STUDY REPORT

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
Version identifier of the study report	1.0
Date	27 January 2025
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.
Country(-ies) of study	Global
Author	[redacted]

TABLE OF CONTENTS

LIST OF IN-TEXT TABLES AND FIGURES

Annex 1. List of stand-alone documents

Appendix 1. SIGNATURES

Not applicable

Appendix 2.1 PROTOCOL

Not applicable

Appendix 2.2 Protocol administrative change letter (PACL)

Not applicable

Appendix 3. INVESTIGATORS AND CORRESPONDING INDEPENDENT ETHICS COMMITTEES (IECs) OR INSTITUTIONAL REVIEW BOARDS (IRBs)

Not applicable

Appendix 3.1. List of Investigators by Country

Not applicable

Appendix 3.2. List of Independent Ethics Committee (IEC) or Institutional Review Board (IRB) and Corresponding Protocol Approval Dates

Not applicable

Appendix 4. STATISTICAL ANALYSIS PLAN

Not applicable

Appendix 5. SAMPLE CASE REPORT FORM (CRF) / DATA COLLECTION TOOL (DCT)

Not applicable

Appendix 6. SAMPLE STANDARD SUBJECT INFORMATION SHEET AND INFORMED CONSENT DOCUMENT (ICD)

Not applicable

Appendix 7. LIST OF SUBJECT DATA LISTINGS

Not applicable

Appendix 7.1 Withdrawn Subjects

Not applicable

Appendix 7.2 Protocol Deviations

Not applicable

Appendix 7.3 Subjects Excluded from the Analysis

Not applicable

Appendix 7.4 Demographic Data

Not applicable

Appendix 7.5 Medication/Treatment Data

Not applicable

Appendix 7.6 Endpoint Data

Not applicable

Appendix 7.7 Adverse Events

Not applicable

Appendix 7.8 Laboratory listings

Not applicable

Appendix 8. ADDITIONAL DOCUMENTS

Not applicable

1. ABSTRACT (STAND-ALONE DOCUMENT)

2. LIST OF ABBREVIATIONS

Abbreviation	Definition
AE	Adverse event
ASUC	Acute severe ulcerative colitis
AT	Advanced treatment
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
ET	Exposure time
GPP	Good Pharmacoepidemiology Practices
HZV	Herpes zoster virus
IBDQ	Inflammatory Bowel Disease Questionnaire
ICMJE	International Committee of Medical Journal Editors
IR	Incidence rate
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
PY	Person-years
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. INVESTIGATORS

The names, affiliations, and contact information of the investigators at each study site are listed in Appendix 3.1.

Principal Investigator(s) of the Protocol

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

4. OTHER RESPONSIBLE PARTIES

Not applicable

5. MILESTONES

Milestone	Planned Date	Actual Date
Start of data collection	17 January 2024	16 Jan 2024
End of data collection	31 March 2024	30 Apr 2024
Registration in the EU PAS register	02 January 2024	15 Jan 2024
Final study report	15 March 2024	27 January 2025

6. 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 non-interventional study was designated as a Post-Authorization Safety Study (PASS) and was conducted voluntarily by Pfizer.

7. RESEARCH QUESTION AND OBJECTIVES

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?

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.

The secondary objectives for this study are:

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

8. AMENDMENTS AND UPDATES

None.

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.

9.3. Subjects

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 **Intervention** Tofacitinib **Comparators** • Adalimumab • Filgotinib • Golimumab • Infliximab • Ozanimod • Upadacitinib • Ustekinumab • Vedolizumab	• Patients with disease other than UC • UC studies with pediatric population Studies of patients with ASUC Any pharmacological intervention other than reported in the included list Any pharmacological/non-pharmacological treatment other than reported in the included list of comparators

Table 1. Inclusion and exclusion criteria for SLR

Inclusion criteria	Exclusion criteria
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) • 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	Other outcomes

Table 1. Inclusion and exclusion criteria for SLR

Inclusion criteria	Exclusion criteria
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.

Publications meeting any of the criteria listed in **Table 1** in Section 9.3 will not be included in the SLR.

9.4. 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-TNFα 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.

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

9.5. Data sources and measurement

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.

9.6. Bias

This was a NMA of published studies. There is potential bias of studies or findings not published, or those not published in English or outside of the scope of the time frame specified. MEDLINE In-Process was searched to ensure that non-indexed citations were retrieved.

9.7. 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-TNFα agents	1

Table 2. Distribution of studies by treatments

Study type	Details	No. of studies
	Vedolizumab and anti-TNFα agents	20
	Vedolizumab and ustekinumab	1
	Vedolizumab, ustekinumab and anti-TNFα agents	1
	Comparing anti-TNFα agents	15
Single treatment	Tofacitinib	38
	Vedolizumab	60
	Ustekinumab	14
	Upadacitinib	1
	Anti-TNFα agents	85

9.8. Data transformation

Detailed methodology for data transformations, particularly complex transformations (eg, many raw variables used to derive an analytic variable), are documented in the statistical analysis plan (SAP), which is dated, filed and maintained by the sponsor (Appendix 4).

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.

9.9. Statistical methods

9.9.1. Main summary measures

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.

9.9.2. Main statistical methods

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.

Comparative Studies: 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.

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

9.9.3. Missing values

The missing values in the covariates listed above will be accounted for by multiple imputation using predictive mean matching, implemented via the *mice package[26]* in the *R* software.

9.9.4. Sensitivity analyses

Model 4: Arm-based NMA[27, 28]

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

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

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

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

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

$logit(.)$ – 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*
$(v_{i1}, v_{i2}, \dots, v_{iK})^T$ – 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*
p_l – population-averaged absolute probability of an event on treatment *l*
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[29]

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

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
$\mathbf{\Sigma}_x$ – variance-covariance matrix quantifying between-studies heterogeneity, and correlation, across interaction effects

Comparative and single-arm studies with weighing

In the following analysis, a contrast-based NMA is performed as a base case (Model 6), 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[30]

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

$L^{SATs}(p_{ik})$ – 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[29]

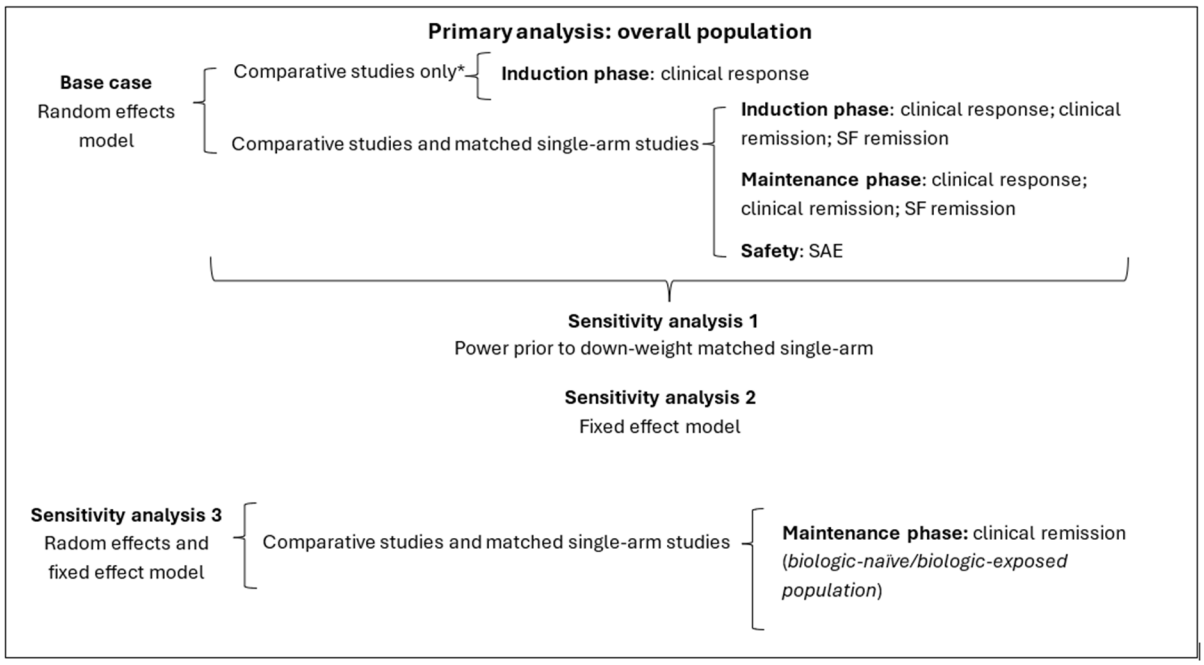
$$L^{SATs}(p_{ik}) = \left(p_{ik}^{r_{ik}}(1-p_{ik})^{n_{ik}-r_{ik}}\right)^{\gamma}$$

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.[31]

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.

Figure 1. Flow diagram of the primary analyses conducted



Secondary analyses

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[32]

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,

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

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.[33] The arm-based models (Models 4-5, 7) will be fit using the *pcnetmeta* package in the *R* software.[27]

9.9.5. Amendments to the statistical analysis plan

None.

9.10. 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.

9.11. Protection of human subjects

Subject information and 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.

Independent Ethics Committee (IEC)/Institutional Review Board (IRB)

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

Ethical conduct of the study

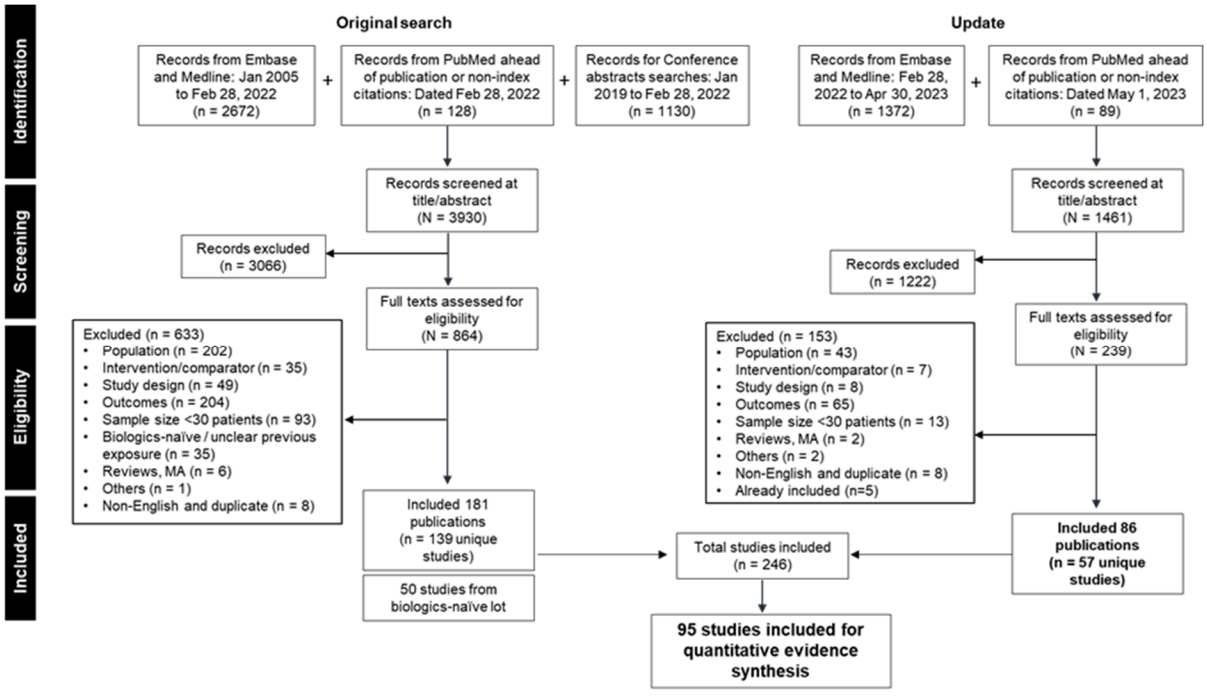
The study was conducted in accordance with legal and regulatory requirements, as well as with scientific purpose, value and rigor and followed generally accepted research practices described in Guidelines for Good Pharmacoepidemiology Practices (GPP) issued by the International Society for Pharmacoepidemiology (ISPE),[34] European Medicines Agency (EMA) European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) Guide on Methodological Standards in Pharmacoepidemiology.[35]

10. RESULTS

10.1. Participants

A total of 5,391 records were retrieved from literature. Of these, 4,288 were excluded during title and abstract screening, and 1,103 full-text publications were assessed for inclusion. After full-text screening, 319 publications were included. After linking the multiple publications of the same study, a total of 246 distinct observational studies were included for data extraction. A list of these studies is provided in Supplementary Table S4. A PRISMA flow diagram showing the study selection process is presented in **Figure 2**.

Figure 2. PRISMA flow diagram of the study selection



10.2. Descriptive data

Most of the included studies (76%) were peer-reviewed full publications. Nearly three-fourths of studies (74%) were published between 2019 and 2023, indicating that publications have increased in recent years. The majority of studies were from Europe (61%), followed by the US and Canada (15%), and Asia (14%). The study design of included studies comprised retrospective (66%), prospective (30%), post-marketing safety studies (2%), ambispective (1%), and unclear (1%).

In terms of ATs assessed across these studies, most studies (80%) were single-treatment studies (including 85 studies of anti-TNFα agents, 60 vedolizumab, 38 tofacitinib, 14 ustekinumab, and one upadacitinib). Forty-eight studies (20%) were comparative, containing 21 of vedolizumab (20 vs. anti-TNFα agents, and 1 vs. ustekinumab), 15 comparing different anti-TNFα agents, 11 of tofacitinib (5 vs. vedolizumab, 4 vs. ustekinumab, 1 vs. upadacitinib, and 1 vs. anti-TNFα agents), and one study of ustekinumab, vedolizumab and anti-TNFα agents. No observational studies were retrieved for filgotinib or ozanimod.

10.3. Outcome data

Ninety-five studies were included in the quantitative evidence synthesis. The baseline characteristics of patients in these studies are summarized in Supplementary Table S5. In the included studies, 68% had mixed population of biologic-naïve and biologic-exposed, 21% of studies had biologic-naïve patients, and 11% of studies had biologic-exposed patients. The median age of patients across studies ranged from 26 to 55 years. The use of concomitant steroid and immunomodulators varied considerably across studies, between 6% to 100% and 0% to 92% of patients, respectively.

10.4. Main results

The network diagram for each analysis is shown in Supplementary Figure S1 to Figure S10. **Table 3** summarizes the number of arms and patients per outcome. The results of a random effects NMA for overall population are summarized in **Figure 2** to **Figure 5** using adalimumab as the reference treatment, with subgroup results presented in Supplementary Figures S11 and S12. Note that for ease of interpretation, the anti-TNFα class was removed from the results presentation as individual ATs are of interest. The head-to-head pairwise results from the random effects and fixed effect NMA for all outcomes are presented in Supplementary Table S7 to Table S26.

Table 3. Summary of the number of studies and patients included in each analysis

		[Number of studies, Number of events, Total number of patients]						
Endpoint (Phase | Population)	**Studies included**	**Adalimumab**	**Golimumab**	**Infliximab**	**Tofacitinib**	**Ustekinumab**	**Vedolizumab**	**Anti-TNFα Class**

Table 3. Summary of the number of studies and patients included in each analysis

		[Number of studies, Number of events, Total number of patients]						
Clinical Response (Induction | Overall)	Comparative only	[3, 216, 387]	[2, 103, 187]	[1, 159, 251]	[1, 24, 38]	[1, 37, 101]	[5, 304, 469]	[3, 197, 300]
Clinical Response (Induction | Overall)	Comparative + Matched*	[10, 590, 1034]	[9, 462, 785]	[3, 314, 452]	[14, 668, 1027]	[4, 148, 295]	[21, 1084, 1678]	[3, 197, 300]
Clinical Response (Induction | Biologic Naïve)	Comparative + Matched*	[2, 108, 197]	[2, 23, 46]	NR	[1, 7, 9]	NR	[4, 52, 76]	[1, 63, 86]
Clinical Remission (Induction | Overall)	Comparative + Matched*	[6, 187, 492]	[6, 128, 402]	[1, 166, 251]	[13, 498, 1213]	[8, 201, 653]	[18, 790, 1732]	[3, 123, 300]
SF Remission (Induction | Overall)	Comparative + Matched*	[4, 159, 473]	[3, 66, 231]	[1, 69, 134]	[3, 74, 202]	[3, 66, 190]	[10, 251, 922]	[2, 58, 203]
Clinical Response (Maintenance | Overall)	Comparative + Matched*	[7, 291, 607]	[4, 192, 330]	[4, 301, 437]	[7, 360, 695]	NR	[11, 505, 768]	[2, 112, 194]
Clinical Remission (Maintenance | Overall)	Comparative + Matched*	[10, 903, 1666]	[4, 129, 418]	[5, 335, 625]	[9, 433, 903]	[1, 26, 33]	[17, 1134, 2273]	[4, 153, 394]

Table 3. Summary of the number of studies and patients included in each analysis

		[Number of studies, Number of events, Total number of patients]						
Clinical Remission (Maintenance | Biologic Naïve)	Comparative + Matched*	[2, 20, 41]	[1, 3, 15]	NR	[1, 18, 44]	NR	[6, 180, 316]	[2, 90, 244]
Clinical Remission (Maintenance | Biologic Pre-treated)	Comparative + Matched*	[2, 34, 83]	[1, 7, 35]	[1, 14, 29]	[1, 75, 220]	NR	[6, 262, 608]	[1, 35, 108]
SF Remission (Maintenance | Overall)	Comparative + Matched*	[3, 100, 345]	[2, 44, 139]	[2, 66, 157]	[7, 417, 1001]	NR	[9, 319, 792]	[2, 33, 143]
Serious AE (Overall)	Comparative + Matched*	[2, 6, 177]	[2, 1, 161]	[2, 32, 225]	[1, 4, 74]	[1, 4, 66]	[4, 42, 1079]	[2, 17, 160]

*: Matched refers to matched single-arm studies.

Abbreviations: SF: steroid-free; AE: adverse event; NR: not reported.

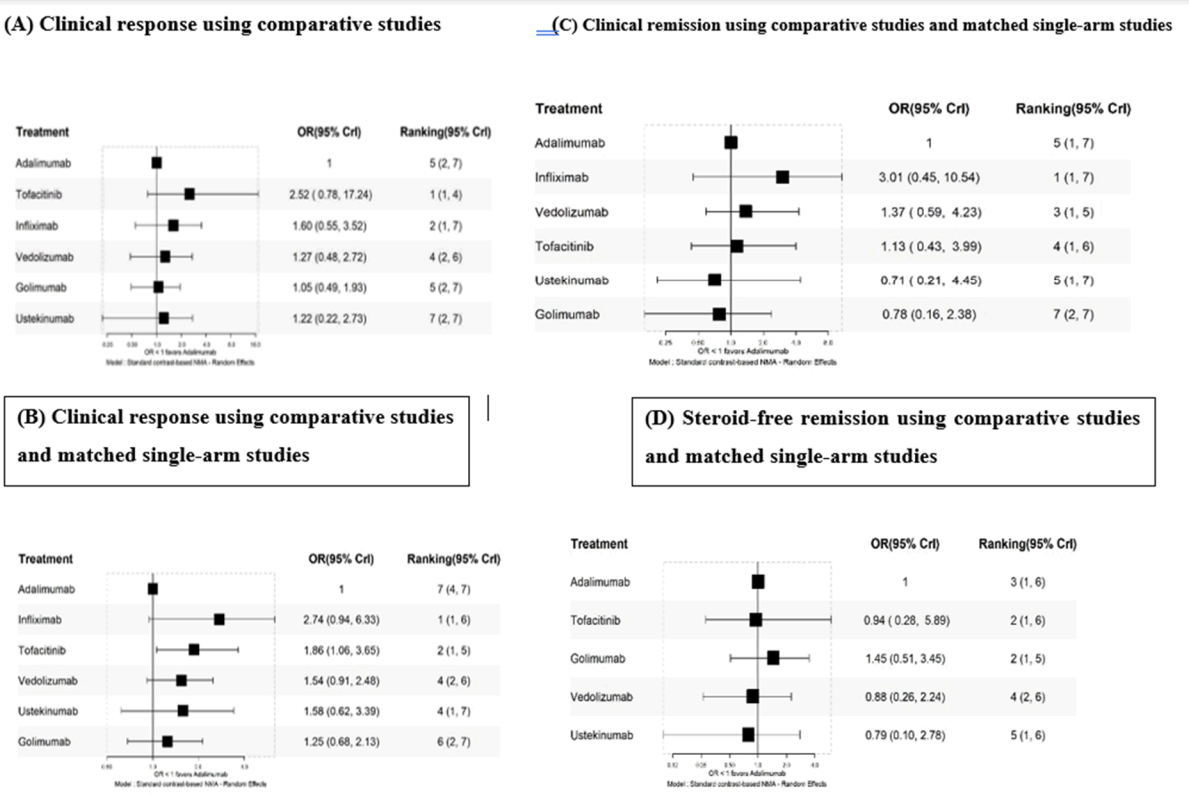
Induction phase

Clinical response: A total of 32 studies (comparative and matched single-arm) were included in the evidence synthesis for clinical response in the induction phase. The evidence from seven comparative studies formed a connected network for clinical response in the induction phase. Tofacitinib had the highest probability of being the most effective treatment, whereas infliximab had the second highest probability based on comparative studies only (Figure 3A). Tofacitinib was associated with significantly higher odds of inducing a clinical response compared to ustekinumab (Odds ratio [OR]: 3.34 with 95% credible interval [CrI] 1.02 to 24.11); but there were no significant differences between other ATs (Supplementary Table S7). After incorporating matched single-arm studies into the NMA, infliximab was ranked the first and tofacitinib being ranked the second (Figure 3B). Tofacitinib was associated with significantly higher odds of inducing a clinical response compared to adalimumab (OR: 1.86 with 95% CrI 1.06 to 3.65) and there were no significant differences between other ATs (Supplementary Table S9).

Clinical remission: In the evidence synthesis of clinical remission in the induction phase, a total of 28 comparative and matched single-arm studies were included. The results of NMA revealed that infliximab had the highest probability of being the most effective treatment for clinical remission in induction phase, followed by vedolizumab and tofacitinib (Figure 3C). No significant differences were noted between the ATs for clinical remission in the induction phase (Supplementary Table S11).

SF remission: The evidence network for SF remission was relatively sparse compared with clinical response and remission. Eleven comparative and matched single-arm studies were included in the evidence synthesis of SF remission in the induction phase. The results of NMA demonstrated that tofacitinib and golimumab had the highest probability of being the most effective treatment, followed by adalimumab (Figure 3D). There were no significant differences between the ATs (Supplementary Table S13).

Figure 3. Forest plot of relative effectiveness in the induction phase in the overall population



Abbreviations: OR: odds ratio; CrI: credible interval; NMA: network meta-analysis.

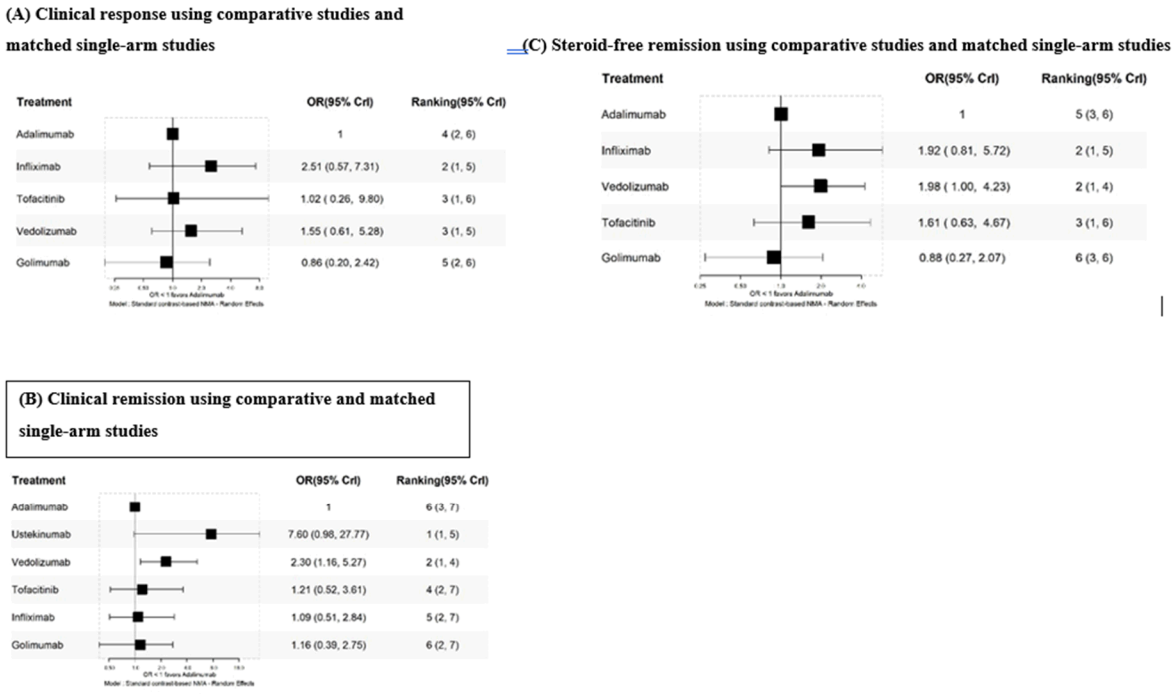
Maintenance phase

Clinical response: Sixteen studies were included in the evidence synthesis for clinical response in the maintenance phase. The results showed that infliximab had the highest probability of being the most effective treatment, followed by tofacitinib and vedolizumab (Figure 4A). There were no significant differences observed between the ATs (Supplementary Table S15).

Clinical remission: In total, 23 comparative and matched single-arm studies comprised in the evidence synthesis of clinical remission in the maintenance phase. The results showed that ustekinumab had the highest probability of being the most effective treatment, followed by vedolizumab and tofacitinib (Figure 4B). Vedolizumab was associated with significantly higher odds of maintaining clinical remission compared to adalimumab (OR: 2.30 with 95% CrI 1.16 to 5.27) and no significant differences were observed between other ATs (Supplementary Table S17).

SF remission: There were ten comparative and matched single-arm studies comprising the evidence synthesis of SF remission in the maintenance phase. Infliximab and vedolizumab had the highest probability of being the most effective treatment, followed by tofacitinib (Figure 4B). No significant differences were observed between the ATs (Supplementary Table S23).

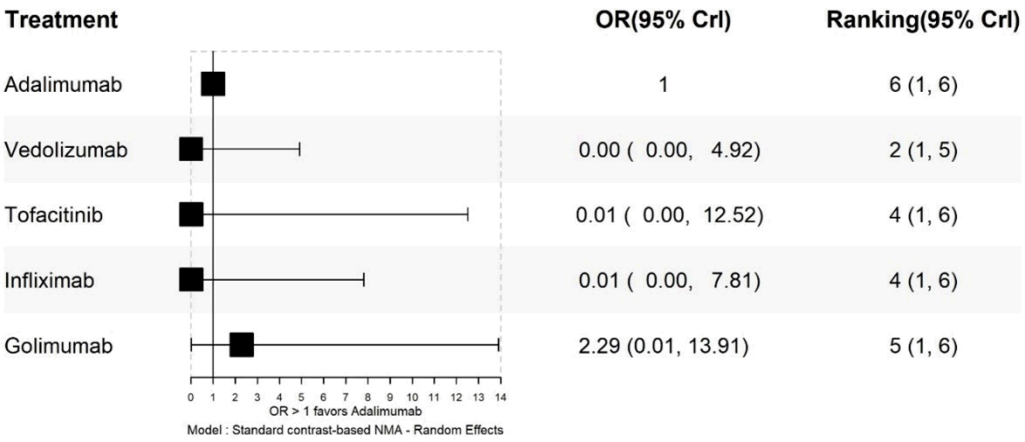
Figure 4. Forest plot of relative effectiveness in the maintenance phase in the overall population



Abbreviations: OR: odds ratio; CrI: credible interval; NMA: network meta-analysis.

A total of six comparative and matched single-arm studies were included in evidence synthesis of SAE. No significant differences were observed between the ATs (Supplementary Table S25). However, there are too few data to make any conclusions regarding the relative difference in odds of experiencing a Serious Adverse Events between the treatments.

Figure 5. Forest plot of serious adverse event in the overall population using comparative and matched single-arm studies



Abbreviations: OR: odds ratio; CrI: credible interval; NMA: network meta-analysis.

10.5. Other analyses

Sensitivity analyses were conducted by down-weighting the matched single-arm studies using power prior (Supplementary Table S27 to Table S50). The results were consistent with the base case analysis without down-weighting. Results from the fixed effect model provided similar point estimates as the random effects results with narrower credible intervals (Supplementary Table S8, S10, S12, S14, S16, S18, S20, S22, S24, and S26). Subgroup analysis in the biologic-naïve or -exposed subpopulations was only feasible for clinical remission in the maintenance phase. Vedolizumab and tofacitinib had the highest probability of being the most effective treatment in the biologic-naïve and biologic-exposed subpopulations, respectively (Supplementary Figure S11 to Figure S12).

10.6. Adverse events / adverse reactions

In total, 34 studies evaluated serious infection, including nine of infliximab, nine of tofacitinib, seven of adalimumab, six of vedolizumab, and one each of golimumab, ustekinumab and upadacitinib. Eight studies reported HZV infection comprising six of tofacitinib, one of vedolizumab, and one of golimumab. There were seven studies having VTE data, with six of tofacitinib and one of upadacitinib. Six studies of tofacitinib and one study of vedolizumab had data for MACE. Fourteen studies evaluated malignancies (4 infliximab, 4 tofacitinib, 3 adalimumab, and 3 vedolizumab). Mortality

data was provided in 12 studies (5 infliximab, 3 tofacitinib, 3 vedolizumab, and 1 adalimumab). Table 4 and 5 summarizes the exposure time and pooled incidence rate per 100 PYs for these safety events.

Table 4. Exposure time for adverse events of interest by treatment

AE	Treatment							
	Adalimumab	**Golimumab**	**Infliximab**	**Tofacitinib**	**Upadacitinib**	**Ustekinumab**	**Vedolizumab**	**Subcutaneous anti-TNFα**
Serious infection	7 studies ET=2003 PYs	1 study ET=1568 PYs	9 studies ET=11605 PYs	9 studies ET=1457 PYs	1 study ET=10.85 PYs	1 study ET= 107.83 PYs	6 studies ET=57874 PYs	1 study ET=94.96 PYs
HZV infection	NR	1 study ET=1568 PYs	NR	6 studies ET=1363 PYs	NR	NR	1 study ET=712.5 PYs	NR
VTE	NR	NR	NR	6 studies ET=1375 PYs	1 study ET=10.85 PYs	NR	NR	NR
MACE	NR	NR	NR	6 studies ET=1363 PYs	NR	NR	1 study ET=712.5 PYs	NR
Malignancies	3 studies ET=1644 PYs	NR	4 studies ET=10225 PYs	4 studies ET=1110 PYs	NR	NR	3 studies ET=57103 PYs	NR
Mortality	1 study ET=35 PYs	NR	5 studies ET=10878 PYs	3 studies ET=677 PYs	NR	NR	3 studies ET=56912 PYs	NR

ET = exposure time; PY= person-years.

Table 5. Pooled incidence rates for safety events of interest, by treatment

		Treatment [IR per 100 PY (95%CI)]							
AE		**Adalimumab**	**Golimumab**	**Infliximab**	**Tofacitinib**	**Upadacitinib**	**Ustekinumab**	**Vedolizumab**	**Subcutaneous anti-TNFα**
Serious infection	Fixed effect	2.1 (1.55 - 2.84)	0.57 (only one data point, ET=1568 PYs)	2.6 (2.33 - 2.92)	3.43 (2.6 - 4.53)	0 (only one data point, ET=10.85 PYs)	0.93 (only one data point, ET=107.83 PYs)	0.81 (0.74 - 0.88)	8.42 (only one data point, ET=94.96PYs)
	Random effects	2.7 (1.16 - 6.31)	0.	3.14 (1.8 - 5.48)	3.03 (1.17 - 7.86)			1.45 (0.6 - 3.5)	
HZV infection	Fixed effect	NR	0.13 (only one data point, ET=1568 PYs)	NR	2.2 (1.54 - 3.15)	NR	NR	0 (only one data point, ET=712.5 PYs)	NR
	Random effects	NR		NR	2.57 (1.36 - 4.84)	NR	NR		NR
VTE	Fixed effect	NR	NR	NR	0.73 (0.39 - 1.35)	0 (only one data point, ET=10.85 PYs)	NR	NR	NR
	Random effects	NR	NR	NR	0.73 (0.39 - 1.35)		NR	NR	NR
MACE	Fixed effect	NR	NR	NR	0.07 (0.01 - 0.52)	NR	NR	0.28 (only one data point, ET=712.5 PYs)	NR
	Random effects	NR	NR	NR	0.07 (0.01 - 0.52)	NR	NR		NR
Malignancies	Fixed effect	0.43 (0.2 - 0.89)	NR	0.79 (0.64 - 0.98)	0.72 (0.36 - 1.44)	NR	NR	0.21 (0.18 - 0.25)	NR
	Random effects	0.43 (0.2 - 0.89)	NR	0.83 (0.47 - 1.44)	0.72 (0.36 - 1.44)	NR	NR	0.21 (0.18 - 0.25)	NR
Mortality	Fixed effect	0 (only one data point, ET=35.27 PYs)	NR	0.13 (0.08 - 0.22)	0.44 (0.14 - 1.37)	NR	NR	0.09 (0.07 - 0.12)	NR
	Random effects		NR	0.12 (0.04 - 0.34)	0.44 (0.14 - 1.37)	NR	NR	0.24 (0.04 - 1.29)	NR

11. DISCUSSION

11.1. Key results

Treatment algorithms in UC have become more complex with newer therapies being approved and to be approved in the future, making it more difficult to determine the adequate drug ordering/sequencing for a specific patient. Thus, knowing the relative effectiveness and safety profile of ATs in UC will help to sufficiently weigh the benefit/risk of ATs against each other to provide a broad perspective that can be used to guide treatment decision-making. However, there is very limited or no evidence assessing the comparative effectiveness and safety of ATs in UC using real-world data from observational studies conducted in clinical practice. In the present study conducted as NMA, we compared the ATs for moderate-to-severe UC based on data from observational comparative and single-arm studies. Single-arm studies were matched with one another based on similarity in baseline characteristics to mitigate bias in the comparison between the studies.

In the induction phase, tofacitinib and infliximab were shown to have the highest probability of being the most effective treatments for clinical response; infliximab was also ranked the first for clinical remission. Tofacitinib was associated with statistically significantly better clinical response than ustekinumab and adalimumab based on comparative studies only and comparative studies plus matched single-arm studies, respectively.

In the maintenance phase, infliximab was ranked first for clinical response, ustekinumab was ranked first for clinical remission and infliximab and vedolizumab were ranked the first for SF remission. Vedolizumab was associated with significantly better clinical remission than adalimumab during maintenance.

Data were available for biologic-naïve and biologic-exposed subpopulations for clinical remission in the maintenance phase. NMA for these two subgroups also showed that no statistically significant difference between the ATs except in the case comparing vedolizumab with adalimumab in the biologic-naïve subpopulation in which vedolizumab was associated with significantly higher odds of inducing a clinical response compared to adalimumab (OR: 3.69; 95% CrI 1.23 to 21.16). The treatment ranking varied slightly: vedolizumab and tofacitinib have the highest probability of being the most effective treatment in the biologic-naive subpopulation and the biologic-exposed subpopulation, respectively.

No reliable network could be formed to compare the risk of adverse events among ATs. The IR per 100 PY for each AT for other safety events and mortality were generated. The purpose of this exploratory analysis was to contextualize with clinical trials rather than to compare treatment. Where reported, ATs safety profiles analyzed from this exploratory analysis are consistent with the findings from clinical trials.

Controlled clinical trials are regarded as the gold standard of evidence collection. Several NMAs have been published comparing the efficacy and safety of small molecules and biologics in patients with moderate-to-severe UC based on the data from RCTs.[10-12, 36-39] Observations in our study about the efficacy of ATs in evaluated outcomes based on observational studies appear to be consistent with those reported in NMAs based on RCTs. Singh et al. compared the efficacy and safety of different first-line (biologic-naïve) and second-line (prior exposure to TNF antagonists) ATs. Infliximab was ranked highest in biologic-naïve patients, and ustekinumab and tofacitinib were ranked highest in anti-TNFα-exposed patients, for induction of remission and endoscopic improvement in patients with moderate-to-severe ulcerative colitis.[39] Attauabi et al. reported tofacitinib and upadacitinib to be ranked the highest for the induction of clinical response and clinical remission.[36] Many analyses based on observational studies for tofacitinib[9, 40] vedolizumab,[41] and Ustekinumab.[14, 15] have reported effectiveness and safety of these ATs that were consistent with findings from RCTs.

One of the strengths of this study is a comprehensive evidence synthesis of data from observational studies to estimate the comparative effectiveness and safety of the ATs for UC patients in clinical practice. Extensive sensitivity analyses were conducted to investigate robustness of the results. In addition, potential biases from the single-arm studies were controlled by matching the covariates on the study level. This is equivalent to controlling observed confounding at the aggregate level. Random effects models taking into account heterogeneity among the studies was used as the base case to allow for appropriate propagating of uncertainty in the estimates.

The sensitivity analysis using the fixed effect model demonstrated consistent point estimates as the results from the random effects model, with narrower credible intervals as the model assumes the common treatment effect. The results from the sensitivity analysis using power prior to down-weight the matched single arm studies were also consistent with the results from the analysis without down-weighting, demonstrating the robustness of the NMAs conducted.

11.2. Limitations

This study has certain limitations that should be acknowledged. The limitations inherent to NMA methodology, such as assumptions and susceptibility to methodological quality of included studies, should be applied. This NMA is based on data collected in real-world settings which follows less stringent inclusion and exclusion criteria than the RCTs. Differences in study designs and patient characteristics across observational studies may result in considerable heterogeneity in reported outcomes. The results suggested moderate to high heterogeneity among the studies, reflecting the inclusion of observational studies which are potentially more heterogeneous than RCTs. The lack of patient-level data limited the investigations on the impact of differences between the study populations. There were a limited number of comparative studies reporting effectiveness outcomes, resulting in only one NMA

being feasible for the clinical response in the induction phase based on comparative studies only.

There were also limited safety data which were reported in the observational studies included in this analysis, which makes it difficult to compare safety among ATs. Some studies had small sample sizes and sparse networks for certain outcomes, increasing uncertainty in the results. The definitions of outcomes varied across included studies but were assessed to be comparable enough to synthesize. Despite these limitations, this analysis highlights the lack of comparative RWE studies comparing ATs directly. Further research using large registries or electronic health record data that capture these outcomes should focus on addressing this issue.

11.3. Interpretation

This study utilized NMA on real-world data from observational comparative and single-arm studies of ATs in moderate-to-severe UC. There is very limited or no evidence assessing the comparative effectiveness and safety of ATs in UC using real-world data from observational studies conducted in clinical practice, making it difficult to contextualize findings with previous evidence. This study found that in the induction phase, tofacitinib and infliximab were shown to have highest probability of being the most effective treatments for clinical response; infliximab was also ranked the first for clinical remission. In the maintenance phase, infliximab was ranked first for clinical response; ustekinumab was ranked first for clinical remission. These observations on the comparative effectiveness based on observational studies are consistent with findings from NMA based on RCTs. No significant differences were observed between the ATs for SAE. However, there are too few data to make any conclusions regarding the relative difference in odds of experiencing a Serious Adverse Events between the treatments. No network can be built for specific AESI to compare among ATs, but the IR of AESIs for tofacitinib reported in RW is generally consistent with that reported from clinical trial.

12. OTHER INFORMATION

Not Applicable.

13. CONCLUSIONS

In conclusion, this NMA based on real-world data from observational comparative and single-arm studies showed that in the induction phase, tofacitinib and infliximab were shown to have highest probability of being the most effective treatments for clinical response; infliximab was also ranked the first for clinical remission. In the maintenance phase, infliximab was ranked first for clinical response; ustekinumab was ranked first for clinical remission. These observations on the comparative effectiveness based on observational studies are consistent with findings from NMA based on RCTs. No significant differences were observed between the ATs for SAE. However, there are too few data to make any conclusions regarding the relative difference in odds of experiencing a Serious Adverse Events between the

treatments. No network can be built for specific AESI to compare among ATs, but the IR of AESIs for tofacitinib reported in RW is generally consistent with that reported from clinical trial. Overall, the findings from this NMA using observational studies, taken together with evidence from RCTs NMA, will help clinicians in decision-making to select the most appropriate therapy for the treatment of patients with moderate-to-severe UC in clinical practice.

14. REFERENCES

1. Feuerstein, J.D. and A.S. Cheifetz, *Ulcerative colitis: epidemiology, diagnosis, and management.* Mayo Clin Proc, 2014. **89**(11): p. 1553-63.
2. Kayal, M. and S. Shah, *Ulcerative Colitis: Current and Emerging Treatment Strategies.* J Clin Med, 2019. **9**(1).
3. America, C.s.C.F.o., *The Facts About Inflammatory Bowel Diseases*. 2014.
4. Rubin, D.T., et al., *ACG Clinical Guideline: Ulcerative Colitis in Adults.* Am J Gastroenterol, 2019. **114**(3): p. 384-413.
5. FDA, U. *Zeposia® (ozanimod)*. Highlights of prescribing information [cited 2021.
6. EMA, *Jyseleca® (filgotinib)*. 2021.
7. FDA, U., *Rinvoq® (upadacitinib)*. 2022.
8. Sandborn, W.J., et al., *Tofacitinib as Induction and Maintenance Therapy for Ulcerative Colitis.* New England Journal of Medicine, 2017. **376**(18): p. 1723-1736.
9. Taxonera, C., D. Olivares, and C. Alba, *Real-World Effectiveness and Safety of Tofacitinib in Patients With Ulcerative Colitis: Systematic Review With Meta-Analysis.* Inflamm Bowel Dis, 2022. **28**(1): p. 32-40.
10. Lasa, J.S., et al., *Efficacy and safety of biologics and small molecule drugs for patients with moderate-to-severe ulcerative colitis: a systematic review and network meta-analysis.* Lancet Gastroenterol Hepatol, 2022. **7**(2): p. 161-170.
11. Burr, N.E., et al., *Efficacy of biological therapies and small molecules in moderate to severe ulcerative colitis: systematic review and network meta-analysis.* Gut, 2021.
12. Lu, X., et al., *Comparative efficacy of advanced treatments in biologic-naïve or biologic-experienced patients with ulcerative colitis: a systematic review and network meta-analysis.* Int J Clin Pharm, 2023. **45**(2): p. 330-341.
13. Jin, Y., et al., *Meta-analysis of the effectiveness and safety of vedolizumab for ulcerative colitis.* World J Gastroenterol, 2015. **21**(20): p. 6352-60.
14. Uchida, G., et al., *Real-world effectiveness of ustekinumab for patients with ulcerative colitis: a systematic review and meta-analysis.* Nagoya J Med Sci, 2023. **85**(3): p. 402-427.
15. Taxonera, C., et al., *Meta-analysis: Real-world effectiveness and safety of ustekinumab in patients with ulcerative colitis.* Aliment Pharmacol Ther, 2023. **57**(6): p. 610-619.
16. Cumpston, M., et al., *Updated guidance for trusted systematic reviews: a new edition of the Cochrane Handbook for Systematic Reviews of Interventions.* Cochrane Database Syst Rev, 2019. **10**(10): p. Ed000142.
17. Page, M.J., et al., *The PRISMA 2020 statement: An updated guideline for reporting systematic reviews.* Int J Surg, 2021. **88**: p. 105906.
18. Sterne, J.A.C., et al., *RoB 2: a revised tool for assessing risk of bias in randomised trials.* BMJ, 2019. **366**: p. l4898.
19. Shea, B.J., 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**: p. j4008.
20. Dias, S. and A.E. Ades, *Absolute or relative effects? Arm-based synthesis of trial data.* Res Synth Methods, 2016. **7**(1): p. 23-8.
21. Hong, H., et al., *Rejoinder to the discussion of "a Bayesian missing data framework for generalized multiple outcome mixed treatment comparisons," by S. Dias and A. E. Ades.* Res Synth Methods, 2016. **7**(1): p. 29-33.

22. Dias S, W.N., Sutton AJ, et al, *NICE DSU Technical Support Document 2: A Generalised Linear Modelling Framework for Pairwise and Network Meta-Analysis of Randomised Controlled Trials [Internet]*. 2014: National Institute for Health and Care Excellence (NICE).
23. Dias, S., et al., *Network Meta-Analysis for Decision-Making*. 2018: Wiley.
24. Thom, H.H., et al., *Network meta-analysis combining individual patient and aggregate data from a mixture of study designs with an application to pulmonary arterial hypertension.* BMC Med Res Methodol, 2015. **15**: p. 34.
25. Schmitz, S., 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 Med Res Methodol, 2018. **18**(1): p. 66.
26. van Buuren, S. and K. Groothuis-Oudshoorn, *mice: Multivariate Imputation by Chained Equations in R.* Journal of Statistical Software, 2011. **45**(3): p. 1 - 67.
27. Lin, L., et al., *Performing Arm-Based Network Meta-Analysis in R with the pcnetmeta Package.* Journal of Statistical Software, 2017. **80**(5): p. 1 - 25.
28. Chu, H., et al., *Bivariate random effects models for meta-analysis of comparative studies with binary outcomes: methods for the absolute risk difference and relative risk.* Stat Methods Med Res, 2012. **21**(6): p. 621-33.
29. Hong, H., H. Fu, and B. Carlin, *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**.
30. Ibrahim, J.G., et al., *The power prior: theory and applications.* Stat Med, 2015. **34**(28): p. 3724-49.
31. Spiegelhalter, D.J., et al., *Bayesian measures of model complexity and fit.* Journal of the Royal Statistical Society: Series B (Statistical Methodology), 2002. **64**(4): p. 583-639.
32. G, S. *meta: An R package for meta-analysis.* 2007.
33. Sturtz, S., U. Ligges, and A. Gelman, *R2WinBUGS: A Package for Running WinBUGS from R.* Journal of Statistical Software, 2005. **12**(3): p. 1 - 16.
34. *Guidelines for good pharmacoepidemiology practice (GPP).* Pharmacoepidemiol Drug Saf, 2016. **25**(1): p. 2-10.
35. (EMA), E.M.A., *European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) Guide on Methodological Standards in Pharmacoepidemiology*. 2015.
36. Attauabi, M., et al., *Comparative onset of effect of biologics and small molecules in moderate-to-severe ulcerative colitis: a systematic review and network meta-analysis.* EClinicalMedicine, 2023. **57**: p. 101866.
37. Chu, X., et al., *Network meta-analysis on efficacy and safety of different biologics for ulcerative colitis.* BMC Gastroenterol, 2023. **23**(1): p. 346.
38. Panaccione, R., et al., *Efficacy and Safety of Advanced Therapies for Moderately to Severely Active Ulcerative Colitis at Induction and Maintenance: An Indirect Treatment Comparison Using Bayesian Network Meta-analysis.* Crohns Colitis 360, 2023. **5**(2): p. otad009.
39. Singh, S., et al., *First- and Second-Line Pharmacotherapies for Patients With Moderate to Severely Active Ulcerative Colitis: An Updated Network Meta-Analysis.* Clin Gastroenterol Hepatol, 2020. **18**(10): p. 2179-2191.e6.
40. Lucaciu, L., et al., *Real-world experience with tofacitinib in ulcerative colitis: a systematic review and meta-analysis.* Therapeutic Advances in Gastroenterology, 2021. **14**: p. 175628482110640.

41. Schreiber, S., et al., *Systematic review with meta-analysis: real-world effectiveness and safety of vedolizumab in patients with inflammatory bowel disease.* J Gastroenterol, 2018. **53**(9): p. 1048-1064.

15. LIST OF SOURCE TABLES AND FIGURES

Supplementary Document 1: Head-to-head pairwise results.docx

Supplementary Document 2: Network Plots.docx

Supplementary Document 3: Safety exploratory analysis.docx

Supplementary Document 4: Sensitivity Analysis.docx

Supplementary Document 5: SLR supplementary_15JUL24.docx

Supplementary Document 6: Subgroup Analysis.docx

Document Approval Record

Document Name:	Xeljanz_A3921447 Study Report v2 2Jan2025 1
Document Title:	A3921447

Signed By:	Date(GMT)	Signing Capacity
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]