporting effectiveness and/or safety outcomes of advanced therapies for moderate-to-severe UC.[18,19] A comprehensive literature search was performed using the Embase® and MEDLINE® databases through the Embase.com platform

from 01 January 2005 to 30 April 2023. The MEDLINE® Epub ahead of print, in-process, and other nonindexed citations were searched on PubMed (01 May 2023). The selection of the time frame for searches starting January 2005 was based on the earliest approval of infliximab for ulcerative colitis in September 2005. Only products that were approved at the time of search were included.

9.2.1. Inclusion Criteria

Publications must meet all of the following inclusion criteria to be eligible for inclusion in the SLR:

Table 1. Inclusion and exclusion criteria for SLR

Inclusion criteria	Exclusion criteria
Patient population • Adult patients (≥18 years) with moderate-to-severe ulcerative colitis	• Patients with disease other than UC • UC studies with pediatric population • Studies of patients with ASUC
Intervention • Tofacitinib	• Any pharmacological intervention other than reported in the included list
Comparators • Adalimumab • Filgotinib • Golimumab • Infliximab • Ozanimod • Upadacitinib • Ustekinumab • Vedolizumab	• Any pharmacological/non-pharmacological treatment other than reported in the included list of comparators
Outcomes **Efficacy/effectiveness:** • Response (all definitions, including sustained response) • Remission (all definitions, i.e., sustained remission, steroid-free remission) • MH/endoscopic improvement (all definitions, including sustained MH) • Histological changes/remission • Relapse or loss of response/remission • Treatment duration • Mayo score / Disease activity index (including changes in Mayo score from baseline) • Fecal Calprotectin • C-reactive protein **Safety:** • Rates of surgical intervention • Time to surgical intervention • Hospitalization • Mortality • Serious infection • Herpes zoster • Venous thromboembolism • Malignancies (including NMSC)	• Other outcomes

Inclusion criteria	Exclusion criteria
• AEs • Serious AEs • Anemia • Fatigue • Headache • Nausea • Nasopharyngitis • Pyrexia • Worsening ulcerative colitis • Discontinuation (any reason, AE, LOE) • Opportunistic infections • Major adverse cardiovascular events • UC-related surgery/colectomy • UC-related hospitalization **Patient-reported outcomes:** • EQ-5D • SF-36 • IBDQ • PROMIS • PGA	
Study designs • Real-world studies, including observational studies such as cohort study/follow-up study, longitudinal study, cross-sectional study, prospective study, retrospective study, case-control study, population-based study, registry, survey	• Randomized controlled trials • Interventional clinical studies • Systematic / narrative reviews • Case reports • Editorial / Opinions / Commentary / Letters
Species • Humans	• Animal studies • In-vitro / In-vivo studies • PK/PD studies
Language • English	• Studied published in language other than English
Country • No restriction	• Not applicable
Sample size • Studies with 30 or more patients	• Studies with <30 patients
Publication type • Peer reviewed full-text journal articles • Conference abstracts	• Not applicable
Search timeframe • January 01, 2005, to till April 30, 2023	• Full-text studies published before 2005 • Conference abstracts published before 2019 (of last 4 years)

AE, adverse event; EQ-5D, EuroQol- 5 Dimension questionnaire; IBDQ, Inflammatory Bowel Disease Questionnaire; NMSC, non-melanoma skin cancer; PD, Pharmacodynamics; PGA, Physician Global Assessment; PK, Pharmacokinetics; PROMIS, Patient-Reported Outcomes Measurement Information System; SF-36, 36-Item Short Form Survey; UC, ulcerative colitis.

9.2.2. Exclusion Criteria

Publications meeting any of the criteria listed in **Table 1** in Section 9.2.1 will not be included in the SLR.

9.3. Variables

Data on the following baseline characteristics will be used to perform the covariate adjustment, and to measure the similarity of single-arm studies for matching.

- Age (mean, SD)
- Male (n, %)
- Smoking status (current, former, never; n, %)
- Disease duration (median, IQR)
- Extent of disease (E1, E2, E3; n, %)
- Disease severity (moderate-to-severe; n, %)
- C-reactive protein (mg/L; median, IQR)
- Extraintestinal manifestations (n, %)
- Previous steroid use (n, %)
- Previous immunosuppressive agent use (n, %)
- Previous biologic exposure (n, %)
- Previous anti-TNF exposure (n, %)
- Concomitant steroid use (n, %)
- Concomitant immunosuppressive agent use (n, %)

Given that data are not available from each arm in every study, multiple imputation via predictive mean matching will be used to address the missing data in covariates as outlined in Section 9.8.1. Owing to the availability of aggregate data, the adjustment of the covariates will be restricted at the study-level.

9.3.1. Outcomes

The clinical effectiveness outcomes for NMA in this study are:

- Clinical remission
- Clinical response
- Steroids-free remission

The clinical safety outcomes for NMA in this study are:

- Serious AE

The safety AEs for IR assessment in this study are:

- Serious infection
- Herpes zoster virus (HZV) infection
- Venous thromboembolism (VTE)
- Major adverse cardiovascular events (MACE)
- Malignancies
- Mortality

9.4. Data Sources

For the SLR, the key electronic biomedical literature databases (Medical Literature Analysis and Retrieval System Online [MEDLINE®] and Excerpta Medica Database [Embase®]) were searched to identify the published evidence on the effectiveness and/or safety of advanced therapies in moderate-to-severe UC. MEDLINE In-Process was searched to ensure that non-indexed citations were retrieved. The time frame of searches was from 01 January 2005 30 April 2023 for full publications, and from 01 January 2019 to 30 April 2023 for the conference abstracts. Embase® was searched using the embase.com interface, whereas MEDLINE® In-Process and other non-indexed citations was searched using the Pubmed.com interface.

Following the standard process of SLR as described in Cochrane Collaboration, a total of 246 distinct studies have been included that will provide the data to conduct NMA for effectiveness outcomes and meta-analysis for safety events.

9.5. Study Size

In this retrospective assessment, as the data for the present study have been collected in the form of a SLR, there are no priori hypotheses to test and sample size calculations are not applicable.

The primary data source for this study are 246 individual studies that have been included in the SLR. Of these, 48 are comparative treatment studies and 198 are single-treatment studies. In terms of population in the included studies, 26 studies have patients previously treated with biologics, 52 studies with biologics-naïve patients, 167 studies with mixed populations, and one study with biologics exposure unclear. Only aggregated study-level data is available. The individual participant-level data is not available for any of the included studies.

The distribution of these studies by study design and treatments is provided below.

Table 2. Distribution of studies by treatments

Study type	Details	No. of studies
Comparative	Tofacitinib and vedolizumab	5
	Tofacitinib and ustekinumab	4
	Tofacitinib and upadacitinib	1
	Tofacitinib and anti-TNF agents	1
	Vedolizumab and anti-TNF agents	20
	Vedolizumab and ustekinumab	1
	Vedolizumab, ustekinumab and anti-TNF agents	1
	Comparing anti-TNF agents	15

Single treatment	Tofacitinib	38
	Vedolizumab	60
	Ustekinumab	14
	Upadacitinib	1
	Anti-TNF agents	85

9.6. Data Management

All study data exist as structured data by the time of study. Analyses will be conducted using *R* statistical software. Versions of packages will be documented to assure reproducibility.

9.7. Data Analysis

An overview of the planned analyses is provided below. As the planned analysis is described with sufficient details, no separate statistical analysis plan (SAP) will be developed for this study.

In the following, two approaches have been planned to performing the NMA, (i) contrast-based models which perform the synthesis of data on relative treatment effects between study arms, and (ii) arm-based models which perform the synthesis of data on absolute effects across study arms. Both approaches can be applied to estimate an overall pooled relative effect. There has been much discussion in the literature on which approach makes more efficient use of data,[20,21] thus both approaches will be implemented to identify any sensitivity in the results, especially in the context of incorporating single-arm studies into the synthesis.

9.7.1. Primary Analysis

The following models assume a binomial likelihood for the number of responders, r_{ik}, on treatment *k* in study *i*,

$$r_{ik} \sim Bin(p_{ik}, n_{ik})$$

with number of participants, n_{ik}, and response probability p_{ik}.

The pooled relative treatment effects will be estimated on the odds ratio (OR) scale.

9.7.1.1. Comparative Studies Only

The following models will be used for comparative studies.

- **Model 1:** the standard contrast-based NMA[22,23]

$$logit(p_{ik}) = \mu_i + \delta_{i,bk} I_{k \neq b}$$
$$\delta_{i,bk} \sim N(d_{bk}, \sigma^2)$$
$$d_{bk} = d_{1k} - d_{1b}$$
$$d_{1k} \sim N(0, 100^2)$$
$$\sigma \sim U(0,2)$$

$logit(.)$ – link function to perform synthesis on the log OR scale
μ_i – response on the baseline treatment in study *i*
$\delta_{i,bk}$ – log OR comparing treatment *k* with the baseline treatment in study *i*
$I_{k\neq 1}$ – indicator variable that is equal to 1 if $k \neq 1$, and is equal to 0 if $k = 1$
d_{bk} – pooled relative effect comparing treatment *k* with the baseline treatment
σ^2 – common between-studies heterogeneity parameter across all treatment contrasts
d_{1k} – pooled relative effect comparing treatment *k* with the reference treatment
d_{1b} – pooled relative effect comparing the baseline treatment with the reference treatment

A random-effects model will be implemented as a base case (due to expectation of between-studies heterogeneity). A fixed-effects (i.e., $\sigma^2 = 0$) model will be implemented as a sensitivity analysis. The fit of both models will be reported in terms of the deviance information criterion (DIC), and residual deviance, values.

- **Model 2:** Model 1 + covariate adjustment on baseline response[24]

$$logit(p_{ik}) = \mu_i + \sum_{m=1}^{M} \alpha_i^m . x_{ik}^m + \delta_{i,bk} I_{k\neq b}$$
$$\alpha_i^m \sim N(a^m, \sigma_a^2)$$
$$a^m \sim N(0, 100^2)$$
$$\sigma_a^2 \sim U(0,2)$$

x_{ik}^m – mean value of covariate *m*, in treatment arm *k* of study *i*
α_i^m – association between the baseline response and the mean value of covariate *m*

- **Model 3:** Model 2 + covariate adjustment on treatment effects[24]

$$logit(p_{ik}) = \mu_i + \sum_{m=1}^{M} \alpha_i^m . x_{ik}^m + (\delta_{i,bk} + \sum_{m=1}^{M} \beta_{i,bk}^m . x_{ik}^m) I_{k\neq b}$$
$$\beta_{i,bk}^m \sim N(B_{bk}^m, \sigma_B^2)$$
$$B_{bk}^m = B_{1k}^m - B_{1b}^m$$
$$B_{1k}^m \sim N(0, 100^2)$$

$\beta_{i,bk}^m$ – treatment-covariate interaction, quantifying the association between covariate *m* and the relative effect comparing treatment *k* with the baseline treatment in study *i*

A fixed covariate effects model (i.e., setting $\sigma_B^2 = 0$) will also be implemented as a sensitivity analysis.

9.7.1.2. Comparative and single-arm studies

Base case

- Matching of single-arm studies using similarity measure[25]

$$\Delta_{total}[i_1, i_2] = \frac{\sum_{m=1}^{M} w_m \Delta_m[i_1, i_2]}{\sum_{m=1}^{M} w_m}$$

$\Delta_{total}[i_1, i_2]$ – the total distance between single-arm studies i_1 and i_2 in terms of mean covariate values (i.e., baseline characteristics)
w_m – the weight associated with covariate *m*

$\Delta_m[i_1, i_2]$ – the difference between single-arm studies i_1 and i_2 in covariate *m*
M – total number of covariates

- Models 1, 2, and 3

Sensitivity analysis

- **Model 4:** Arm-based NMA[26,27]

$$logit(p_{ik}) = \mu_k + v_{ik}$$

$$(v_{i1}, v_{i2}, \ldots, v_{iK})^T \sim MVN(\mathbf{0}, \mathbf{\Sigma}_K)$$

$$E(p_k) \approx expit\left(\mu_k / \sqrt{1 + C^2\sigma_k^2}\right)$$

$$C = 16\sqrt{3}/(15\pi)$$

$$OR_{kl} = \frac{p_k/(1-p_k)}{p_l/(1-p_l)}$$

$logit(.)$ – link function to perform synthesis on log odds scale
$expit(.)$ – back transformation from the log odds scale to the probability scale
μ_k – mean absolute effect on treatment *k*
v_{ik} – random effect for treatment *k* in study *i*
$(v_{i1}, v_{i2}, \ldots, v_{iK})^T$ – vector of random effects for treatments in study *i*
$\mathbf{\Sigma}_K$ – variance-covariance matrix quantifying between-studies heterogeneity, and within-study correlation, for all *K* treatments
p_k – population-averaged absolute probability of an event on treatment *k*
σ_k^2 – *k*th diagonal element of Σ_K, representing between-studies heterogeneity associated with treatment *k*
OR_{kl} – odds ratio comparing treatments *k* and *l*

- **Model 5:** Model 4 + covariate adjustment[28]

$$g(p_{ik}) = \mu_k + v_{ik} + \sum_{m=1}^{M} \beta_{ik}^{m} . x_{ik}^{m}$$
$$(\beta_{i1}^{m}, \dots, \beta_{iK}^{m})^T \sim N(\boldsymbol{B^m}, \boldsymbol{\Sigma}_x)$$

β_{ik}^{m} – interaction between the absolute treatment effect and the covariate *m*
x_{ik}^{m} – mean value for covariate *m*, on treatment *k* in study *i*
$\boldsymbol{B^m} = (B_1^m, \dots, B_K^m)^T$ – vector of mean interaction effects
$\boldsymbol{\Sigma_x}$ – variance-covariance matrix quantifying between-studies heterogeneity, and correlation, across interaction effects

9.7.1.3. Comparative and single-arm studies with weighting

In the following analysis, a contrast-based NMA is performed as a base case (Model 7), whilst an arm-based approach is implemented as a sensitivity analysis (Model 7). The models will be fit to the full data set consisting of comparative studies and matched single-arm studies (i.e., after applying the study matching approach described in 9.7.1.2).

Base case

- **Model 6:** Model 3 + power prior[29]

$$L^{SATs}(p_{ik}) = \left(p_{ik}^{r_{ik}}(1-p_{ik})^{n_{ik}-r_{ik}}\right)^{\gamma}$$

$L^{SATs}(p_{ik})$ – the likelihood (data) corresponding to matched single-arm studies
γ – power parameter, taking values between 0 and 1

The power parameter values will be varied in increments of 0.2 between 0 and 1 to illustrate the change in results as the data from the single-arm studies are down-weighted relative to the comparative studies, where 0.2 represents a large down-weighting and 1 gives no down-weighting (both sets of evidence contribute equally to the overall pooled estimates).

Sensitivity analysis

- **Model 7:** Model 5 + power prior[28]

$$L^{SATs}(p_{ik}) = \left(p_{ik}^{r_{ik}}(1-p_{ik})^{n_{ik}-r_{ik}}\right)^{\gamma}$$

9.7.1.3.1. Assessment of model fit

Within each data set (i.e., comparative studies only, comparative and single-arm studies, and comparative and single-arm studies with weighting), models will be compared based on the residual deviance and DIC values. Smaller DIC values suggest better model fit, with a difference in DIC of two, or more, units considered meaningful.[30]

NMA results will be reported as basic parameters (i.e., quantifying the effect of each comparator therapy versus one another), on the OR scale, in terms of mean and 95% credible interval (CrI) estimates. A forest plot will be used to illustrate the treatment effect estimates (mean and 95% CrIs) of tofacitinib vs. comparator therapies. Treatment rankings will also be reported based on the probability of each therapy being the most effective treatment.

9.7.2. Secondary analysis

The IRs will be estimated for each of the safety events as described above. A meta-analysis will be performed to pool IR estimates (calculated using data on the numbers of events and person-years per study) across studies for each outcome, in terms of mean and 95% CI estimates for each therapy. This meta-analysis will be implemented using the *metarate* function in the *meta R* package.[31]

The likelihood model[22] assumes data are available on the number of events, r_{ik}, in arm *k* of study *i* during the follow-up period, and on the exposure time in person-months at risk, E_{ik}. Then, the number of events is assumed to follow a Poisson distribution,

$$r_{ik} \sim Poisson(\lambda_{ik} E_{ik})$$

where λ_{ik} represents the rate at which events occur in arm *k* of study *i*.

The synthesis model[22] is then given by,

$$\begin{gathered} log(\lambda_{ik}) = \mu_i + \delta_{i,bk} I_{k \neq b} \\ \delta_{i,bk} \sim N(d_{bk}, \sigma^2) \\ d_{bk} = d_{1k} - d_{1b} \\ d_{1k} \sim N(0, 100^2) \\ \sigma \sim U(0,2) \end{gathered}$$

9.7.3. Software implementation

All models will be implemented under a Bayesian framework, using Markov chain Monte Carlo (MCMC) simulation to estimate posterior distributions for model parameters. Each implementation will consist of three MCMC chains, for which the effective sample size and $\hat{R}$ statistics will be used to assess non-convergence, and posterior estimates will be checked for sensitivity to initial values. Contrast-based models (Models 1-3, 6) will be fit using the *R2OpenBUGS* package in the *R* software.[32] The arm-based models (Models 4-5, 7) will be fit using the *pcnetmeta* package in the *R* software.[26]

9.8. Quality Control

Analyses are programmed according to the specifications in the protocol, and if applicable, the SAP, and documented in a programming plan. Final deliverables will be reviewed and verified by a second, independent programmer who may also perform double programming. All quality checks will be documented in the programming plan.

9.8.1. Methods to Address Missing Data

The missing values in the covariates listed above will be accounted for by multiple imputation using predictive mean matching, implemented via the *mice* package[33] in the *R* software.

9.9. Limitations of the Research Methods

The potential confounding caused by single-arm studies in the traditional arm-based method could be one of the limitations of the present analysis. However, the current analysis plan appears to have already incorporated measures to overcome this, such as the use of covariate adjustment and power priors from single-arm studies. Nevertheless, there is the possibility of residual bias due to unmeasured confounders.

In terms of data, certain outcomes might have a relatively small number of studies, while other outcomes might have more studies. In the former case, the parameter estimation might be unstable. A potential plan is to use multivariate models to jointly analyze multiple outcomes (in addition to the joint analysis of multiple treatments). Of note, the data are available as aggregate study-level data (not individual participant data) which are not necessarily free of bias despite the covariate adjustment. In addition, publication bias might be another problem; the comparison-adjusted funnel plots to be used to assess publication bias and validate the results.

Regarding the analytic methods, a potential problem is the lack of convergence of the MCMC algorithm for the Bayesian analyses. This is likely associated with limited data, improper initial values for the MCMC, and inappropriate prior distributions. However, this problem can be overcome by using initial values that are clinically sensible and various choices of priors, including informative priors based on external data.

9.10. Other Aspects

Not applicable.

10. PROTECTION OF HUMAN PARTICIPANTS

10.1. Patient Information

This study involves data that exist in deidentified/anonymized structured format and contain no patient personal information.

10.2. Patient Consent

As this study involves deidentified/anonymized structured data, which according to applicable legal requirements do not contain data subject to privacy laws, obtaining informed consent from patients by Pfizer is not required.

10.3. Institutional Review Board (IRB)/ Ethics Committee (EC)

As this study involves only the use of de-identified data, and is conducted internal to Pfizer, no IRB/IEC approval is necessary.

10.4. Ethical Conduct of the Study

The study will be conducted in accordance with legal and regulatory requirements, as well as with scientific purpose, value and rigor and follow generally accepted research practices described in Guidelines for Good Pharmacoepidemiology Practices (GPP) issued by the International Society for Pharmacoepidemiology (ISPE),[34] European Medicines Agency (EMA) European Network of

Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) Guide on Methodological Standards in Pharmacoepidemiology.[35]

11. MANAGEMENT AND REPORTING OF ADVERSE EVENTS/ADVERSE REACTIONS

This study involves data that exist as structured data by the time of study start.

In these data sources, individual patient data are not retrieved or validated, and it is not possible to link (i.e., identify a potential association between) a particular product and medical event for any individual. Thus, the minimum criteria for reporting an adverse event (i.e., identifiable patient, identifiable reporter, a suspect product, and event) cannot be met.

12. PLANS FOR DISSEMINATING AND COMMUNICATING STUDY RESULTS

One or more abstracts may be developed and submitted to relevant scientific conference(s) and one manuscript will be developed and submitted to relevant peer-reviewed medical journals. Authorship will follow the guidelines proposed by the International Committee of Medical Journal Editors (ICMJE; www.icmje.org). All authors should meet the criteria for authorship, and all people who meet the criteria should be authors. Any potential conflicts of interest will be disclosed.

In the event of any prohibition or restriction imposed (e.g., clinical hold) by an applicable competent authority in any area of the world, or if the party responsible for collecting data from the participant is aware of any new information which might influence the evaluation of the benefits and risks of a Pfizer product, Pfizer should be informed immediately.

13. REFERENCES

1. Feuerstein JD, Cheifetz AS. Ulcerative colitis: epidemiology, diagnosis, and management. Mayo Clinic Proceedings; 2014: Elsevier; 2014. p. 1553-63.

2. Kayal M, Shah S. Ulcerative colitis: current and emerging treatment strategies. *Journal of clinical medicine* 2019; **9**(1): 94.

3. Crohon's & Colitis Foundation of America. The Facts About Inflammatory Bowel Diseases. 2014. Updated IBD Factbook.pdf (crohnscolitisfoundation.org). (accessed May 4 2023).

4. Rubin DT, Ananthakrishnan AN, Siegel CA, Sauer BG, Long MD. ACG clinical guideline: ulcerative colitis in adults. *Official journal of the American College of Gastroenterology| ACG* 2019; **114**(3): 384-413.

5. US FDA. Zeposia® (ozanimod). Highlights of prescribing information. , 2021.

6. EMA. Jyseleca® (filgotinib). Prescribing information. 2021.

7. US FDA. Rinvoq® (upadacitinib). Highlights of prescribing information. 2022.

8. Sandborn WJ, Su C, Sands BE, et al. Tofacitinib as induction and maintenance therapy for ulcerative colitis. *New England Journal of Medicine* 2017; **376**(18): 1723-36.

9. Taxonera C, Olivares D, Alba C. Real-world effectiveness and safety of tofacitinib in patients with ulcerative colitis: systematic review with meta-analysis. *Inflammatory Bowel Diseases* 2022; **28**(1): 32-40.

10. Lasa JS, Olivera PA, Danese S, Peyrin-Biroulet L. Efficacy and safety of biologics and small molecule drugs for patients with moderate-to-severe ulcerative colitis: a systematic review and network meta-analysis. *The Lancet Gastroenterology & Hepatology* 2022; **7**(2): 161-70.

11. Burr NE, Gracie DJ, Black CJ, Ford AC. Efficacy of biological therapies and small molecules in moderate to severe ulcerative colitis: systematic review and network meta-analysis. *Gut* 2022; **71**(10): 1976-87.

12. Lu X, Jarrett J, Sadler S, Tan M, Dennis J, Jairath V. Comparative efficacy of advanced treatments in biologic-naïve or biologic-experienced patients with ulcerative colitis: a systematic review and network meta-analysis. *International Journal of Clinical Pharmacy* 2023; **45**(2): 330-41.

13. Jin Y, Lin Y, Lin L-J, Zheng C-Q. Meta-analysis of the effectiveness and safety of vedolizumab for ulcerative colitis. *World Journal of Gastroenterology: WJG* 2015; **21**(20): 6352.

14. Uchida G, Nakamura M, Yamamura T, Tsuzuki T, Kawashima H. Real-world effectiveness of ustekinumab for patients with ulcerative colitis: a systematic review and meta-analysis. *Nagoya Journal of Medical Science* 2023; **85**(3): 402.

15. Taxonera C, Olivares D, López-García ON, Alba C. Meta-analysis: Real-world effectiveness and safety of ustekinumab in patients with ulcerative colitis. *Alimentary Pharmacology & Therapeutics* 2023; **57**(6): 610-9.

16. Cumpston M, Li T, Page MJ, et al. Updated guidance for trusted systematic reviews: a new edition of the Cochrane Handbook for Systematic Reviews of Interventions. *The Cochrane database of systematic reviews* 2019; **2019**(10).

17. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *International journal of surgery* 2021; **88**: 105906.

18. Sterne JA, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *bmj* 2019; **366**.

19. Shea BJ, Reeves BC, Wells G, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *bmj* 2017; **358**.

20. Dias S, Ades A. Absolute or relative effects? Arm-based synthesis of trial data. *Research synthesis methods* 2016; **7**(1): 23.

21. Hong H, Chu H, Zhang J, Carlin BP. Rejoinder to the discussion of "a Bayesian missing data framework for generalized multiple outcome mixed treatment comparisons," by S. Dias and AE Ades. *Research synthesis methods* 2016; **7**(1): 29.

22. Dias S, Welton NJ, Sutton AJ, Ades A. NICE DSU technical support document 2: a generalised linear modelling framework for pairwise and network meta-analysis of randomised controlled trials. 2011.

23. Dias S, Ades AE, Welton NJ, Jansen JP, Sutton AJ. Network meta-analysis for decision-making: John Wiley & Sons; 2018.

24. Thom HH, Capkun G, Cerulli A, Nixon RM, Howard LS. Network meta-analysis combining individual patient and aggregate data from a mixture of study designs with an application to pulmonary arterial hypertension. *BMC medical research methodology* 2015; **15**(1): 1-16.

25. Schmitz S, Maguire Á, Morris J, et al. The use of single armed observational data to closing the gap in otherwise disconnected evidence networks: a network meta-analysis in multiple myeloma. *BMC medical research methodology* 2018; **18**: 1-18.

26. Lin L, Zhang J, Hodges JS, Chu H. Performing arm-based network meta-analysis in R with the pcnetmeta package. *Journal of statistical software* 2017; **80**.

27. Chu H, Nie L, Chen Y, Huang Y, Sun W. Bivariate random effects models for meta-analysis of comparative studies with binary outcomes: methods for the absolute risk difference and relative risk. *Statistical methods in medical research* 2012; **21**(6): 621-33.

28. Hong H, Fu H, Carlin BP. Power and commensurate priors for synthesizing aggregate and individual patient level data in network meta-analysis. *Journal of the Royal Statistical Society Series C: Applied Statistics* 2018; **67**(4): 1047-69.

29. Ibrahim JG, Chen MH, Gwon Y, Chen F. The power prior: theory and applications. *Statistics in medicine* 2015; **34**(28): 3724-49.

30. Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linde A. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society Series B: Statistical Methodology* 2002; **64**(4): 583-639.

31. Schwarzer G. meta: An R package for meta-analysis. *R news* 2007; **7**(3): 40-5.

32. Sturtz S, Ligges U, Gelman A. R2WinBUGS: a package for running WinBUGS from R. *Journal of Statistical software* 2005; **12**: 1-16.

33. Van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate imputation by chained equations in R. *Journal of statistical software* 2011; **45**: 1-67.

34. International Society of Pharmacoepidemiology. Guidelines for good pharmacoepidemiology practice (GPP). *Pharmacoepidemiology and drug safety* 2016; **25**(1): 2-10.

35. European Medicines Agency (EMA). European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) Guide on Methodological Standards in Pharmacoepidemiology 2015.

14. LIST OF TABLES

15. LIST OF FIGURES

None

ANNEX 1. LIST OF STANDALONE DOCUMENTS

Number	Document reference number	Date	Title
1	Section 4	21 December 2023	Abstract

ANNEX 2. ENCEPP CHECKLIST FOR STUDY PROTOCOLS

Not required.

ANNEX 3. ADDITIONAL INFORMATION

Not applicable.