NON-INTERVENTIONAL (NI) STUDY PROTOCOL

Study Information

Title	**RE**al **L**ife Safety and effectiveness of tof**A**citinib in comparison to **T**NF **I**nhibit**O**rs using the French **N**ational Healthcare Database (RELATION)
Protocol number	A3921399
Protocol version identifier	V1.0
Date	30 April 2021
EU Post Authorization Study (PAS) register number	EUPAS 41054
Active substance	Tofacitinib citrate
Medicinal product	Xeljanz
Research question and objectives	**Main objectives** To compare the risk of acute cardiovascular events between patients initiating a treatment by tofacitinib and patients initiating a treatment by TNFi* or vedolizumab (separately for UC, RA, and PsA patients). To compare the risk of malignancies, excluding non-melanoma skin cancer (NMSC) between patients initiating a treatment by tofacitinib and patients initiating a treatment by TNFi or vedolizumab* (separately for UC, RA, and PsA patients). **Secondary objectives** To compare the risk of malignancies, excluding non-melanoma skin cancer (NMSC) between patients initiating a

	treatment by tofacitinib and patients initiating a treatment by TNFi or vedolizumab*according to age [below and above 65 years old] (separately for UC, RA and PsA patients). To compare the risk of acute cardiovascular events between patients initiating a treatment by tofacitinib and patients initiating a treatment by TNFi* or vedolizumab according to age [below and above 65 years old] (separately for UC, RA and PsA patients). To compare the risk of serious infections between patients initiating a treatment by tofacitinib and patients initiating a treatment by TNFi* or vedolizumab (separately for UC, RA, and PsA patients and according to age [below and above 65 years old]). To compare the risk of thromboembolic events between patients initiating a treatment by tofacitinib and patients initiating a treatment by TNFi* or vedolizumab (separately for UC, RA, and PsA patients and according to age [below and above 65 years old]). To describe the risk factors/patients profiles related to the occurrence of each safety and effectiveness event of interest among patients initiated a treatment by tofacitinib To describe tofacitinib, TNFi* or vedolizumab treatment duration and persistence (includes drug survival, separately for UC, RA and PsA patients). To compare the effectiveness of tofacitinib vs TNFi or vedolizumab * (separately for UC, RA, and PsA patients). *TNFi are the reference comparators in RA/PSA indication. However in UC,

	vedolizumab is also a reference comparator and it will be added as comparator.
Author	**Amira BOUZIDI** Inflammation & Immunology France Senior medical Manager 23 25 avenue du Dr Lannelongue 75668 Paris cedex 14 France

1. TABLE OF CONTENTS

2. LIST OF ABBREVIATIONS

Abbreviation	Term
AE	adverse event
ALD	Affection de Longue Durée (long-term disease)
AS	ankylosing spondylarthritis
ATC	Anatomical Therapeutic Chemical
ATE	Arterial thrombEembolisms
ATIH	Technical Agency for Hospitalization Information
BID	twice daily
CABG	Coronary Artery Bypass Graft Surgery
CCAM	Classification Commune des Actes Médicaux (French medical classification for clinical procedure)
C. diff	Clostridium *difficile*
CépiDc	The cause of deaths database
CESREES	Comité éthique et scientifique pour les recherches, les études et les évaluations dans le domaine de la santé (Ethics and Scientific Committee for Research, Studies and Evaluations in the Health Field)
CI	confidence interval
CIP	pack identifier code
CMUc	Couverture maladie Universelle-complémentaire (state funded free complementary healthcare)
CNAM	Caisse nationale de l'assurance maladie
CNIL	Commission Nationale de l'Informatique et des Libertés (National Commission for Information Technology and Civil Liberties)

Abbreviation	Term
COVID	coronavirus disease
COVID-19	coronavirus disease-19
CRO	contract research organization
CRP	C-reactive protein
CT	clinical trial
CV	Cardiovascular
DAG	Directed Acyclic Graph
DAS	DiseaseActivity Score
DCIR	outpatient healthcare consumption data
DLP	Data Leak Prevention
DP	Principal Diagnosis
DVT	deep vein thrombosis
EMA	European Medicines Agency
EU	European Union
FDA	Food and Drug Administration
GHS	hospital stay group
GP	general practitioners
GPP	Good Pharmacoepidemiology Practices
HAD	homecare units
HDH	Health Data Hub
HDL	high-density lipoprotein
HDS	Health Data Host

Abbreviation	Term
HIV	human immunodeficiency virus
HR	hazard ratio
HZ	herpes zoster
ICD-10	International Classification of Diseases and Related Health Problems (10th revision)
IEC	Independent Ethics Committee
IL	interleukin
IMID	Immune-Mediated Inflammatory Disease
IR	Incidence Rates
IRB	Institutional Review Board
IV	intravenous
JAK	Janus Kinase
LTE	long-term extension
LPP	List of products and services
MACE	Major Adverse Cardiovascular Events
MCO	Medicine-Surgery-Obstetric
MRI	magnetic resonance imaging
MSM	marginal structural Cox proportional hazard regression models
MTX	Methotrexate
N	number of patients
NA	Not applicable
NABM	nomenclature of laboratory tests

Abbreviation	Term
NHI	National Health Insurance
NIR	Social Security Number
NMSC	non-melanoma skin cancer
OCTAVE	Oral Clinical Trials for tofAcitinib in ulceratiVE colitis (OCTAVE) clinical program
OLE	open-label extension
OPAL	Oral Psoriatic Arthritis Trial tofacitinib phase 3 trials
ORAL	Oral Rheumatoid Arthritis trial tofacitinib phase 3 trials.
OVH	Online Virtual Hosting
PAS	Post Authorization Study
PASS	Post-Authorization Safety Study
PE	pulmonary embolism
PMSI	hospital discharge database
PsA	psoriatic arthritis
Psy	psychiatric institutions
PY	patient-years
RA	rheumatoid arthritis
RCT	Randomized controlled trials
RSA	Anonymous Hospital Discharge Summary
SAP	Statistical Analysis Plan
SC	subcutaneous
SLM	Local Mutualist Section

Abbreviation	Term
SNDS	Système National des Donées de Santé (National Health Data System)
SNIIRAM	National Healthcare Insurance Database
SSR	rehabilitation units
TNF	tumor necrosis factor
TNFi	tumor necrosis factor inhibitor
TyK2	tyrosine kinase 2
UC	ulcerative colitis
UCD	Common classification of dispensable units
US	United States
VTE	Venous thromboembolism

3. RESPONSIBLE PARTIES

Principal Investigator(s) of the Protocol

Name, degree(s)	Job Title	Affiliation	Address
Jacques-Eric Gottenberg, M.D., Ph.D.	Rheumatologist	Strasbourg University Hospital	Hôpital de Hautepierre Avenue Molière 67200 Strasbourg
Julien Kirchgesner, M.D., Ph.D.	Gastroenterologist and Epidemiologist	AP-HP – Sorbonne University	Hôpital Saint-Antoine - AP-HP 184 Rue du Faubourg Saint-Antoine, 75012 Paris
Amira Bouzidi, M.D., M.Sc.	Senior Medical Manager Gastroenterology NI Study Lead	Pfizer France	23-25 avenue du Dr Lannelongue, Paris
Meriem Kessouri, M.D., M.Sc.	Senior Medical Manager Rheumatology	Pfizer France	23-25 avenue du Dr Lannelongue, Paris
Nadir Mammar, M.D.	Medical Director	Pfizer France	23-25 avenue du Dr Lannelongue, Paris
Jérémie Rudant, M.D., Ph.D.	Real World Data Medical Manager	Pfizer France	23-25 avenue du Dr Lannelongue, Paris
Pauline Lemeille, M.Sc.	Data Scientist	HEVA	186 Avenue Thiers, 69006 Lyon
Eleonore Herquelot, Ph.D.	Senior Biostatistician	HEVA	186 Avenue Thiers, 69006 Lyon
Fanny Raguideau, Pharm.D., M.Sc.	Senior Pharmacoepidemiologist	HEVA	186 Avenue Thiers, 69006 Lyon
Aurélie Schmidt, M.Sc.	Senior Project Manager	HEVA	186 Avenue Thiers, 69006 Lyon
Benjamin Grenier, Pharm.D., M.Sc.	Project Manager	HEVA	186 Avenue Thiers, 69006 Lyon

4. ABSTRACT

Refer to ANNEX 1.

5. AMENDMENTS AND UPDATES

None.

6. MILESTONES

Milestone	Planned date
Start of data collection	01 November 2021
End of data collection	01 December 2022
Registration in the EU PAS register	15 October 2021
Final study report	30 November 2023

7. RATIONALE AND BACKGROUND

7.1. Janus Kinases and Tofacitinib

Janus Kinases (JAK 1, 2 et 3, and TYK2) enzymes are part of the intracellular tyrosine kinases family that transduce cytokine-mediated signals via the JAK-STAT pathway. They physiologically regulate immunity and play a part in the pathological process of chronic inflammation.[1]

Tofacitinib is a potent, selective inhibitor of the JAK family. In enzymatic assays, tofacitinib inhibits JAK1, JAK2, JAK3, and to a lesser extent tyrosine kinase 2 (TyK2). In human cells, tofacitinib preferentially inhibits signaling by heterodimeric cytokine receptors that associate with JAK3 and/or JAK1 with functional selectivity over cytokine receptors that signal via pairs of JAK2. Inhibition of JAK1 and JAK3 by tofacitinib attenuates signaling of interleukins (IL-2, -4, -6, -7, -9, -15, -21) and type I and type II interferons, which will result in modulation of the immune and inflammatory response.[2,5]

Tofacitinib, has been marketed under the brand name Xeljanz and is registered as a 5 and 10 mg film coated tablet.

Xeljanz obtained its European marketing authorization on 23 March 2017 for **r**heumatoid **a**rthritis (RA) and its reimbursement authorization in France on 29 November 2017. **ps**oriatic **a**rthritis (PsA) and **u**lcerative **c**olitis (UC) indications for Xeljanz were later authorized, in June and July 2018, respectively, and reimbursed for both indications in July 2019 in France. The recommended dose for RA and PsA is 5 mg twice daily, and for UC, 10 mg twice a day for the first 8 weeks and thereafter 5 mg twice a day. Tofacitinib's target population in France reaches around 30,000 patients for RA, 12,000 patients for PsA, and 5,000 patients for UC.[2,4] In 2018 and 2019, France National Health Insurance's open data reported 5,000 patients treated by tofacitinib 5 mg (rheumatology indicated dose).

Tofacitinib efficacy and safety have been widely investigated in numerous randomized clinical trials (CT) through all its indications, These studies demonstrated sustained efficacy and manageable safety profile. Identified risks associated with tofacitinib include infection, lipid elevations, anemia, neutropenia and malignancies.[2,5]

Based on this safety profile, The United States (US) Food and Drug Administration (FDA) requested to further define the safety profile of tofacitinib 5 mg twice daily (BID) and 10 mg BID.

ORAL-Surveillance (A3921133; NCT02092467) is a Post-Authorization Safety Study (PASS) initiated in 2014 in response to the requirements of the FDA.

The primary objective of this clinical trial was to evaluate the safety of tofacitinib at two doses BID versus a tumor necrosis factor inhibitor (TNFi) in more than 4300 subjects with RA who were 50 years of age or older and had at least one additional cardiovascular risk factor (defined in the protocol as current cigarette smoker, high blood pressure, high-density lipoprotein [HDL] <40 mg/dL, diabetes mellitus, history of coronary artery disease, family history of premature coronary heart disease, extraarticular RA disease).

The co-primary endpoints of this study were adjudicated Major Adverse Cardiovascular Events (MACE) and adjudicated malignancies (excluding non-melanoma skin cancer[NMSC]).[15]

In February 2019 while the study A3921133 was still ongoing, during routine analysis of all safety events, the external, independent tofacitinib Rheumatology Data Safety Monitoring Board reported that, among other findings, the frequency of PE and all-cause mortality in patients receiving tofacitinib 10 mg twice daily was higher than in patients treated with a TNFi (full data set is not yet available).[7]

More recently, the co-primary endpoint results have been disclosed. Results showed for these co-primary endpoints, prespecified non-inferiority criteria were not met. Preliminary results suggest that these risks are associated with both approved dosage/dosing regimens (5 mg twice daily, and 10 mg twice daily which is approved only in UC).[8]

7.2. Medicinal Comparators

Xeljanz is recommended either in combination with Methotrexate (MTX) for RA and PsA[16,17] or as a monotherapy for RA and UC.[16,18] Randomized controlled trials (RCT) assessing tofacitinib (the Oral Rheumatoid Arthritis Trial(ORAL) RCTs for RA, the Oral Psoriatic Arthritis Trial (OPAL) RCTs for PsA, and the Oral Clinical Trials for tofAcitinib in ulceratiVE colitis (OCTAVE) RCTs for UC) have been conducted which included between 100 to 1,200 treated patients depending on the trial design. In France, this molecule is positioned as a second or third line of treatment depending on the pathology. The 10 mg dosage is reimbursed for the UC indication and the 5 mg dosage is reimbursed for UC, RA, and PsA.[2,4] Tofacitinib's main medicinal comparators are TNFi in all three indications with also vedolizumab in UC as detailed in Table 1.

Table 1. Tofacitinib Medicinal Comparators (TNFi and Vedolizumab)

Line of treatment	Rheumatoïd arthritis (RA)	Psoriatic Arthritis (PsA)	Ulcerative colitis (UC)
2nd line	**In combination with MTX** • **TNFi: Adalimumab, Infliximab, Etanercept, Certolizumab, and Golimumab** **Monotherapy** • **TNFi: Adalimumab, etanercept and certolizumab**		**TNFi: Adalimumab, infliximab and golimumab** **Vedolizumab**
3rd line	**TNFi : Adalimumab, Infliximab, Etanercept, Certolizumab, and Golimumab**	• **TNFi: Adalimumab, Infliximab, Etanercept, Certolizumab, and Golimumab**	**Vedolizumab**

7.3. Generic Adverse Events

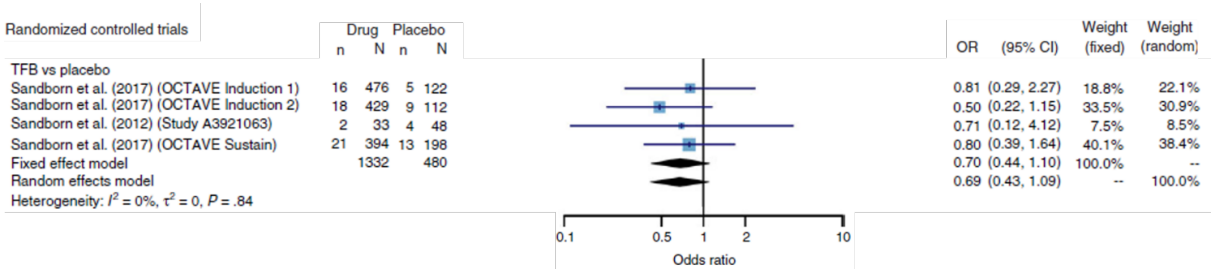
Tofacitinib safety profile has been described previously in the RA indication trials. As stated in the summary of product characteristics, the most commonly (≥1/100 to <1/10 patients) reported adverse reactions during the first 3 months in the controlled trials were headaches, upper respiratory infections, nasopharyngitis, diarrhea, nausea and hypertension. Overall, the safety profile of tofacitinib in patients affected by PsA or UC was consistent with the safety profile in patients affected by RA.[19]

7.4. Specific Adverse Events

The most recent published work assessing the safety of tofacitinib with a large number of patients is the Olivera et al. systematic review and meta-analysis which collected data from clinical trials assessing JAKi (JAK inhibitors) treatments in RA, UC and PsA.[6] This work aimed to assess adverse events, serious adverse events and adverse events of interest (mortality, serious infections, herpes zoster infections, non-melanoma skin cancer, other type of malignancy, and major adverse cardiovascular events including venous thromboembolism) in JAKi treated patients. The analyses, including more than 66,000 JAKi exposed patients, showed no increased risk in JAKi patients for almost all outcomes assessed compared with patients given a placebo. Herpes zoster risk of infection was significantly increased in patients receiving JAKi.

The results found by Olivera et al. were in accordance with the Bonovas et al. systematic review and meta-analysis, notably concerning serious adverse events in moderate-to-severe UC tofacitinib treated patients.[20] (Figure 1). A more recent network meta-analysis performed by Singh et al. in patients with moderate to severely active UC found similar results comparing tofacitinib 5 mg with ustekinumab 90 mg, vedolizumab, golimumab, adalimumab, Infliximab, and placebo for risk of infections and risk of serious adverse events. Mainly, a higher odds ratio of 1.78 (confidence interval [CI]95%: 1.02, 3.09) for risk of infections in tofacitinib 10 mg patients compared to vedolizumab patients has been found.[21]

Figure 1. Tofacitinib Serious Adverse Events in Moderate-to-severe Ulcerative Colitis (Odds ratio <1 favours tofacitinib and odds ratio >1 favours placebo) (20)



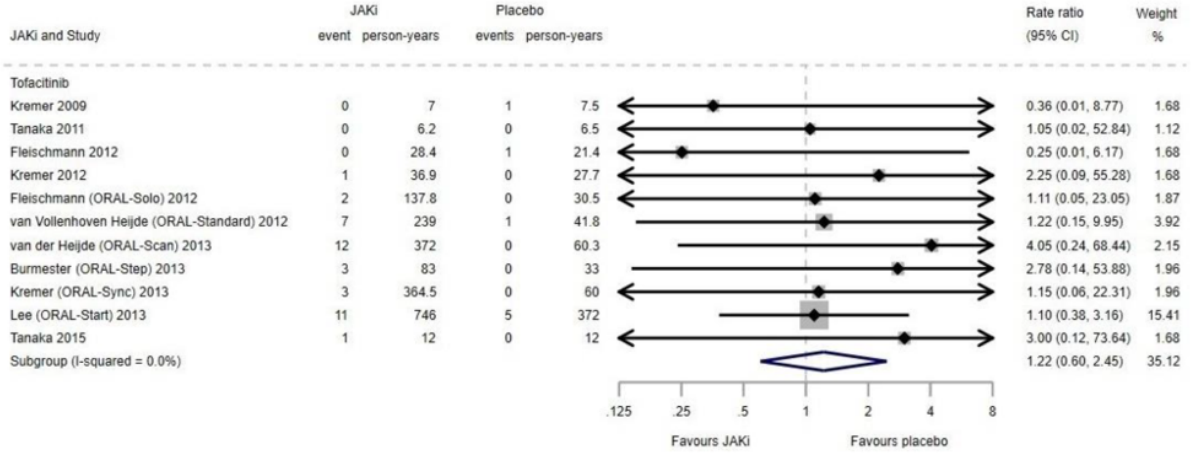
7.4.1. Malignancies (excluding non-melanoma skin cancer)

Immune-Mediated Inflammatory Disease (IMID) patients, and particularly RA patients are at increased risk of cancer compared with the general population.[14] Among all the types of cancer, lymphoma and lung cancer are the ones most associated with patients affected by RA. It is therefore important to assess the influence that immunomodulating and biological treatments (ie, Tofacitinib and TNFi) would have on the cancer outcome in already at risk patients. Data compiled from tofacitinib phase I-III trials and long-term extension (LTE) studies, analysing around 7,000 RA patients (22,785 patient-years exposure), presented a malignancy (excluding NMSC) IR of 0.8 per 100 patient-years.[9] This outcome was also analyzed in the ORAL surveillance study, the Hazard ratio for the risk of malignancies exluding NMSC was 1.48 (1.04, 2.09) [The non-inferiority criterion was not met for the primary comparison of the combined tofacitinib doses to TNF-alpha inhibitors since the upper limit of the 95% CI exceeded the pre-specified non-inferiority criterion of 1.8, ie, 2.09 >1.8].[22]

7.4.2. Serious Infections

Infections associated with death, hospitalization or intravenous antibiotic use are defined as serious infections. In the Olivera et al. study, the pooled global incidence rate (IR)of serious infections in more than 42,000 JAKi exposed patients reached 3.36 per 100 patient-years. The Bechman et al. systematic review and meta-analysis of infection risk with JAKi in RA found a slightly lower but still comparable crude incidence rate of serious infections of 1.97 per 100 patient-years in tofacitinib treated patients.[23] The serious infection incidence rate ratio in tofacitinib patients compared to placebo equaled 1.22 but did not reach statistical significance (Figure 2).

Figure 2. Serious Infection Incidence Rate Ratio in RA Patients Treated by Tofacitinib (23)



7.4.3. Herpes Zoster Infections

Recent data from tofacitinib clinical trials in either UC or RA patients allowed to compute a dose-dependent herpes zoster (HZ) infection incidence rates in tofacitinib patients ranging from 3.4 to 4.25 events for 100 patients-years.[24,25] A retrospective cohort of RA patients treated by tofacitinib in the MarketScan and Medicare databases (2011-2016), found comparable HZ infection incidence rates. The patients were treated either by tofacitinib alone or in combination with methotrexate and/or glucocorticoids.[26] The Olivera *et al.* systematic review and meta-analysis found lower but comparable HZ infection incidence rates of 2.11 per 100 patient-years.[6]

7.4.4. Opportunistic Infections

Results from pooled phase I, II, III, IIIb/IV trials and 2 open label LTE studies of RA patients treated by tofacitinib showed that opportunistic infections including viral, fungal and bacterial infections were reported with an IR (95% CI) of 0.4 (0.3-0.5) per 100 patients-years.[9] Similar results were found for PsA patients treated by tofacitinib were an IR (95% CI) of 0.3 (0.1-0.9) per 100 patients-years.[10] In UC patients treated by tofacitinib, data gathered from the OCTAVE Induction 1 & 2, and Sustain phase III trials and OCTAVE open-label extension (OLE) study showed an IR (95% CI) of 1.3 (0.8-2.0) per 100 patients-years.[11] To compare, data from French claim databases reported an IR of 0.21 per 100 patients-years for opportunistic infections in inflammatory bowel disease patients treated by TNFi.[27]

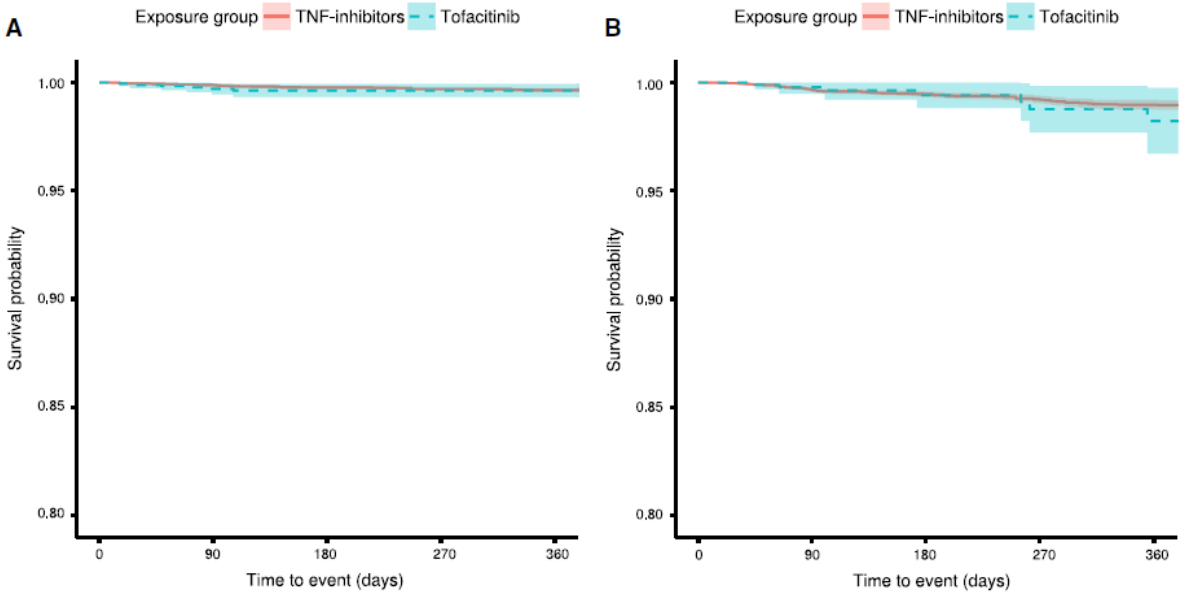
7.4.5. Thromboembolic Events

Venous thromboembolism (VTE) incidence rates has been shown to be around 2.4 times higher in RA patients than in age-matched and sex-matched non-RA patients and to be around 2.2 times higher in inflammatory bowel disease patients than in the general population.[28,29] An observational cohort study of RA patients assessing the VTE risk was performed on the Truven MarketScan and Medicare databases. The VTE risk in tofacitinib

RA patients was compared to the VTE risk in TNFi RA patients and no statistical difference was found. Propensity score–adjusted HRs showed no significant differences in the risk of VTE between the tofacitinib and TNF inhibitor treatment groups in either database.[12]

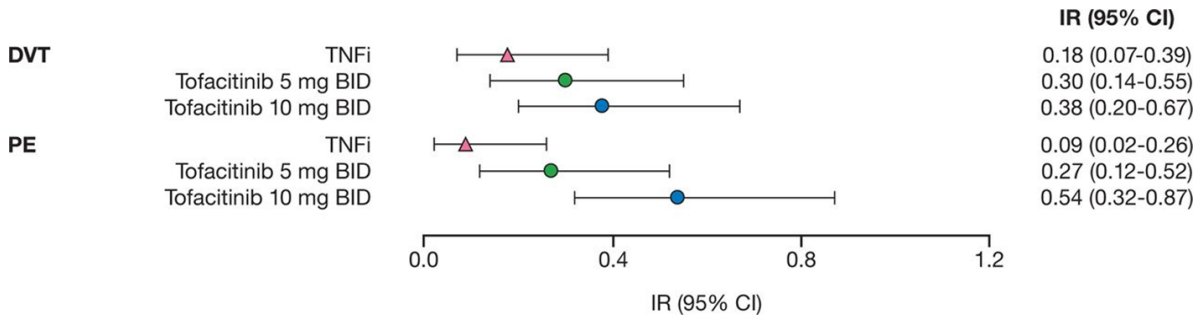
In addition, propensity score–weighted Kaplan-Meier plots (Figure 3) indicated no appreciable difference in the risk of VTE over the follow-up period between the 2 exposure groups, with widely overlapping 95% CIs for survival probabilities (P = 0.13 for the Truven MarketScan database and P = 0.70 for the Medicare database, by log rank test).[12]

Figure 3. Adjusted Kaplan-Meier Plots for the Risk of Venous Thromboembolism After Exposure to Tofacitinib or Tumor Necrosis Factor (TNF) Inhibitors in Rheumatoid Arthritis Patients in the Truven MarketScan database (2012-2016) (A) and the Medicare Database (2012–2015) (B). Shading Indicates 95% Confidence Intervals for Survival Probabilities (P = 0.13 for the Truven MarketScan Database and P = 0.70 for the Medicare Database, Tofacitinib *vs* TNF Inhibitors, by Log Rank Test).[12]



However, the interim results of ORAL Surveillance study thrombotic events (notably PE) was observed at an increased and dose-dependent incidence in patients treated with tofacitinib compared to TNFi (Figure 4).[30] Compared with TNFi, the HR for PE was 5.96 (1.75-20.33) and 2.99 (0.81-11.06) for tofacitinib 10 mg twice daily and tofacitinib 5 mg twice daily, respectively.[15,19]

Figure 4. Incidence Rates (95% CI) for DVT and PE Among Patients in Study A3921133 (ad hoc safety analysis). BID, Twice Daily; DVT, Deep Vein Thrombosis; IR, Incidence Rate (number of patients with an event per 100 PY of exposure); PE, Pulmonary Embolism; PY, Patient-years; TNFi, Tumour Necrosis Factor Inhibitor[30]



Of note, no difference in VTE risk was reported between tofacitinib and TNFi RA patients in an observational study based on US claims databases RR (95% CI) = 1.33 (0.78–2.24).[12]

Arterial thromboembolisms (ATE) were reported in the Cohen et al. comprehensive analysis of RA patients treated by tofacitinib and their IR (95% CI) was 0.4 (0.3-0.5) per 100 patients-years.[9] ATE IRs were similar according to tofacitinib doses. A similar IR (95% CI) ranging from 0.34 (0.13-0.69) to 0.4 (0.1-1.0) per 100 patient-years was computed for PsA patients treated by tofacitinib.[10,30]

7.4.6. Major Adverse Cardiovascular Events

Adjudicated Major Adverse Cardiovascular Events (MACE) from tofacitinib phase III and OLE trials included myocardial infarction, stroke, and/or cardiovascular (CV) death. For RA patients treated by tofacitinib, adjudicated MACE IR (95% CI) was 0.4 (0.3 to 0.5) per 100 patients-years.[9] In active PsA patients treated by tofacitinib, adjudicated MACE IR (95% CI) was 0.3 (0.1-0.8) per 100 patients-years and were myocardial infarction, ischemic stroke, and stress cardiomyopathy.[10] In UC patients treated by tofacitinib, IR (95% CI) was 0.2 (0.1-0.6) per 100 patients-years and encompassing hemorrhagic stroke, aortic dissection, acute coronary syndrome and myocardial infarction).[11]

Compared to TNF inhibitors Hazard ratio for the risk of adjudicated MACE reported in the preliminary results of the ORAL Surveillance study was 1.33 (0.91, 1.94). The non-inferiority criterion was not met for the primary comparison of the combined tofacitinib doses to TNF-alpha inhibitors since the upper limit of the 95% CI exceeded the pre-specified non-inferiority criterion of 1.8, ie, 1.94 >1.8.[22]

Cardiovascular deaths are taken into account for the MACE definition. However, we will not consider them in the study and therefore will use this composite endpoint (included myocardial infarction and/or stroke) “Acute cardiovascular event”.

The preliminary results of the post authorization safety study (A392113,[30] led the FDA and the European Medicine Agency (EMA) to amend in 2020 the Xeljanz US/European Union (EU) labelling.

The current Xeljanz EU label includes:

- A special warning regarding the use of tofacitinib in patients with known risk factors for VenousThromboembolic Events (VTE), regardless of indication and dosage ;
- Restriction of use of tofacitinib 10 mg twice daily for maintenance treatment in patients with UC who have known VTE risk factors, only if no suitable alternative treatment is available ;
- Restrtiction of use of tofacitinib in patients over 65 years only if no suitable alternative treatment is available.[7]

Further evaluation of the results of the recently completed study A3921133. and their potential impact on tofacitinib product information by EMA is currently ongoing.[31]

Given the discrepancy between the preliminary findings of Oral-Surveillance study (A3921133) and the safety data accumulated to date on tofacitinib, by using insightful data sources, Pfizer would like to assess tofacitinib effectiveness and safety through a PASS real-life data in France, and compare it to TNFi, separately for UC, RA, and PsA patients. In France those results will help to better understand the safety profile of tofacitinib in a wild and heterogenous IMID population

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

8. RESEARCH QUESTION AND OBJECTIVES

The objectives of this study are mainly to assess tofacitinib safety compared to TFNi* or vedolizumab through French national healthcare databases. This study will be performed separately for the rheumatology (RA, and PsA), and for the gastroenterology patients (UC).

8.1. Primary Objectives

- To compare the risk of malignancies, excluding non-melanoma skin cancer (NMSC) between patients initiating a treatment by tofacitinib and patients initiating a treatment by TNFi or vedolizumab* (separately for UC, RA, and PsA patients).
- To compare the risk of acute cardiovascular events between patients initiating a treatment by tofacitinib and patients initiating a treatment by TNFi* or vedolizumab (separately for UC, RA, and PsA patients).

8.2. Secondary Objectives

- To compare the risk of malignancies, excluding non-melanoma skin cancer (NMSC) between patients initiating a treatment by tofacitinib and patients initiating a treatment by TNFi or vedolizumab* (separately for UC, RA, and PsA patients and according to age [below and above 65 yo]).
- To compare the risk of acute cardiovascular events between patients initiating a treatment by tofacitinib and patients initiating a treatment by TNFi* or vedolizumab (separately for UC, RA, and PsA patients and according to age [below and above 65 yo]).
- To compare the risk of serious infections between patients initiating a treatment by tofacitinib and patients initiating a treatment by TNFi* or vedolizumab (separately for UC, RA, and PsA patients and according to age [below and above 65 yo]).
- To compare the risk of thromboembolic events between patients initiating a treatment by tofacitinib and patients initiating a treatment by TNFi* or vedolizumab (separately for UC, RA, and PsA patients and according to age [below and above 65 yo]).
- To describe the risk factors/patients profiles related to the occurrence of each safety and effectiveness event of interest among patients initiated a treatment by tofacitinib.
- To describe tofacitinib, TNFi* or vedolizumab treatment duration and persistence (includes drug survival, separately for UC, RA and PsA patients).
- To compare the effectiveness of tofacitinib vs TNFi or vedolizumab * (separately for UC, RA, and PsA patients).

*TNFi are the reference comparators in RA/PSA indication. However in UC, vedolizumab is also a reference comparator and it will be added as comparator.

9. RESEARCH METHODS

9.1. Study Design

We will perform a retrospective observational cohort study using the National Health Data System (SNDS).

This is a secondary data collection structured data analysis.

This study's main strength is its large and nation-wide coverage made possible by the use of two comprehensive and complementary databases (DCIR and PMSI), combined into the National Healthcare Insurance Database (SNIIRAM) database, whose data are retrospectively and independently collected. In addition, as data are collected for reimbursement purposes and not for the purpose of this study, there will be no patient selection bias.

9.2. Setting

All adults initiating tofacitinib or TNFi* for RA, PsA or UC or vedolizumab for UC between 01 January 2010 and 31 December 2021 will be included.

9.2.1. Inclusion Criteria

Patients must meet all of the following inclusion criteria, during the inclusion period, to be eligible for inclusion in the study:

- RA and PsA cohorts:

1. Adult (aged ≥18 years)

2. TNFi naïve and experienced patients initiating a treatment by tofacitinib or a new TNFi for RA and/or PsA between 01 November 2017 and 31 December 31 2020 and patients initiating a treatment by TNFi for RA and/or PsA between 01 January 2010 and 31 December 2020. Identification of RA or PsA patients will be based on long-term diseases (ALDs) and/or hospitalization discharges including International Classification of Diseases and Related Health Problems (10th revision) (ICD-10) codes of RA or PsA reported in Table 2.

   For patients with multiple hospitalizations with ICD-10 codes related to either RA or PsA, we propose to retain the most recent diagnosis at cohort entry.[32]

- UC cohort:

1. Adult (aged ≥ 18 years).

2. TNFi naïve and experienced patients initiating a treatment by tofacitinib or a new TNFi or vedolizumab for UC between 01 July 2019 and 31 December 2021 and patients initiating a treatment by TNFi or vedolizumab for UC between 01 January 2010 and 31 December 2021. Identification of UC patients will be based on ALDs and/or hospitalization discharges including ICD-10 codes of UC reported in Table 2.

For patients with multiple hospitalizations with ICD-10 codes related to either UC or Crohn disease (TNFi being also indicated for Crohn disease), we propose to retain the most recent diagnosis at cohort entry.[32]

1. The index date will be the initiation of TNFi*, or tofacitinib, or vedolizumab.

The initiation of a treatment will be defined by the absence of dispensing/delivery of that molecule during the year prior to the index date.

1. For a given subject, only codes entered in the database before the index date will be considered.

Table 2. ICD-10 Codes Selected for Study Population Inclusion Through ALDs and/or Hospitalizations

Pathology	ICD-10 code	Data source(s)
RA	M05*	Seropositive rheumatoid arthritis
	M06*	Other rheumatoid arthritis
PsA	L405	Arthropathic psoriasis
	M07.0	Distal interphalangeal psoriatic arthopathy
	M07.1	Arthritis mutilans
	M07.2	Psoriatic spondylitis
	M07.3	Other psoriatic arthropathies
UC	K51*	Ulcerative colitis

* TNFi: etanercept, infliximab, adalimumab, certolizumab pegol, golimumab for RA, and PsA, and infliximab, adalimumab, and golimumab for UC.

9.2.2. Exclusion Criteria

Patients meeting any of the following criteria will not be included in the study:

1. For UC cohort: Patients for whom the most recent hospital diagnosis at cohort entry is for Crohn disease (TNFi being also indicated for Crohn disease)

2. Patients with the safety outcome of interest at time of inclusion: as a severe infection within the 60 days prior to the index date;[33] and/or an acute cardiovascular or thromboembolic events within the whole pre index date) [See ANNEX 4. EXCLUSION CRITERIA ICD-10 AND CCAM CODES].

3. Patients with immunodeficiency status within the whole pre-index date or an active cancer (because chemotherapy may lead to immunosuppression and increased the risk of infection, and cancer increases VTE risk) during the 2 years prior to index date (See ANNEX 4. EXCLUSION CRITERIA ICD-10 AND CCAM CODES).

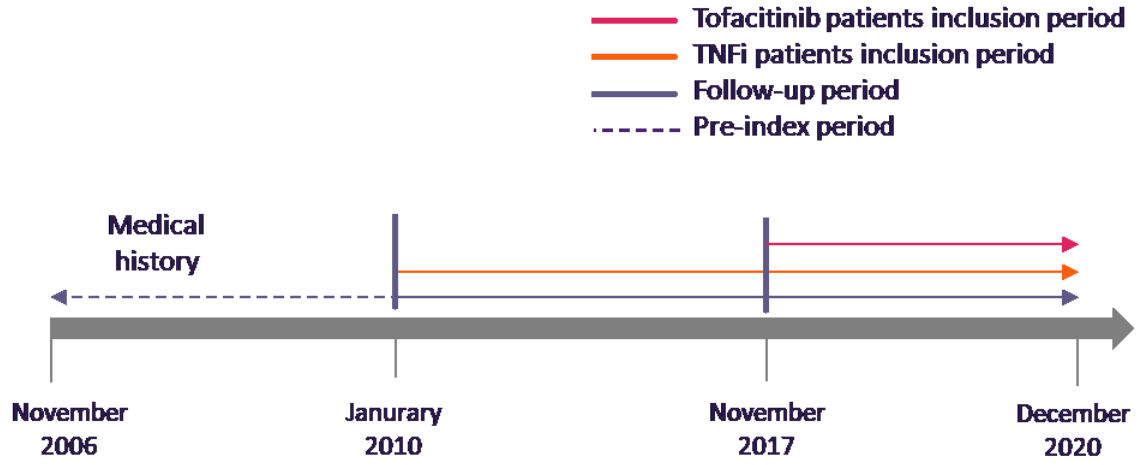
4. Due to SNIIRAM technical constraints: patients with non-unique identifiers (patients with mock, temporary or multiple identifiers)

9.2.3. Study Periods

For comparative analyses, two different study periods according to dates of French marketing authorizations will be proposed. As few patients will be treated with TNFi or vedolizumab in third line over the period 2017 to 2021 (due to tofacitinib availability), the inclusion period for this cohort will be extended to 2010 with a pre index period until 2006. This would enable us to capture patients treated with TNFi or vedolizumab in the third line and have comparable patients.

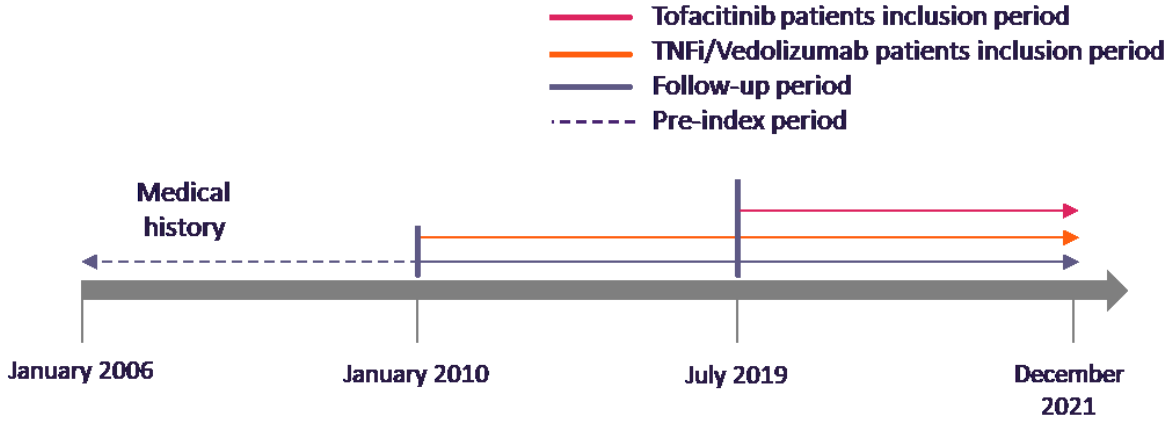
Study period 1: Patients initiating a treatment by tofacitinib or TNFi for RA and/or PsA between 01 November 2017 and 31 December 2020 and patients initiating a treatment by TNFi for RA and/or PsA between 01 January 2010 and 31 December 2020 (Figure 5).

Figure 5. Study Period 1



Study period 2 : Patients initiating a treatment by tofacitinib for UC between 01 July 2019 and 31 December 2021 and patients initiating a treatment by TNFi or vedolizumab for UC between 01 January 2010 and 31 December 2021 (Figure 6).

Figure 6. Study Period 2



9.2.4. End of Follow-up

Detailed end of follow-up motives are presented below.

9.2.4.1. Primary Objectives

- Malignancy (excluding NMSC):

- RA and PsA: Patients will be followed up until: 31 December 2020, death, treatment discontinuation, loss to follow-up (defined as the absence of healthcare consumption in the database for one year), occurrence of cancer, whichever occurs first. In case of loss to follow-up, end of follow-up will be the last know contact date, defined by the last claim in the database.

- UC: Patients will be followed up until: 31 December 2021, death, treatment discontinuation, loss to follow-up (defined as the absence of healthcare consumption in the database for one year), occurrence of acute cardiovascular events or cancer, whichever occurs first. In case of loss to follow-up, end of follow-up will be the last know contact date, defined by the last claim in the database.

- Acute cardiovascular events:

  - RA and PsA: Patients will be followed up until: 31 December 2020, death, treatment discontinuation, loss to follow-up (defined as the absence of healthcare consumption in the database for one year), occurrence of acute arterial event, whichever occurs first. In case of loss to follow-up, end of follow-up will be the last know contact date, defined by the last claim in the database.

  - UC: Patients will be followed up until: 31 December 2021, death, treatment discontinuation, loss to follow-up (defined as the absence of healthcare consumption in the database for one year), occurrence of acute arterial event, whichever occurs first. In case of loss to follow-up, end of follow-up will be the last know contact date, defined by the last claim in the database.

9.2.4.2. Secondary Objectives

- Malignancy (excluding NMSC): See Primary Objectives section above.

- Acute cardiovascular events: See Primary Objectives section above.

- Serious infections:

  - RA and PsA: Patients will be followed up until: 31 December 2020, death, treatment discontinuation, loss to follow-up (defined as the absence of healthcare consumption in the database for one year), occurrence of serious infection or cancer, whichever occurs first. In case of loss to follow-up, end of follow-up will be the last know contact date, defined by the last claim in the database.

- UC: Patients will be followed up until: 31 December 2021, death, treatment discontinuation, loss to follow-up (defined as the absence of healthcare consumption in the database for one year), occurrence of serious infection or cancer, whichever occurs first. In case of loss to follow-up, end of follow-up will be the last know contact date, defined by the last claim in the database.

- Venous thromboembolic events:

  - RA and PsA: Patients will be followed up until: 31 December 2020, death, treatment discontinuation, loss to follow-up (defined as the absence of healthcare consumption in the database for one year), occurrence of venous thromboembolic event, whichever occurs first. In case of loss to follow-up, end of follow-up will be the last know contact date, defined by the last claim in the database.

  - UC: Patients will be followed up until: 31 December 2021, death, treatment discontinuation, loss to follow-up (defined as the absence of healthcare consumption in the database for one year), occurrence of venous thromboembolic event, whichever occurs first. In case of loss to follow-up, end of follow-up will be the last know contact date, defined by the last claim in the database.

9.3. Variables

9.3.1. Drug Exposure

In France, infliximab, adalimumab, golimumab, etanercept, certolizumab pegol, tofacitinib, and vedolizumab are approved by the regulatory authorities for the treatment of RA, PsA, UC (Table 3) and are prescribed in hospitals and private clinics.

Etanercept, adalimumab, certolizumab pegol, and golimumab, are administered subcutaneously every two weeks. Tofacitinib is administered per os following different dosages depending on the indication: 5 mg twice a day for RA or PsA and 10 mg twice a day for UC. The 5 mg and 10 mg tofacitinib dosage will be taken into account by counting the number of treatment boxes between two pharmacy dispensing. All are dispensed by pharmacies for one month.[34] Infliximab and vedolizumab are administered intravenously every two months as a maintenance dosage.

Therefore, patients who received infliximab or vedolizumab (administered intravenously) will be considered exposed for two months following an intravenous (IV) administration. Those who received adalimumab, etanercept, certolizumab pegol, golimumab or tofacitinib will be considered exposed for one month following delivery.[35]

Table 3. TNFi, Vedolizumab, and Tofacitinib Marketing Authorizations for RA, PsA, and UC.

Molecule	RA	PsA	UC
Etanercept	L1/L2/L3	L2/L3	Without MA
Infliximab	L1/L2/L3	L2/L3	L2
Adalimumab	L1/L2/L3	L2/L3	L2
Certolizumab pegol	L1/L2/L3	L2/L3	Without MA
Golimumab	L1/L2/L3	L2/L3	L2
Tofacitinib 5 mg 10 mg	 L2/L3 Without MA	 L3 Without MA	 L2/L3 L2/L3
Vedolizumab	Without MA	Without MA	L2/L3
MA: Marketing Authorization, L1: First line of treatment, L2: Second line of treatment, L3: Third Line of treatment			

9.3.1.1. Treatment Duration

Treatment discontinuation will be defined as no dispensing within 3 months after the last estimated tofacitinib exposure (the exposure period will be estimated based on dosing schedules for individual agents). Switching between exposure groups will resulted in immediate censoring.

Treatment persistence will be assessed using two indicators: median duration of treatment and rate of patients under treatment at 1 year.

9.3.2. Outcomes

9.3.2.1. Safety

- Malignancy (excluding NMSC): defined as a malignancy (excluding NMSC) primary discharge diagnosis. Malignancies will be classified according to the following categories: breast cancer, lung cancer, lymphoma, melanoma, and other active cancer. Table 7 provides malignancies ICD-10 codes.
- Acute cardiovascular event: Acute cardiovascular events will be defined as a composite outcome of first hospital admission or procedures specifically related to ischemic heart disease (including myocardial infarction), cerebrovascular disease (including stroke), and peripheral artery disease (excluding acute mesenteric ischemia, related ICD-10 codes as primary diagnosis or procedure code.) Table 11 provides Acute cardiovascular events ICD-10 and procedure codes.

- Serious infection: defined as a diagnosis of infection requiring hospitalization (related ICD-10 codes as primary diagnosis). Within the SNIIRAM database, the diagnoses of infection requiring hospitalization and the type of infection have been shown to be accurate in 97% and 98% of the cases, respectively.[36] Serious infections will be classified according to infection sites. These include pulmonary; gastrointestinal; skin; urinary tract; ear, nose and throat; musculoskeletal; and other infections (including sepsis, non-classified opportunistic, and mycobacterial infections). Opportunistic infections will be classified according to pathogens. These will include viral, mycobacterial, bacterial, fungal, and parasitic infections. Table 8 and Table 9 provide infection diagnoses and related ICD-10 codes according to infection sites and pathogens.
- Venous thromboembolic events (VTE): including pulmonary embolism (PE) or deep vein thrombosis (DVT) (related ICD-10 codes as primary diagnosis). VTE outcome will be validated via a sensitivity analysis consisting in combining patients having undergone Doppler ultrasounds during the inclusion stay and initiating an anticoagulant therapy within the 4 months following the hospitalization. Table 10 provides VTE-related ICD-10 codes:
- For VTE identified during hospital stays thanks to relevant ICD-10 codes, an anticoagulant treatment dispensing during the month following the end of the hospital stay will be sought.
- For VTE occurring outside of hospital stays and without relevant ICD-10 codes, the identification will consist in a combination of doppler ultrasound followed by an anticoagulant treatment dispensing during the 7 days after the test. The absence of anticoagulant treatment during the 3 months before the ultrasound will be necessary to avoid poor characterization of the event.

9.3.2.2. Effectiveness

9.3.2.2.1. Ulcerative Colitis

Effectiveness will be assessed via a composite endpoint encompassing: UC-related hospitalization, UC-related surgery, all-cause hospitalization except childbirth for at least one night or switch to another drug (tofacitinib, TNFi, or vedolizumab[35,37] or corticosteroids exposure. Only the first event will be considered (Table 13 and Table 15).

9.3.2.2.2. Rheumatoid Arthritis

Effectiveness will be assessed via a composite endpoint encompassing: revascularization or Coronary Artery Bypass Graft Surgery (CABG) procedures, intra-articular injection, all-cause hospitalization except childbirth for at least one night or switch to another drug (tofacitinib and TNFi) (Table 12, Table 14, and Table 15).

9.3.2.2.3. Psoriatic Arthritis

Effectiveness will be assessed via a composite endpoint encompassing: revascularization or CABG procedures, intra-articular injection, all-cause hospitalization except childbirth for at least one night or switch to another drug (tofacitinib and TNFi) (Table 12, Table 14, and Table 15).

9.3.3. Covariates

9.3.3.1. Time-fixed covariates

The following covariates will be measured at index date (Table 6):

- Age, Gender;
- French universal healthcare (CMUc);
- Costly ALD;
- Disease duration defined as the time since an ALD or first hospitalization for the IMID (Immune Mediated Inflammatory Disease) of interest;
- Comorbidities of interest: History of cancer, cirrhosis and portal hypertension, respiratory chronic disease, chronic kidney disease, atherosclerosis, cardiovascular risk factors (diabetes, dyslipidemia, alcohol use disorder, hypertension, smoking behavior, obesity, and oral contraceptives), other inflammatory bowel disease (IBD), RA, PsA, ankylosing spondylarthritis (AS).
- Complications related to the IMID (using a 6 months pre-index period):
    - UC: UC-related hospitalization, UC-related surgery;
    - RA: RA-related hospitalization, RA-related surgery;
    - PsA: PsA -related hospitalization, PsA-related surgery.
- General Practitioner consultations (within the last 6 months).
- Specialists consultations (gastroenterologist and/or rheumatologist within the last 6 months).
- Number of C-reactive protein (CRP) dosage (within the last 6 months).
- Clostridium *difficile* (C. diff) and standard coproculture (within the last 6 months).
- CornaVirus Disease (COVID-19) hospitalization (within the last 6 months).

- Previous (within the last 6 months) and concomitant treatment (MTX, azathioprine, aminosalycilates, 6-mercaptopurin, ciclosporin).

- Disease activity assessment at index date (digestive endoscopy, radiology tests).

9.3.3.2. Time-varying Covariates

The following covariates will be updated (Table 6):

- Corticosteroid use (updated each month);

- Dosage tofacitinib (updated each month);

- Disease activity (updated each 6 months) defined as IMID-related hospitalizations or surgery.

9.4. Data Sources

This study will be carried out on data from the National Health Data System (SNDS). The SNDS collates and pseudoanonymises health data collected by public bodies. To date, this pool of data concerns three pre-existing databases:

1. The National Health Insurance (NHI) Database (SNIIRAM).

2. Hospital and other healthcare facilities (the hospital discharge database [PMSI]).

3. The cause of deaths database (CépiDc).

For this study we will ask for access to the NHI Database (SNIIRAM)

The French national health insurance system

The majority of medical expenses are reimbursed by the French NHI,[38,39] usually about 70% of doctors' fees, 80% of hospital fees and 65% of prescribed medications considered to be useful and 100% of very expensive drugs administered during a hospital stay.[40,41] Certain serious, costly or long-term illnesses, called ALD (Affections de Longue Durée), are eligible for 100% reimbursement of related healthcare expenditure by French NHI. The remaining medical fees constitute out-of-pocket expenses for the patient, who usually subscribe to mutual funds or private health insurance or are eligible for state funded free complementary healthcare (CMUc [Couverture Médicale Universelle complémentaire]).[40] The CMUc allocation is provided to people with an annual income below 50% of the poverty threshold.

The French NHI comprises of several insurance schemes. Individuals and their children are covered by a scheme based on their occupational branch and remain covered by this scheme after retirement. The largest scheme is the General Scheme (RG [Régime Général]). Together with the Local Mutualist Sections (SLM [Sections Locales Mutualistes], eg, students' and

civil servants' health insurances) they cover about 87% of the population living in France.[42] There are no major differences between schemes in terms of health expenditure coverage.

9.4.1. The SNIIRAM – The NHI Database

The SNIIRAM contains individual-level data for outpatient and private healthcare facilities health expenditure billing and reimbursement purposes (**DCIR** [Outpatient healthcare consumption data]), linked to the hospital discharge database (**PMSI** [hospital discharge database]) with a unique, anonymous identifier, the Social Security Number (NIR). Therefore, it encompasses anonymous, individual-level data for all healthcare claims for more than 99% of the population residing in France, regardless of the insurance scheme, ie, close to 65 million people.[42-45]

9.4.2. The DCIR - Outpatient Healthcare Consumption Data

The DCIR contains nationwide individual-level healthcare claims for all outpatient and private healthcare facilities care. It holds both administrative and medical data.

The main collected administrative information are (*sensitive variables):

- Socio-demographic data: date of birth*, gender, CMUc, department and commune* of residence, geographical social disadvantage index and date of death;*
- Health insurance scheme of the beneficiary;
- Sick leave and disability pension;
- Primary and secondary care (number of visits, date of care*) provided by general practitioners (GPs), other specialists, nurses, physiotherapists etc.;
- Outpatient medical and surgical procedures (number of procedures and date of care*), coded according to the French medical classification for clinical procedure [CCAM (Classification Commune des Actes Médicaux)];
- Laboratory tests (number of procedures and date of care*), coded according to the nomenclature of laboratory tests [NABM (Nomenclature des Actes de Biologie Médicale)].
- Drugs, coded according to Anatomical Therapeutic Chemical (ATC) classification system (number of dispensed drugs and date of care*):
    - for drugs delivered by retail pharmacies, coded according to the pack identifier code (CIP [Code Identifiant de Présentation]);
    - for drugs dispensed by hospital pharmacies, coded according to common classification of dispensable units (UCD [Unité Commune de Dispensation]);

- Medical devices (number of dispensed devices and date of care*), coded according to the list of products and services (LPP [Liste des Produits et Prestations]);

- ALD, coded according to the 10th revision of the International Classification of Diseases and Related Health Problems (ICD-10);

- Transportation expenditure.

Although the DCIR does not offer the rational for the healthcare nor a diagnosis for each claim, it does include the name of the ALD.

9.4.3. The PMSI – The Hospital Discharge Database

The PMSI is a database with medical and administrative data managed by the Technical Agency for Hospitalization Information (ATIH [Agence Technique de l'Information sur l'Hospitalisation]) which collects all the activities performed by the four types of healthcare facilities: the Medicine-Surgery-Obstetric (MCO [Médecine-Chirurgie-Obstétrique]), rehabilitations units (SSR [Soins de Suite et de Réadaptation]), homecare units (HAD [Hospitalisation à Domicile]), and psychiatric institutions (Psy). This database is accessible directly or through the SNIIRAM.

Since 2007, for the General Scheme, 97% of the outpatient and hospitalization data are matched. The PMSI-MCO informs about hospital stays (both in public and private facilities) with the Anonymous Hospital Discharge Summary (RSA [Résumé de Sortie Anonyme]) and additional medical data (with the date of care*), including:

- Reason for hospitalization with the principal diagnosis (DP) or the related diagnosis , coded in ICD-10;

- Medical procedures (coded with CCAM) performed during the hospital stay;

- Comorbidities which required to be cared for during the hospital stay (disease activity score [DAS] or DP);

- Expensive drugs and implantable medical devices;

- Other medical information such as organ harvesting, transfusion, transplant, intensive care units stays.

Sensitive variables

The SNDS databases contain sensitive variables (city of residence, month and year of birth, date of care, date of death). Cross-linkage of these four variables is forbidden, as well as cross-linkage of "city of residence" with some medical information (ALD reference, ALD ICD-10 code or PMSI diagnosis, hospital stay group [GHS] number).

Using one of these sensitive variables triggers the loss of precise information on the other three variables. For instance, access to the exact date of health care implies that only the department of residence, year of birth and year of death will be available.

9.5. Study Size

The SNDS captures all outpatient healthcare reimbursements and all hospital stays, in both public and private institutions, in France.[42] Through open data, tofacitinib's target population in France reaches around 30,000 patients for RA, 12,000 patients for PsA, and 5,000 patients for UC. Power calculation and thus minimum detectable hazard ratios for events of interest (malignancy excluding NMSC, serious infections, VTE, and acute cardiovascular events) have been calculated for each pathology based on incidence rates per 100 patients-years (IR per 100 PY) from published literature, when available. Minimum detectable HRs have been computed with an α of 5.0%, a power of 80% and the sample sizes for each pathology from open data cited before with a ratio of 20% treated by tofacitinib and 80% treated by TNFi.

Details for other outcomes/pathologies are provided in Table 4.

Table 4. Minimum Detectable Hazard Ratios and Corresponding Incidence Rates per 100 Patients-years.

IMID	Outcome	IR per 100 PY for TNFi patients	Minimum detectable hazard ratio	Sources (IR for TNFi patients)
RA N=30,000	Malignancy (excluding NMSC)	0.77	1.57	46
	Serious infections	1.63	1.35	47
	Acute cardiovascular event	0.62	1.65	46
	VTE	1.71	1.34	48
PsA N=12,000	Malignancy (excluding NMSC)	0.41	2.81	46
	Serious infections	2.70	1.47	33
	Acute cardiovascular event	0.54	2.45	46
	VTE	0.31	3.59	46
UC N=5,000	Malignancy (excluding NMSC)			
	Serious infections	2.20	1.91	49
	Acute cardiovascular event	2.37	1.87	50
	VTE	1.36	2.23	51
Power = 0.8, α = 0.05, 20% of patients treated by tofacitinib and 80% of patients treated by TNFi.				

9.6. Data Management

After Commission Nationale de l'Informatique et des Libertés (CNIL) approval, data will made available to the authorized HEVA team on the secure SNDS portal or on HEVA's private cloud computing (in agreement with CNAM - see description in ANNEX 3. HEVA'S HEALTH DATA HOST PRIVATE CLOUD COMPUTING). HEVA's private cloud computing complies with the regulatory obligations of the 22 March 2017 decree regarding the SNDS security framework.

Confidentiality of patient records will be maintained at all times. All study reports will contain only aggregated data and will not identify individual patients or physicians. At no time during the study will the data controller receive patient identifying information.

9.7. Data Analysis

9.7.1. General Considerations

Detailed methodology for summary and statistical analyses of data collected in this study will be documented in a Statistical Analysis Plan (SAP), which will be dated, filed and maintained by the sponsor. The SAP may modify the plans outlined in the protocol; any major modifications of primary endpoint definitions or their analyses would be reflected in a protocol amendment.

Analyses will be performed separately per type of IMID, RA, PsA, UC. Depending on number of patients, RA and PsA patients can be grouped together.

Continuous, quantitative, variable summaries will include: the number of patients (N) (with non-missing values), mean, standard deviation, median, minimum and maximum, 1st, and 3rd quartiles.

Categorical, qualitative, variables summaries will include the frequency and percentage of patients per category. The denominator for percentage calculations will be based on the number of observed data unless otherwise specified.

All applicable statistical tests will be two-sided and will be performed using a 5% significance level. All confidence intervals presented will be 95% and two-sided. The confidence interval for prevalence and incidence will be calculated using the normal approximation or the exact Poisson distribution, as appropriate.

At the beginning of the analysis, a diagram will summarize the conduct of the study: eligible subjects, inclusions and follow-up of the protocol.

For both primary and secondary safety objectives, a Cox model will be used. In case of sufficient power, a marginal structural Cox proportional hazard regression models (MSM) will also be implemented.

The statistical analyses will be performed with SAS® software (version 9.4 or later) and Python (version 3.7 or later).

9.7.2. Baseline Patients Characteristics

Refer to Section 9.3 for the list of variables.

Patient characteristics will be summarized using descriptive statistics.

9.7.3. Primary Objective

9.7.3.1. Cox Model (Marginal structural Cox proportional hazard regression models (MSM) if sufficient power)

- To compare the risk of malignancies, excluding non-melanoma skin cancer (NMSC) between patients initiating a treatment by tofacitinib and patients initiating a treatment by TNFi or vedolizumab* (separately for UC, RA, and PsA patients);

- To compare the risk of acute cardiovascular events between patients initiating a treatment by tofacitinib and patients initiating a treatment by TNFi* or vedolizumab (separately for UC, RA, and PsA patients)

9.7.4. Secondary Objectives

9.7.4.1. Cox Model (Marginal Structural Cox Proportional Hazard Regression Models (MSM) if Sufficient Power)

- To compare the risk of malignancies, excluding non-melanoma skin cancer (NMSC) between patients initiating a treatment by tofacitinib and patients initiating a treatment by TNFi or vedolizumab* (separately for UC, RA, and PsA patients and according to age [below and above 65 years old]).

- To compare the risk of acute cardiovascular events between patients initiating a treatment by tofacitinib and patients initiating a treatment by TNFi* or vedolizumab (separately for UC, RA, and PsA patients and according to age [below and above 65 years old]).

- To compare the risk of serious infections between patients initiating a treatment by tofacitinib and patients initiating a treatment by TNFi* or vedolizumab (separately for UC, RA, and PsA patients and according to age [below and above 65 years old]).

- To compare the risk of thromboembolic events between patients initiating a treatment by tofacitinib and patients initiating a treatment by TNFi* or vedolizumab (separately for UC, RA, and PsA patients and according to age [below and above 65 years old]).

Cox model:

In order to compare the risk of the different events of interest between tofacitinib treatment and TNFi treatment, a Cox regression model will be implemented on the statistical unit of per patient, to provide an estimation of association between tofacitinib and TNFi and the rate of events of interest, presented as Hazard Ratios. The Cox model models time to the event of interest and will explain the hazard of having the event of interest for each patient. The hazard ratio can be defined as, the ratio of (risk of outcome in one treatment group)/(risk of outcome in another treatment group), occurring at a given interval of time. A hazard ratio of

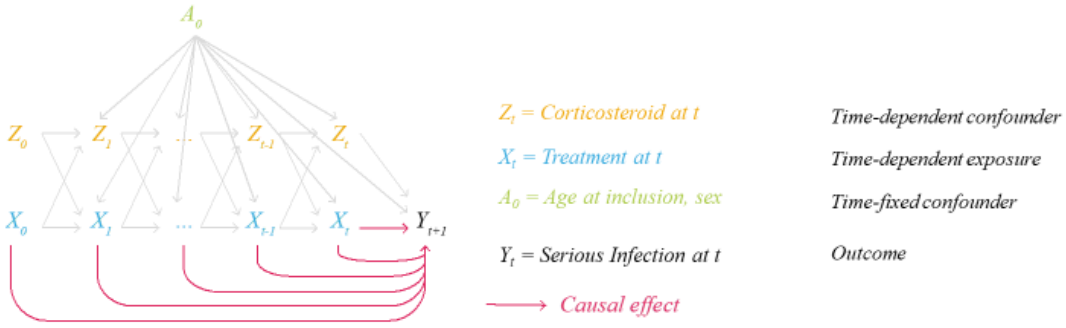
1 means lack of association, a hazard ratio greater than 1 suggests an increased risk, and a hazard ratio below 1 suggests a smaller risk.

MSM:

For comparative purposes, Cox proportional hazard regression models will be used to compare the risks of adverse event of interest during the follow-up between with each treatment line (TNFi) compared with tofacitinib. This model allows to consider the time of the follow-up of each patients as well as censoring of patients who have not experienced a serious infection and to quantify the difference between treatment lines with a hazard ratio (HR) and the corresponding 95% confidence. The model will account for potential confounders such as patient characteristics and known risk factors. The HR will be interpreted as risk ratios in the patients treated by tofacitinib vs TNFi.

During follow-up, patients could be exposed successively to different treatment sequences and could therefore contribute to more than 1 group of drug exposure. If the treatment variable is considered as a time-fixed covariate (known as an intention to treat analysis), when a patient changed treatment regimes, a patient would be considered with their original allocation and hence continue to contribute to the risk associated with that treatment regime, and thus introduce bias to the estimates and reduce power to detect a true difference in risk associated with treatments. To allow patients to switch treatments from their treatment regime at inclusion, the covariate of treatment line (tofacitinib, TNFi) will be treated as a time-varying covariate in the Cox regression model. Corticoids are hypothesized to be associated with treatment allocation as well as the risk of serious infections and hence are hypothesized to be time-dependent confounders. In the presence of time-dependent covariates (Corticosteroids) that might be associated with both exposure (treatment regimens) and outcomes (serious infections) and could also be affected by past exposure to treatment regimes, marginal structural Cox proportional hazard regression models (MSM) adjusted for the time-fixed and time-varying covariates can be used (Figure 7).[52] Briefly, this entails modeling the hazard ratio of the relationship of treatment and serious infections, with weights to account for the time-dependent confounding (referred to as the inverse probability of treatment weights) and censoring (the inverse probability of censoring weights). MSM models ascertains temporality of the relationships (causal modeling).

Figure 7. Directed Acyclic Graph (DAG) Describing Time-dependent Confounding of the Relationship of Treatment and Risk of Serious Infections.



9.7.4.2. Kaplan-Meier Estimation Method

- To estimate the risk of malignancy (excluding NMSC), serious infections, acute cardiovascular and thromboembolic events and describe the time to each first event among patients treated by tofacitinib and patients treated by TNFi.

- To describe tofacitinib treatment duration (drug survival) and persistence during the follow-up.

To describe the time to the first event of each outcome, a Kaplan-Meier estimation method will be used in order to quantify the median, mean, q1 and q3 treatment duration by treatment (tofacitinib and TNFi). The time scale will be the time since the treatment start.

For the subset of patients prescribed tofacitinib, a Kaplan-Meier estimation method will be used in order to quantify the median, mean, q1 and q3 tofacitinib treatment duration. The time scale will be the time since the treatment start.

To address the persistence of tofacitinib, a description of the percentage of treatment discontinuation and treatment switch since the start of each treatment will be provided at 1 year since the inclusion will be provided.

9.7.4.3. Machine Learning Profiling

To describe the risk factors/patients profiles related to the occurrence of each safety and effectiveness event of interest among patients treated by tofacitinib.

To determine if individual-like characteristics are linked to the occurrence of each of the events of interest, we propose to use a profiling analysis. It will result in the identification of the patient profiles the most likely to overexpress this outcome. The objective will be to search for combinations of patient characteristics.

The discovery of these homogeneous groups of patients will be made possible with the use of supervised machine learning methods (algorithm based on decision trees). In the end, this profiling analysis by Machine Learning will produce the following three results:

- The list of identified profiles (including number of patients and their characteristics);
- Histogram of the contribution of the explanatory variables to the target variable;
- The comparative table of each profile vs the rest of the cohort in terms of outcomes;
- The main advantage of this approach is to identify subgroups of patients the most at risk. Each subgroup is considered a profile (= a combination of several variables, ex: women over 65 years old, treated by tofacitinib 10 mg twice daily, with one cardiovascular risk factor). Details of the profiling analysis technique are presented in ANNEX 12. COHORT PROFILING ANALYSIS WITH MACHINE LEARNING.

9.7.4.4. Descriptive statistics

- To describe characteristics of patients initiating a treatment by tofacitinib, TNFi* or vedolizumab.

9.7.4.5. Competing Risk Model – Fine & Gray

To compare the effectiveness of tofacitinib vs TNFi. Competing risks of occurrence of hospitalizations and surgeries related to complications, and death will be assessed.

9.8. Quality Control

The data for this study will be retrieved from the SNIIRAM provided by CNAM. The validity of extracted dataset is ensured by CNAM. The designated contract research organization (CRO), HEVA will perform the analysis according to the contract agreement.

The statistical analyses will be programmed by a statistician and checked by another statistician. All variables used in the analyses will be described in order to evaluate the validity of derived variables.

All SAS programs are archived by HEVA.

9.9. Limitations of the Research Methods

This study has limitations that must be considered when interpreting the results. Some of the study limitations are related to the use of claims data. First, only treatments and procedures covered by the NHI are recorded. Over-the-counter drugs or drugs prescribed but not reimbursed, procedures not reimbursed [and their costs] are not considered in the study. Besides, we only have access to the dispensing data. We therefore have to assume that the patient is taking the drug, and that he/she does so on dispensing day. However, this assumption is reasonable and all the more likely that the patients has recurring dispensing for the dame drug. Additionally, the database does not contain medical data such as laboratory tests results or imaging procedures results. For this reason, we are using algorithms to identify patients with rheumatoid arthritis, psoriatic arthritis, and ulcerative colitis and outcomes. These algorithms will be developed with the scientific committee.

9.10. Other Aspects

Not applicable (NA).

10. PROTECTION OF HUMAN SUBJECTS

10.1. Patient information

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

10.2. Patient Consent

As this study involves 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.

Re-use of secondary health data, including from big data, without patient consent is allowed provided that the study has a public interest and that patients' rights and freedom are protected. These protections imply technical and organizational measures.[53]

In accordance with the regulations in force, the study protocol will be submitted to the committee for research, studies and evaluations in the field of health (Comité éthique et scientifique pour les recherches, les études et les évaluations dans le domaine de la santé, CESREES) and for authorization to the French data protection authority (Commission Nationale de l'Informatique et des Libertés, CNIL).

HEVA is also committed to comply with the standard setting the criteria of confidentiality, expertise and independence for research laboratories and consulting firms provided for by the decree of 17 July 2017 (RERC171002).

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

Due to its retrospective scheme based on secondary data collection, the study is not subject to the obligation of review by an ethics committee according to national legislation (Article R1121-1 of the Public Health Code).

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). Public Policy Committee, International Society of Pharmacoepidemiology. Pharmacoepidemiology and Drug Safety 2016; 25:2-10.

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

12. PLANS FOR DISSEMINATING AND COMMUNICATING STUDY RESULTS

In accordance with the regulations in force, the results of the study will be sent to the Health Data Hub (HDH) for information. The results will be used for discussions with the authority. The data controller also commits to publish the study results within one year of the submission of the final study report.

In the event of any prohibition or restriction imposed (eg, 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. Wilks AF. Two putative protein-tyrosine kinases identified by application of the polymerase chain reaction. Proc Natl Acad Sci. 1 mars 1989;86(5):1603-7.

2. Commission de la Transparence. Avis CT - Xeljanz 5 mg [Internet]. Haute Autorité de Santé; 2017 sept [cité 10 déc 2020]. Disponible sur: https://www.has-sante.fr/upload/docs/evamed/CT-16198_XELJANZ_PIC_INS_Avis2_CT16198.pdf.

3. Commission de la Transparence. Avis CT - Xeljanz 5 mg [Internet]. Haute Autorité de Santé; 2018 mai [cité 10 déc 2020]. Disponible sur: https://www.has-sante.fr/upload/docs/evamed/CT-17188_XELJANZ_PIC_EIT_RP_Avis2_CT17188.pdf.

4. Commission de la Transparence. Avis CT - Xeljanz 5 et 10 mg [Internet]. Haute Autorité de Santé; 2019 mars [cité 10 déc 2020]. Disponible sur: https://www.has-sante.fr/upload/docs/evamed/CT-17229_XELJANZ_RCH_PIC_INS_Avis3_CT17198&17229.pdf.

5. Pfizer. Xeljanz : EPAR - Product Information [Internet]. 2021 janv [cité 10 mai 2021]. Disponible sur: https://www.ema.europa.eu/en/documents/product-information/xeljanz-epar-product-information_en.pdf.

6. Olivera PA, Lasa JS, Bonovas S, Danese S, Peyrin-Biroulet L. Safety of Janus Kinase Inhibitors in Patients With Inflammatory Bowel Diseases or Other Immune-mediated Diseases: A Systematic Review and Meta-Analysis. Gastroenterology. mai 2020;158(6):1554-1573.e12.

7. EMA. Xeljanz: Scientific conclusions [Internet]. 2020 juin [cité 9 juill 2020]. Disponible sur: https://www.ema.europa.eu/en/documents/scientific-conclusion/xeljanz-epar-scientific-conclusions_en.pdf.

8. Pfizer Shares Co-Primary Endpoint Results from Post-Marketing Required Safety Study of XELJANZ® (tofacitinib) in Subjects with Rheumatoid Arthritis (RA) | pfpfizeruscom [Internet]. [cité 10 mai 2021]. Disponible sur: https://www.pfizer.com/news/press-release/press-release-detail/pfizer-shares-co-primary-endpoint-results-post-marketing.

9. Cohen SB, Tanaka Y, Mariette X, Curtis JR, Lee EB, Nash P, et al. Long-term safety of tofacitinib up to 9.5 years: a comprehensive integrated analysis of the rheumatoid arthritis clinical development programme. RMD Open. oct 2020;6(3):e001395.

10. Nash P, Coates LC, Kivitz AJ, Mease PJ, Gladman DD, Covarrubias-Cobos JA, et al. Safety and Efficacy of Tofacitinib in Patients with Active Psoriatic Arthritis: Interim Analysis of OPAL Balance, an Open-Label, Long-Term Extension Study. Rheumatol Ther. sept 2020;7(3):553-80.

11. Sandborn WJ, Panés J, D'Haens GR, Sands BE, Su C, Moscariello M, et al. Safety of Tofacitinib for Treatment of Ulcerative Colitis, Based on 4.4 Years of Data From Global Clinical Trials. Clin Gastroenterol Hepatol. juill 2019;17(8):1541-50.

12. Desai RJ, Pawar A, Weinblatt ME, Kim SC. Comparative Risk of Venous Thromboembolism in Rheumatoid Arthritis Patients Receiving Tofacitinib Versus Those Receiving Tumor Necrosis Factor Inhibitors: An Observational Cohort Study. Arthritis Rheumatol. juin 2019;71(6):892-900.

13. Kremer JM, Bingham CO, Cappelli LC, Greenberg JD, Madsen AM, Geier J, et al. Postapproval Comparative Safety Study of Tofacitinib and Biological Disease-Modifying Antirheumatic Drugs: 5-Year Results from a United States-Based Rheumatoid Arthritis Registry. ACR Open Rheumatol. mars 2021;3(3):173-84.

14. Simon TA, Thompson A, Gandhi KK, Hochberg MC, Suissa S. Incidence of malignancy in adult patients with rheumatoid arthritis: a meta-analysis. Arthritis Res Ther. déc 2015;17(1):212.

15. Pfizer. Phase 3B/4 randomized safety endpoint study of 2 doses of tofacitinib in comparaison to a Tumor Necrosis Factor (TNF) inhibitor in subjects with rheumatoid arthritis [Internet]. clinicaltrials.gov; 2021 févr [cité 4 févr 2021]. Report No.: NCT02092467. Disponible sur: https://clinicaltrials.gov/ct2/show/NCT02092467.

16. Smolen JS, Landewé RBM, Bijlsma JWJ, Burmester GR, Dougados M, Kerschbaumer A, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2019 update. Ann Rheum Dis. 22 janv 2020;annrheumdis-2019-216655.

17. Gossec L, Baraliakos X, Kerschbaumer A, de Wit M, McInnes I, Dougados M, et al. EULAR recommendations for the management of psoriatic arthritis with pharmacological therapies: 2019 update. Ann Rheum Dis. juin 2020;79(6):700-12.

18. Rubin DT, Ananthakrishnan AN, Siegel CA, Sauer BG, Long MD. ACG Clinical Guideline: Ulcerative Colitis in Adults. Am J Gastroenterol. mars 2019;114(3):384-413.

19. Pfizer. Xeljanz: Summary of product characteristics [Internet]. 2020. Disponible sur: https://www.ema.europa.eu/en/documents/product-information/xeljanz-epar-product-information_en.pdf.

20. Bonovas S, Lytras T, Nikolopoulos G, Peyrin-Biroulet L, Danese S. Systematic review with network meta-analysis: comparative assessment of tofacitinib and biological therapies for moderate-to-severe ulcerative colitis. Aliment Pharmacol Ther. févr 2018;47(4):454-65.

21. Singh S, Murad MH, Fumery M, Dulai PS, Sandborn WJ. First- and Second-Line Pharmacotherapies for Patients With Moderate to Severely Active Ulcerative Colitis: An Updated Network Meta-Analysis. Clin Gastroenterol Hepatol. janv 2020;S1542356520300446.

22. EMA. Direct Healthcare Professional Communication (DHCP) - Xeljanz (tofacitini) initial clinical trial results [Internet]. 2021 [cité 10 mai 2021]. Disponible sur: https://www.ema.europa.eu/en/documents/dhpc/direct-healthcare-professional-communication-dhpc-xeljanz-tofacitinib-initial-clinical-trial-results_en.pdf.

23. Bechman K, Subesinghe S, Norton S, Atzeni F, Galli M, Cope AP, et al. A systematic review and meta-analysis of infection risk with small molecule JAK inhibitors in rheumatoid arthritis. Rheumatology. 1 oct 2019;58(10):1755-66.

24. Winthrop KL, Melmed GY, Vermeire S, Long MD, Chan G, Pedersen RD, et al. Herpes Zoster Infection in Patients With Ulcerative Colitis Receiving Tofacitinib. Inflamm Bowel Dis. 15 sept 2018;24(10):2258-65.

25. Wollenhaupt J, Lee E-B, Curtis JR, Silverfield J, Terry K, Soma K, et al. Safety and efficacy of tofacitinib for up to 9.5 years in the treatment of rheumatoid arthritis: final results of a global, open-label, long-term extension study. Arthritis Res Ther. déc 2019;21(1):89.

26. Curtis JR, Xie F, Yang S, Bernatsky S, Chen L, Yun H, et al. Risk for Herpes Zoster in Tofacitinib-Treated Rheumatoid Arthritis Patients With and Without Concomitant Methotrexate and Glucocorticoids. Arthritis Care Res. sept 2019;71(9):1249-54.

27. Kirchgesner J, Lemaitre M, Carrat F, Zureik M, Carbonnel F, Dray-Spira R. Risk of Serious and Opportunistic Infections Associated With Treatment of Inflammatory Bowel Diseases. Gastroenterology. août 2018;155(2):337-346.e10.

28. Kim SC, Schneeweiss S, Liu J, Solomon DH. The risk of venous thromboembolism in patients with rheumatoid arthritis. Arthritis Care Res. mai 2013;NA-NA.

29. Zezos P, Kouklakis G, Saibil F. Inflammatory bowel disease and thromboembolism. World J Gastroenterol. 14 oct 2014;20(38):13863-78.

30. Mease P, Charles-Schoeman C, Cohen S, Fallon L, Woolcott J, Yun H, et al. Incidence of venous and arterial thromboembolic events reported in the tofacitinib rheumatoid arthritis, psoriasis and psoriatic arthritis development programmes and from real-world data. Ann Rheum Dis. nov 2020;79(11):1400-13.

31. EMA. Meeting highlights from the Pharmacovigilance Risk Assessment Committee (PRAC) 8-11 March 2021 [Internet]. European Medicines Agency. 2021 [cité 10 mai 2021]. Disponible sur: https://www.ema.europa.eu/en/news/meeting-highlights-pharmacovigilance-risk-assessment-committee-prac-8-11-march-2021.

32. Kirchgesner J, Lemaitre M, Carrat F, Zureik M, Carbonnel F, Dray-Spira R. Risk of Serious and Opportunistic Infections Associated With Treatment of Inflammatory Bowel Diseases. Gastroenterology. août 2018;155(2):337-346.e10.

33. Li X, Andersen KM, Chang H-Y, Curtis JR, Alexander GC. Comparative risk of serious infections among real-world users of biologics for psoriasis or psoriatic arthritis. Ann Rheum Dis. févr 2020;79(2):285-91.

34. Légifrance. Article R5123-2 - Code de la santé publique - Légifrance [Internet]. [cité 1 déc 2020]. Disponible sur: https://www.legifrance.gouv.fr/codes/article_lc/LEGIARTI000006915040

35. Kirchgesner J, Nyboe Andersen N, Carrat F, Jess T, Beaugerie L. Risk of acute arterial events associated with treatment of inflammatory bowel diseases: nationwide French cohort study. Gut. mai 2020;69(5):852-8.

36. Sahli L, Lapeyre-Mestre M, Derumeaux H, Moulis G. Positive predictive values of selected hospital discharge diagnoses to identify infections responsible for hospitalization in the French national hospital database. Pharmacoepidemiol Drug Saf. 2016;25(7):785-9.

37. Meyer A, Rudant J, Drouin J, Weill A, Carbonnel F, Coste J. Effectiveness and Safety of Reference Infliximab and Biosimilar in Crohn Disease: A French Equivalence Study. Ann Intern Med. 15 janv 2019;170(2):99.

38. Global Burden of Disease Health Financing Collaborator Network. Evolution and patterns of global health financing 1995-2014: development assistance for health, and government, prepaid private, and out-of-pocket health spending in 184 countries. Lancet Lond Engl. 20 mai 2017;389(10083):1981-2004.

39. Centre des liaisons européennes et internationales de sécurité sociale, République Française. The French social security system - Sickness, maternity, paternity, disability and death branch [Internet]. 2020 [cité 19 nov 2020]. Disponible sur: https://www.cleiss.fr/docs/regimes/regime_france/an_1.html.

40. Chevreul K, Berg Brigham K, Durand-Zaleski I, Hernandez-Quevedo C. France: Health System Review. Health Syst Transit. 2015;17(3):1-218, xvii.

41. Tableaux récapitulatifs des taux de remboursement [Internet]. [cité 19 nov 2020]. Disponible sur: https://www.ameli.fr/assure/remboursements/rembourse/tableau-recapitulatif-taux-remboursement/tableau-recapitulatif-taux-remboursement.

42. Tuppin P, Rudant J, Constantinou P, Gastaldi-Ménager C, Rachas A, de Roquefeuil L, et al. Value of a national administrative database to guide public decisions: From the système national d'information interrégimes de l'Assurance Maladie (SNIIRAM) to the système national des données de santé (SNDS) in France. Rev DÉpidémiologie Santé Publique. oct 2017;65:S149-67.

43. Bezin J, Duong M, Lassalle R, Droz C, Pariente A, Blin P, et al. The national healthcare system claims databases in France, SNIIRAM and EGB: Powerful tools for pharmacoepidemiology. Pharmacoepidemiol Drug Saf. août 2017;26(8):954-62.

44. Moulis G, Lapeyre-Mestre M, Palmaro A, Pugnet G, Montastruc J-L, Sailler L. French health insurance databases: What interest for medical research? Rev Med Interne. juin 2015;36(6):411-7.

45. Tuppin P, de Roquefeuil L, Weill A, Ricordeau P, Merlière Y. French national health insurance information system and the permanent beneficiaries sample. Rev Epidemiol Sante Publique. août 2010;58(4):286-90.

46. Curtis JR, Mariette X, Gaujoux-Viala C, Blauvelt A, Kvien TK, Sandborn WJ, et al. Long-term safety of certolizumab pegol in rheumatoid arthritis, axial spondyloarthritis, psoriatic arthritis, psoriasis and Crohn's disease: a pooled analysis of 11 317 patients across clinical trials. RMD Open. 2019;5(1):e000942.

47. Winthrop KL, Citera G, Gold D, Henrohn D, Connell CA, Shapiro AB, et al. Age-based (<65 vs ≥65 years) incidence of infections and serious infections with tofacitinib versus biological DMARDs in rheumatoid arthritis clinical trials and the US Corrona RA registry. Ann Rheum Dis. 12 oct 2020;

48. Desai RJ, Pawar A, Weinblatt ME, Kim SC. Comparative Risk of Venous Thromboembolism in Rheumatoid Arthritis Patients Receiving Tofacitinib Versus Those Receiving Tumor Necrosis Factor Inhibitors: An Observational Cohort Study. Arthritis Rheumatol. juin 2019;71(6):892-900.

49. Sands BE, Peyrin-Biroulet L, Loftus EV, Danese S, Colombel J-F, Törüner M, et al. Vedolizumab versus Adalimumab for Moderate-to-Severe Ulcerative Colitis. N Engl J Med. 26 2019;381(13):1215-26.

50. Lewis JD, Scott FI, Brensinger CM, Roy JA, Osterman MT, Mamtani R, et al. Increased Mortality Rates With Prolonged Corticosteroid Therapy When Compared With Antitumor Necrosis Factor-α-Directed Therapy for Inflammatory Bowel Disease: Am J Gastroenterol. mars 2018;113(3):405-17.

51. Desai RJ, Gagne JJ, Lii J, Liu J, Friedman S, Kim SC. Comparative risk of incident venous thromboembolism in patients with inflammatory bowel disease initiating tumour necrosis factor-α inhibitors or nonbiologic agents: a cohort study. CMAJ Can Med Assoc J J Assoc Medicale Can. 27 nov 2017;189(47):E1438-47.

52. Robins JM, Hernán MÁ, Brumback B. Marginal Structural Models and Causal Inference in Epidemiology: Epidemiology. sept 2000;11(5):550-60.

53. Cuggia M, Polton D, Wainrib G, Combes S, Health Data Hub. Mission de préfiguration [Internet]. Paris, France: Ministère des Solidarités et de la Santé; 2018 [cité 19 nov 2020]. Disponible sur: https://solidarites-sante.gouv.fr/IMG/pdf/181012_-_rapport_health_data_hub.pdf.

14. LIST OF TABLES

15. LIST OF FIGURES

ANNEX 1. LIST OF STAND ALONE DOCUMENTS

Number	Document reference number	Date	Title
1	Section 4	April 2021	Real Life Safety and effectiveness of tofacitinib in comparison to TNF Inhibitors using the French National Healthcare Database
2	ANNEX 2	June 2021	ENCePP Checklist for Study Protocols *Document Name:* A3921399 ENCePPChecklistforStudyProtocols 16 Jun 2021 *Version:* 1.0;CURRENT *Status:* Final *Location:* /Compounds/CP/CP-690550/Clinical/N-A/A3921399/Protocol http://gdms.pfizer.com/gdms/drl/objectId/090177e1975b78c3

ANNEX 2. ENCEPP CHECKLIST FOR STUDY PROTOCOLS

REFER TO ANNEX 1.

ANNEX 3. HEVA'S HEALTH DATA HOST PRIVATE CLOUD COMPUTING

After being a trailblazer by implementing a first version of a private cloud computing on its premises, HEVA wished to develop this infrastructure towards a new HEVA HDS (Health Data Host, *Hébergement des Données de santé*) private cloud computing. This new dedicated and HDS-certified cloud is hosted in France by the Online Virtual Hosting (OVH) company. This operational solution devoted solely to the hosting and processing of healthcare data is a secure cloud, and therefore ensures:

- The health data's security;
- The health data's integrity;
- The tracking of the access and actions performed on the data.

To support HEVA in this endeavor, the Certilience company has been chosen as a trusted third party and the managing director of the HEVA HDS cloud.

15.1. The Secure Cloud Computing's Components

The infrastructure is a secure cloud because its design and its deployment rely on

- A filtering safety equipment (Firewall Fortigate de Fortinet) allowing only declared IP addresses to get through (HEVA and Certilience network). It is the unique means of communication with the cloud and can filter the results of the databases' extracts (ie, preventing non-anonymous, non-aggregated health data from leaving the cloud).
- A strong authentication equipment (Forti-Authenticator) combined with a Token (FortiToken) for each cloud user.
- An equipment to secure access to the cloud and to the dedicated users' spaces (Bastion Wallix). This allows to record all actions performed by users; the records can be reviewed afterwards.

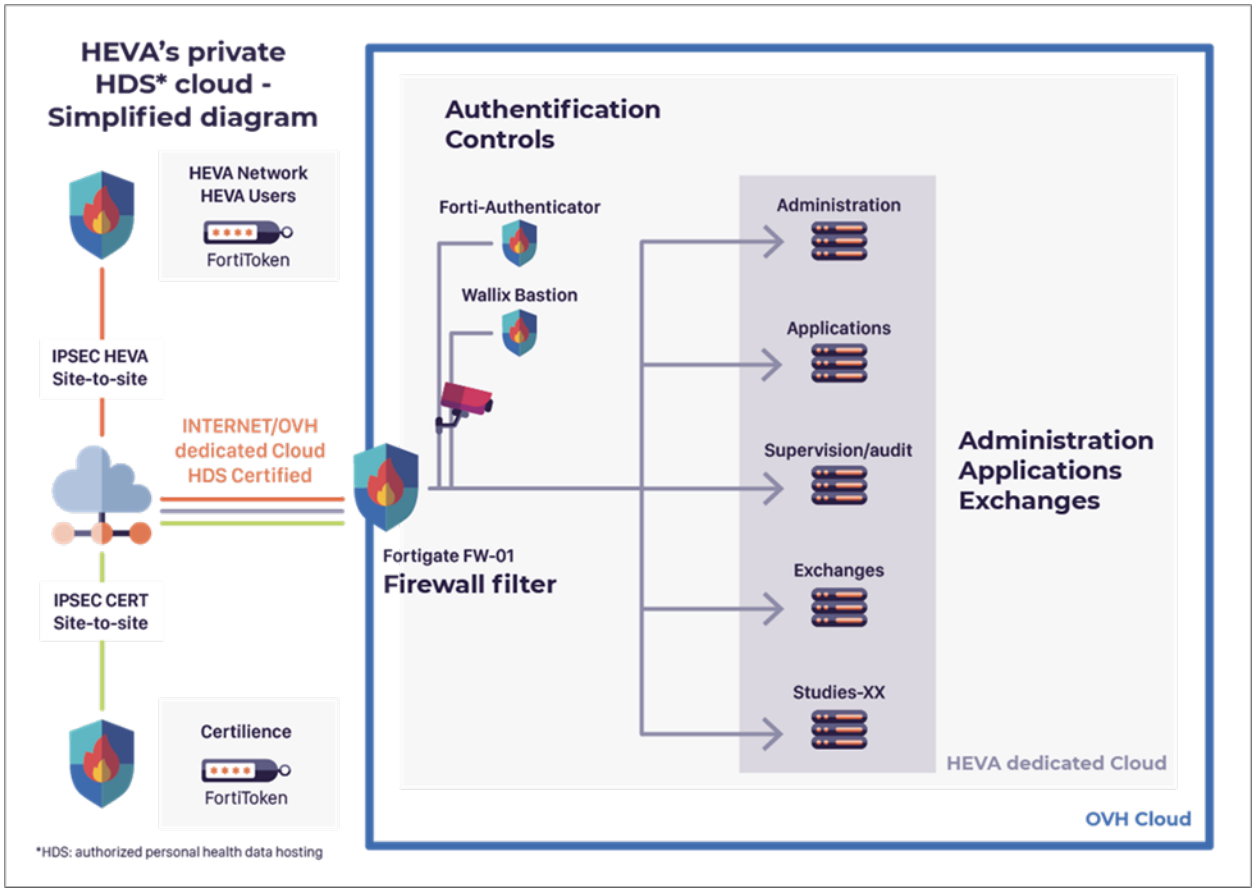
15.2. The Trusted Third Party Linked to the Private Cloud Computing

Our trusted third party, Certilience, plays three roles:

- It integrates the secure solution and advises on necessary evolutions;
- It administers and maintains the cloud computing through a dedicated access, allowing .
    - Its monitoring to ensure its continuous availability
    - Real-time access to all its traces: log-ins, modifications, openings, power shortages, etc.
- It acts as a middle-man for the import of health data provided by authorities (the National Health Insurance and the ATIH).

All actions carried out by Certilience can be fully controlled afterwards and obtained at HEVA's request.

15.3. Simplified Diagram of the HEVA HDS Private Cloud Computing



15.4. The Operational Functioning of the Private Cloud Computing

15.4.1. Entering Health Data into the Private Cloud Computing

Once the administrative and scientific processes have been validated, the health data producer (*Assurance Maladie* or ATIH) routes the health data through a secure process via the creation of a temporary encrypted channel, approved by itself, HEVA and the trusted third party. The trusted third party deposits the data, which is then decrypted by HEVA. All of these operations are recorded by the Bastion and sent to the data producer who can view it if necessary.

15.4.2. HEVA Employees

Employees in charge of health data (Data Managers, Data Scientists, Biostatisticians, and Pharmacoepidemiologists) are authorized to process SNDS data, are regularly reminded of data security procedures, are committed to respecting professional secrecy and the rules related to the use of SNDS data.

From a workstation provided and configured by HEVA and only from the HEVA network, they can connect to the HEVA HDS private cloud computing following this secured procedure:

- Entry of a login and password provided by the HEVA private cloud computing administrator,
- Entry of an additional password generated by an individual physical token (in connection with the Forti-Authenticator).

Once connected, users have the following data processing software at their disposal: Free Office Suite, PHP, SAS, Python, R. The data are stored on different encrypted virtual operating machines with regular encrypted backups.

Upon login, all user actions are recorded in a computerized event log, called a logbook. All health data processing is systematically recorded by the Bastion and can be later reviewed if necessary.

15.4.3. Data Output From the Private Cloud Computing

All the study results of the studies leaving the private could computing are anonymized and aggregated; they are systematically archived for possible future analyses. They pass through filtering equipment that analyzes the content and blocks them if necessary using predefined filtering rules known as DLP (Data Leak Prevention).

15.5. Conclusion

Compliant with the security repository and approved on 05/29/2020, the HEVA HDS private cloud computing makes it possible to systematically identify its various users, to securely host and work with health data, to compartmentalize the various SNDS studies and to monitor all interventions on the data.

ANNEX 4. EXCLUSION CRITERIA ICD-10 AND CCAM CODES (THE DEFINITIVE LIST WILL BE FINALIZED IN THE SAP)

Table 5. Exclusion Criteria ICD-10 Codes and CCAM Codes

Comorbidity	ICD-10 codes	Anatomical Therapeutic Chemical (ATC) Classification System Code	French Medical common Procedure Coding System (CCAM codes)
Ischemic heart disease	I20-I25, T820, T822,T823, T826, T827, Z45.0, Z86.71, Z95 (except Z95.8,Z95.9)	-	Coronary catheterization with arterial dilation; CABG
Cerebrovascular disease	I60-I66, G45, G460-G462, I670 , I671, I68, I69	-	Cerebral angiography with arterial dilation or embolization
Peripheral artery disease	I74	-	Angiography (excluding coronary catheterization and cerebral angiography) with arterial dilation; Peripheral artery bypass
Heart failure	I50	-	-
Atrial fibrillation	I48	-	-
Cardiomyopathies	I42-I43	-	-
Venous thromboembolism	I26, I80-I82, O22.3, O22.9, O87.1, O88.2		
Cancer	C00-C97	-	-
Congenital deficiency	D80-84	J06BA02, J06BA01	-
human immunodeficiency virus (HIV)	B20-24	J05AX07, J05AX08, J05AX09, J05AX10, J05AX11, J05AE (except J05AE12), J05AF (except J05AF10), J05AG, J05AR	-
Organ transplantation	T86.0–T86.4, T86.80–T86.82, T86.9, Z94.0–Z94.4, Z94.803, Z94.804, Z94.809, Z94.81, Z94.82, Z94.88, Z94.9		HNEA002,DZEA001, DZEA002, DZEA003, DZEA004, DZFA004, FELF009, GFEA001, GFEA002, GFEA003, GFEA004, GFEA005, GFEA006, GFEA007, HGEA002, HGEA004, HGEA005, HLEA001, HLEA002, HNEA900, JAEA003
CABG, Coronary Artery Bypass Graft Surgery			

ANNEX 5. COVARIATES ICD-10 AND CCAM CODES.

Table 6. Covariates ICD-10 and CCAM Codes

Comorbidity	ICD-10 codes	French Medical common Procedure Coding System (CCAM codes)
Cancer	C0-C9, E88.3, G53.3, G55.0, G63.1, G73.2, G94.1, J70.0, J70.1, K52.0, K62.7, L58.0, L58.1, L59.8, D63.0, L59.9, M36.0, M36.1, M90.6, M90.7, M96.2, M96.5, N30.4, O35.6, Z08, Z51.1, Z54.2, Z85	Chemotherapy (Z511) and radiotherapy (Z510)
Chronic pulmonary disease	J4–J7, J82–J84, J96.0, J96.1	
Chronic kidney disease	I120, I131, N18–N19, Y84.1, Z49	
Diabetes	E10-E14, M14.2 , M14.6, N08.3, H28.0, H36.0, G59.0, G63.2, G73.0, G99.0, I79.2	
Cirrhosis and portal hypertension	I85; I86.4; I98.2; I98.3; K70.0; K70.3-K70.4; K71.1; K71.7; K72; K74.4-K74.6; K76.6; K76.7;	
Obesity	E66	
Alcohol use disorder	E24.4, G31.2, G62.1, G72.1, I42.6, K29.2, K70, K86.0, Z50.2, Z71.4, Z12.1	
Smoking behavior	F17, Z71.6, Z72.0, T65.2	
Hypertension	I10-I13, I15	
Dyslipidemia	E78.0-E78.5	
Atherosclerosis	I70	
Dyslipidemia	E78.0-E78.5	
Ankylosing Spondylarthritis (AS).	M45	
UC-related hospitalization	K51, K56, K60, K61	
UC-related surgery Colectomy		HHFA002, HHFA004, HHFA005, HHFA006, HHFA008, HHFA009, HHFA010, HHFA014, HHFA017, HHFA018, HHFA021, HHFA022, HHFA023, HHFA024, HHFA026, HHFA028, HHFA029, HHFA030, HHFA031
Intestinal resection		HGCA005, HGCC015, HGFA003, HGFA004, HGFA005, HGFA007, HGFC014, HGFC016, HGFC021
Perineal surgery and minor digestive surgery		HKPA004, HKPA005, HKPA006, HKPA007, HKPA008, HGCA008, HGCC026, HGLA001, HHCA003, HHCC011, HPPA002, HPPC003, ZCJA002, ZCJA004
UC-related imaging tests Abdominal and pelvic CT		HHQH365, HHQK484, ZCQH001, ZCQH002, ZCQK004, ZCQK005
Abdominal and pelvic magnetic resonance imaging (MRI)		ZCQJ004, ZCQJ005, ZCQN001, ZCQN002
Gastrointestinal transit examination (transit radionuclide imaging)		HFQL001, HFQL002, HFQL003, HFQL004
RA-related hospitalization	M05 M06	

Comorbidity	ICD-10 codes	French Medical common Procedure Coding System (CCAM codes)
PsA- related hospitalization	L405, M07.0, M07.1, M07.2, M07.3	
RA or PsA imaging tests		
Superior limb radiography		MAQK001, MAQK002, MAQK003, MDQK001, MDQK002, MEQH001, MFQH001, MFQK001, MFQK002, MGQH001, MGQK001, MGQK003, MHQH001, MZQH001, MZQK001, MZQK003, MZQK004
Superior limb CT		MZQH002, MZQK002
Superior limb MRI		MZQJ001, MZQN001
Joint ultrasound		PBQM001, PBQM002, PBQM003, PBQM004
Neck and torso radiography		LDQK001, LDQK002, LDQK004, LDQK005, LEQK001, LEQK002, LFQK001, LFQK002, LGQK001, LHQH001, LHQH003, LHQH004, LHQH005, LHQK002, LHQK003, LHQK004, LHQK007
Neck and torso CT		LHQH002, LHQH006, LHQK001, LHQK005
Neck and torso MRI		LHQJ001, LHQJ002, LHQN001, LHQN002
Inferior limb radiography		NAQK007, NAQK015, NAQK023, NAQK049, NAQK071, NDQK001, NDQK002, NDQK003, NDQK004, NEQH002, NEQK010, NEQK012, NEQK035, NFQH001, NFQK001, NFQK002, NFQK003, NFQK004, NGQH001, NGQK001, NGQK002, NHQH001, NZQH002, NZQK001, NZQK003, NZQK005, NZQK006
Inferior limb CT		NZQH001, NZQH005, NZQK002
Inferior limb MRI		NZQJ001, NZQN001
Intra-articular injection		
Superior limb intra-articular injection		MZLB001, MZLH001, MZLH002
Neck and torso intra-articular injection		LHLB001, LHLH002, LHLH003
Superior limb intra-articular injection		NZLB001, NZLH001, NZLH002

ANNEX 6. MALIGNANCIES (EXCLUDING NMSC) ICD-10 CODES

Table 7. Malignancies (excluding NMSC) ICD-10 Codes.

Cancer type	ICD-10 code	Title
Active breast cancer	C50	Malignant neoplasm of breast
	D05	Carcinoma in situ of breast
Lymphoma	C81	Hodgkin lymphoma
	C82	Follicular lymphoma
	C83	Non-follicular lymphoma
	C84	Mature T/NK-cell lymphomas
	C85	Other and unspecified types of non-Hodgkin lymphoma
Active lung cancer	C33	Malignant neoplasm of trachea
	C34	Malignant neoplasm of bronchus and lung
	D021	Carcinoma in situ: Trachea
	D022	Carcinoma in situ: Bronchus and lung
Active melanoma	C43	Malignant melanoma of skin
	D03	Melanoma in situ
Other active cancer	C00	Malignant neoplasm of lip
	C01	Malignant neoplasm of base of tongue
	C02	Malignant neoplasm of other and unspecified parts of tongue
	C03	Malignant neoplasm of gum
	C04	Malignant neoplasm of floor of mouth
	C05	Malignant neoplasm of palate
	C06	Malignant neoplasm of other and unspecified parts of mouth
	C07	Malignant neoplasm of parotid gland
	C08	Malignant neoplasm of other and unspecified major salivary glands
	C09	Malignant neoplasm of tonsil
	C10	Malignant neoplasm of oropharynx
	C11	Malignant neoplasm of nasopharynx
	C12	Malignant neoplasm of piriform sinus
	C13	Malignant neoplasm of hypopharynx
	C14	Malignant neoplasm of other and ill-defined sites in the lip, oral cavity and pharynx
	C15	Malignant neoplasm of oesophagus
	C16	Malignant neoplasm of stomach
	C17	Malignant neoplasm of small intestine
	C18	Malignant neoplasm of colon
	C19	Malignant neoplasm of rectosigmoid junction
	C20	Malignant neoplasm of rectum
	C21	Malignant neoplasm of anus and anal canal
	C22	Malignant neoplasm of liver and intrahepatic bile ducts
	C23	Malignant neoplasm of gallbladder
	C24	Malignant neoplasm of other and unspecified parts of biliary tract
	C25	Malignant neoplasm of pancreas
	C26	Malignant neoplasm of other and ill-defined digestive organs
	C30	Malignant neoplasm of nasal cavity and middle ear
	C31	Malignant neoplasm of accessory sinuses
	C32	Malignant neoplasm of larynx
	C37	Malignant neoplasm of thymus
	C38	Malignant neoplasm of heart, mediastinum and pleura
	C39	Malignant neoplasm of other and ill-defined sites in the respiratory system and intrathoracic organs
	C40	Malignant neoplasm of bone and articular cartilage of limbs
	C41	Malignant neoplasm of bone and articular cartilage of other and unspecified sites
	C45	Mesothelioma

Cancer type	ICD-10 code	Title
	C46	Kaposi sarcoma
	C47	Malignant neoplasm of peripheral nerves and autonomic nervous system
	C48	Malignant neoplasm of retroperitoneum and peritoneum
	C49	Malignant neoplasm of other connective and soft tissue
	C51	Malignant neoplasm of vulva
	C52	Malignant neoplasm of vagina
	C53	Malignant neoplasm of cervix uteri
	C54	Malignant neoplasm of corpus uteri
	C55	Malignant neoplasm of uterus, part unspecified
	C56	Malignant neoplasm of ovary
	C57	Malignant neoplasm of other and unspecified female genital organs
	C58	Malignant neoplasm of placenta
	C60	Malignant neoplasm of penis
	C61	Malignant neoplasm of prostate
	C62	Malignant neoplasm of testis
	C63	Malignant neoplasm of other and unspecified male genital organs
	C64	Malignant neoplasm of kidney, except renal pelvis
	C65	Malignant neoplasm of renal pelvis
	C66	Malignant neoplasm of ureter
	C67	Malignant neoplasm of bladder
	C68	Malignant neoplasm of other and unspecified urinary organs
	C69	Malignant neoplasm of eye and adnexa
	C70	Malignant neoplasm of meninges
	C71	Malignant neoplasm of brain
	C72	Malignant neoplasm of spinal cord, cranial nerves and other parts of central nervous system
	C73	Malignant neoplasm of thyroid gland
	C74	Malignant neoplasm of adrenal gland
	C75	Malignant neoplasm of other endocrine glands and related structures
	C76	Malignant neoplasm of other and ill-defined sites
	C77	Secondary and unspecified malignant neoplasm of lymph nodes
	C78	Secondary malignant neoplasm of respiratory and digestive organs
	C79	Secondary malignant neoplasm of other and unspecified sites
	C80	Malignant neoplasm, without specification of site
	C88	Malignant immunoproliferative diseases
	C90	Multiple myeloma and malignant plasma cell neoplasms
	C91	Lymphoid leukaemia
	C92	Myeloid leukaemia
	C93	Monocytic leukaemia
	C94	Other leukaemias of specified cell type
	C95	Leukaemia of unspecified cell type
	C96	Other and unspecified malignant neoplasms of lymphoid, haematopoietic and related tissue
	C97	Malignant neoplasms of independent (primary) multiple sites
	D00	Carcinoma in situ of oral cavity, oesophagus and stomach
	D010	Carcinoma in situ: Colon
	D011	Carcinoma in situ: Rectosigmoid junction
	D012	Carcinoma in situ: Rectum
	D013	Carcinoma in situ: Anus and anal canal
	D014	Carcinoma in situ: Other and unspecified parts of intestine
	D015	Carcinoma in situ: Liver, gallbladder and bile ducts
	D017	Carcinoma in situ: Other specified digestive organs
	D019	Carcinoma in situ: Digestive organ, unspecifie
	D020	Carcinoma in situ: Larynx
	D023	Carcinoma in situ: Other parts of respiratory system

Cancer type	ICD-10 code	Title
	D024	Carcinoma in situ: Respiratory system, unspecified
	D06	Carcinoma in situ of cervix uteri
	D070	Carcinoma in situ: Endometrium
	D071	Carcinoma in situ: Vulva
	D072	Carcinoma in situ: Vagina
	D073	Carcinoma in situ: Other and unspecified female genital organs
	D074	Carcinoma in situ: Penis
	D075	Carcinoma in situ: Prostate
	D076	Carcinoma in situ: Other and unspecified male genital organs
	D09	Carcinoma in situ of other and unspecified sites

ANNEX 7. SERIOUS INFECTIONS ICD-10 CODES

Table 8. Serious Infections ICD-10 Codes

Subgroup of infection	Diagnoses	ICD-10 code
Pulmonary infections	Pneumonia	A48.1, B01.2, B05.2, B25.0, J12–J18, J10–J11
	Other acute lower respiratory infections	A37, A42.0, B39–B40, B44, B58.3, B59, B95.3, J20–J22, U04
	Abscessus pulmonis	J85
	Empyema pleurae	J86
Gastrointestinal infections	Intestinal infectious disease	A00–A08, K93.820
	Viral hepatitis	B15, B17, B25.1
	Cholangitis	K80–K810, K830, K87.00, B25.8
	Liver abscess	K750
	Infectious esophagitis	B00.8 (K23.80)
Skin and subcutaneous (sc) tissue infections	Erysipelas	A46
	Dermatophytosis and other superficial mycoses	B35–B36
	Cellulitis and abscess	L02–L03
	Herpes virus	B00.1–B00.2, B00.7, B00.9, B01.8–B01.9, B02.3–B02.9, B05.3–B05.9, B06.8–B06.9, B08–B09, A60
	Other local infections of skin, oral tissue, and subcutaneous tissue	A36.3, K11.3–K12.2, L00–L01, L04–L05, L08, L30.3, M72.6
Urinary tract infections	Nephritis	N10
	Acute prostatitis and prostate absces	N41.0, N41.2, N41.3
	Cystitis	N30.0
	Salpingitis and oophoritis	N70.0
	Endometritis	N71.0
	Cervicitis uteri	N72
	Syphilis	A50–A53, I98.0
	Gonorrhea	A54
	Chlamydia	A55–A56
	Orchitis and epididymitis	N45
	Other urinary tract infections	N39.0, N73.3, N77.1
ENT infections	Mastoiditis	H70
	Nasopharyngitis	A36.1
	Sinusitis	J01
	Pharyngitis	J02
	Pharyngeal, retropharyngeal, and parapharyngeal abscess	J36, J39.0–J39.1
	Tonsillitis	A36.0, J03
	Laryngitis and tracheitis	A36.2, J04–J05, J37
	Acute upper respiratory infections of multiple and unspecified sites	A36.8–A36.9, J06
	Infection of external ear and acute otitis media	H60.0–H60.3, H65.1–H65.2, H66, H68.0
Musculoskeletal infections	Infectious arthritis	M00–M01
	Infective myositis	M60.0
	Osteomyelitis	M86
Other infections	Infection of the eye	B00.5, B30, H00–H01, H03.1, H06.1, H10.5, H10.8, H13.1, H19.1–H19.2
	Infections in the nervous system	A32.1, A39, A80-A89, B00.3–B00.4, B01.0–B01.1, B02.0–B02.2, B05.0–B05.1, B06.0, G00-G02, G04–G07
	Infections of prosthetic devices, implants, and	T82.6–T82.7, T84.5–T84.7, T85.7

Subgroup of infection	Diagnoses	ICD-10 code
	grafts	
	Sepsis, systemic inflammatory response syndrome (SIRS) of infectious origin and septic shock	A32.7, A40–A41, R57.2, R65.0–R65.1
	Certain bacterial disease	A20–A28, A32, A34–A35, A38, A42–A44, A48.0, A48.2–A49.9, B95.1, B95.2, B95.4–B95.8, B96–B97
	Spirochaetal disease	A65–A69
	Rickettsiosis	A75–A79
	Viral infections	A90–A99, B25.2, B25.9, B26–B27; B33–B34
	Mycoses	B37–B49
	Protozoal diseases	B50–B57, B58.1–B58.2, B58.8–B58.9, B60–B83
	Unspecified infectious diseases	B99.9
	Acute infective pericarditis and endocarditis	I30.1, I33.0
	Mycobacterial infections	A15–A19, A31, K23.0, K67.3, K93.0, M01.1, M49.0, M90.0, N33.0, N74.0, N74.1

ANNEX 8. OPPORTUNISTIC INFECTIONS ICD-10 CODES

Table 9. Opportunistic Infections ICD-10 Codes.

Subgroup of infection	Diagnoses	ICD-10 code
Viral infections	Cytomegalovirus	B25, B27.1
	Herpes virus	B00–B02, A60.0
	Epstein-Barr virus	B27.0
	Progressive multifocal leukoencephalopathy	A81.2
	Acute viral hepatitis unspecified	B17.9
	Viral meningitis	G02.0
	Viral pneumoniae	J17.1
Mycobacterial infections	Mycobacterial infections	A15–A19, A31, K23.0, K67.3, K93.0, M01.1, M49.0, M90.0, N33.0, N74.0, N74.1
Bacterial infections	Bartonellosis	A44
	Legionnaires' disease	A48.1–A48.2
	Pneumonia and sepsis due to Streptococcus pneumoniae	A40.3, J13, B95.3
	Nocardiosis	A43
	Actinomycosis	A42
	Listeriosis	A32
	Salmonella infections	A02
Fungal infections	Candidiasis	B37
	Coccidioidomycosis	B38
	Histoplasmosis	B39
	Blastomycosis	B40
	Aspergillosis	B44
	Cryptococcosis	B45
	Pneumocystosis	B59
	Fungal meningitis	G02.1
	Fungal pneumoniae	J17.2
Parasitic infections	Cryptosporidiosis	A07.2
	Isosporiasis	A07.3
	Leishmaniasis	B55
	Toxoplasmosis	B58
	Strongyloidiasis	B78

ANNEX 9. VENOUS THROMBOEMBOLISM ICD-10 CODES

Table 10. Venous Thromboembolism ICD-10 Codes

Outcome	ICD-10 code
Venous thromboembolism	I26, I80-I82, O22.3, O22.9, O87.1, O88.2

ANNEX 10. ACUTE CARDIOVASCULAR EVENTS ICD-10 CODES

Table 11. Acute Cardiovascular Events ICD-10 Codes

Outcome	ICD-10 codes	French Medical common Procedure Coding System (CCAM codes)
Ichemic heart disease	I20-I25	Coronary catheterization with arterial dilation; CABG
Myocardial infarction	I21-I22	Coronary catheterization with arterial dilation; CABG
Cerebrovascular disease	I60-I64 (except I63.6), G45-G46	Cerebral angiography with arterial dilation or embolization
Stroke	I60-I64 (except I63.6)	Cerebral angiography with arterial dilation or embolization
Peripheral artery disease	I74	Angiography (excluding coronary catheterization and cerebral angiography and mesenteric angiography) with arterial dilation; Peripheral artery bypass
CABG, Coronary Artery Bypass Graft Surgery.		

Table 12. Acute Cardiovascular Events CCAM Codes

Outcome	French Medical common Procedure Coding System (CCAM codes)	Title
Coronary catheterization with arterial dilatation	DDAF001	Dilatation intraluminale d'un vaisseau coronaire sans pose d'endoprothèse, par voie artérielle transcutanée
	DDAF003	Dilatation intraluminale de 3 vaisseaux coronaires ou plus avec pose d'endoprothèse, par voie artérielle transcutanée
	DDAF004	Dilatation intraluminale de 2 vaisseaux coronaires avec pose d'endoprothèse, par voie artérielle transcutanée
	DDAF006	Dilatation intraluminale d'un vaisseau coronaire avec pose d'endoprothèse, par voie artérielle transcutanée
	DDAF007	Dilatation intraluminale de 2 vaisseaux coronaires avec artériographie coronaire, avec pose d'endoprothèse, par voie artériel
	DDAF008	Dilatation intraluminale d'un vaisseau coronaire avec artériographie coronaire, avec pose d'endoprothèse, par voie artériel
	DDAF009	Dilatation intraluminale de 3 vaisseaux coronaires ou plus avec artériographie coronaire, avec pose d'endoprothèse, par voie art
	DDAF010	Dilatation intraluminale d'un vaisseau coronaire avec artériographie coronaire, sans pose d'endoprothèse, par voie artérielle transcutanée
CABG	DDMA003	Revascularisation coronaire par 3 greffons artériels avec 3 anastomoses distales, par thoracotomie avec CEC
	DDMA004	Revascularisation coronaire par greffon veineux avec 4 anastomoses distales ou plus, par thoracotomie avec CEC
	DDMA005	Revascularisation coronaire par 2 greffons artériels et par greffon veineux avec 3 anastomoses distales, par thoracotomie avec CEC
	DDMA006	Revascularisation coronaire par 2 greffons artériels avec 3 anastomoses distales, par thoracotomie avec CEC

Outcome	French Medical common Procedure Coding System (CCAM codes)	Title
	DDMA007	Revascularisation coronaire par greffon veineux avec une anastomose distale, par thoracotomie avec CEC
	DDMA008	Revascularisation coronaire par 2 greffons artériels avec 4 anastomoses distales ou plus, par thoracotomie avec CEC
	DDMA009	Revascularisation coronaire par 2 greffons artériels et par greffon veineux avec 4 anastomoses distales ou plus, par thoracotomie avec CEC
	DDMA011	Revascularisation coronaire par un greffon artériel et par greffon veineux avec 2 anastomoses distales, par thoracotomie avec CEC
	DDMA012	Revascularisation coronaire par 3 greffons artériels et par greffon veineux avec 4 anastomoses distales ou plus, par thoracotomie avec CEC
	DDMA013	Revascularisation coronaire par 3 greffons artériels avec 4 anastomoses distales ou plus, par thoracotomie avec CEC
	DDMA015	Revascularisation coronaire par un greffon artériel avec une anastomose distale, par thoracotomie avec CEC
	DDMA016	Revascularisation coronaire par greffon veineux avec 3 anastomoses distales, par thoracotomie avec CEC
	DDMA017	Revascularisation coronaire par un greffon artériel avec 2 anastomoses distales, par thoracotomie avec CEC
	DDMA018	Revascularisation coronaire par un greffon artériel et par greffon veineux avec 3 anastomoses distales, par thoracotomie avec CEC
	DDMA019	Revascularisation coronaire par greffon veineux avec 2 anastomoses distales, par thoracotomie avec CEC
	DDMA020	Revascularisation coronaire par 2 greffons artériels avec 2 anastomoses distales, par thoracotomie avec CEC
	DDMA021	Revascularisation coronaire par un greffon artériel et par greffon veineux avec 4 anastomoses distales ou plus, par thoracotomie avec CEC
	DDMA022	Revascularisation coronaire par 2 greffons artériels et par greffon veineux avec 3 anastomoses distales, par thoracotomie sans CEC
	DDMA023	Revascularisation coronaire par un greffon artériel avec 2 anastomoses distales, par thoracotomie sans CEC
	DDMA024	Revascularisation coronaire par greffon veineux avec 2 anastomoses distales, par thoracotomie sans CEC
	DDMA025	Revascularisation coronaire par un greffon artériel avec une anastomose distale, par thoracotomie sans CEC
	DDMA026	Revascularisation coronaire par 2 greffons artériels avec 2 anastomoses distales, par thoracotomie sans CEC
	DDMA027	Revascularisation coronaire par greffon veineux avec 3 anastomoses distales, par thoracotomie sans CEC
	DDMA028	Revascularisation coronaire par greffon veineux avec une anastomose distale, par thoracotomie sans CEC
	DDMA029	Revascularisation coronaire par un greffon artériel et par greffon veineux avec 3 anastomoses distales, par thoracotomie sans CEC
	DDMA030	Revascularisation coronaire par 3 greffons artériels avec 3 anastomoses distales, par thoracotomie sans CEC
	DDMA031	Revascularisation coronaire par 2 greffons artériels avec 3 anastomoses distales, par thoracotomie sans CEC

Outcome	French Medical common Procedure Coding System (CCAM codes)	Title
	DDMA032	Revascularisation coronaire par un greffon artériel et par greffon veineux avec 2 anastomoses distales, par thoracotomie sans CEC
	DDMA033	Revascularisation coronaire par 2 greffons artériels avec 4 anastomoses distales ou plus, par thoracotomie sans CEC
	DDMA034	Revascularisation coronaire par 2 greffons artériels et par greffon veineux avec 4 anastomoses distales ou plus, par thoracotomie sans CEC
	DDMA035	Revascularisation coronaire par 3 greffons artériels avec 4 anastomoses distales ou plus, par thoracotomie sans CEC
	DDMA036	Revascularisation coronaire par 3 greffons artériels et par greffon veineux avec 4 anastomoses distales ou plus, par thoracotomie sans CEC
	DDMA037	Revascularisation coronaire par greffon veineux avec 4 anastomoses distales ou plus, par thoracotomie sans CEC
	DDMA038	Revascularisation coronaire par un greffon artériel et par greffon veineux avec 4 anastomoses distales ou plus, par thoracotomie sans CEC
Cerebral angiography with arterial dilation or embolization	EAQH002	Scanographie des vaisseaux encéphaliques [Angioscanner cérébral]
	EAAF002	Dilatation intraluminale du tronc de l'artère carotide interne intracrânienne avec pose d'endoprothèse, par voie artérielle transcutanée
	EAAF004	Dilatation intraluminale du tronc de l'artère carotide interne intracrânienne sans pose d'endoprothèse, par voie artérielle transcutanée
	EAAF900	Dilatation intraluminale de branche de l'artère carotide interne avec pose d'endoprothèse, par voie artérielle transcutanée
	EAAF901	Dilatation intraluminale de branche de l'artère carotide interne sans pose d'endoprothèse, par voie artérielle transcutanée
	EAFA001	Embolectomie ou thromboendartériectomie de vaisseau intracrânien, par craniotomie
	EAJF341	Évacuation de thrombus d'artère intracrânienne par voie artérielle transcutanée
	EANF002	Fibrinolyse in situ suprasélective d'artère intracrânienne, par voie artérielle transcutanée
	EASF004	Embolisation suprasélective unilatérale ou bilatérale de branche de l'artère carotide interne, par voie artérielle transcutanée
	EASF014	Embolisation sélective ou hypersélective unilatérale ou bilatérale de branche de l'artère carotide interne, par voie artérielle transcutanée
Angiography (excluding coronary catheterization and cerebral angiography and mesenteric angiography) with arterial dilation		
Peripheral artery bypass	EDCA004	Pontage artériel iliofémoral pour complication anastomotique sur prothèse de la bifurcation fémorale, par abord direct

Outcome	French Medical common Procedure Coding System (CCAM codes)	Title
	EDCA005	Pontage artériel iliofémoral homolatéral, par abord direct
	EECA001	Pontage artériel fémoropoplité au-dessus de l'interligne articulaire du genou, par abord direct
	EECA002	Pontage artériel fémorofémoral homolatéral, par abord direct
	EECA003	Pontage artériel fémoropoplité au-dessous de l'interligne articulaire du genou, par abord direct
	EECA005	Pontage d'une artère du pied, par abord direct
	EECA006	Pontage artériel subclaviofémoral ou axillobifémoral, par abord direct
	EECA007	Pontage artériel subclaviofémoral ou axillofémoral, par abord direct
	EECA008	Pontage artériel fémorotibial ou fémoropéronier sans collier veineux, par abord direct
	EECA010	Pontage artériel fémorotibial ou fémoropéronier avec collier veineux, par abord direct
	EECA012	Pontage multiple étagé [séquentiel] ou bifurqué des artères des membres inférieurs, par abord direct
CABG, Coronary Artery Bypass Graft Surgery		

ANNEX 11. EFFECTIVENESS ICD-10, UCD13, CIP13, AND CCAM CODES

Table 13. Ulcerative Colitis Effectiveness ICD-10 and CCAM Codes

Comorbidity	ICD-10 codes	French Medical common Procedure Coding System (CCAM codes)
UC-related hospitalization	K51	
Childbirth	O80-84	
UC-related surgery Colectomy		HHFA002, HHFA004, HHFA005, HHFA006, HHFA008, HHFA009, HHFA010, HHFA014, HHFA017, HHFA018, HHFA021, HHFA022, HHFA023, HHFA024, HHFA026, HHFA028, HHFA029, HHFA030, HHFA031
Intestinal resection		HGCA005, HGCC015, HGFA003, HGFA004, HGFA005, HGFA007, HGFC014, HGFC016, HGFC021
Perineal surgery and minor digestive surgery		HKPA004, HKPA005, HKPA006, HKPA007, HKPA008, HGCA008, HGCC026, HGLA001, HHCA003, HHCC011, HPPA002, HPPC003, ZCJA002, ZCJA004

Table 14. Rheumatoid Arthritis Effectiveness ICD-10 and CCAM Codes

Comorbidity	ICD-10 codes	French Medical common Procedure Coding System (CCAM codes)
RA-related hospitalization	M05 M06	
Childbirth	O80-84	
UC-related procedure Intra articular injection		MZLB001, MZLH001, MZLH002, LHLB001, LHLH002, LHLH003, NZLB001, NZLH001, NZLH002

Table 15. TNFi, Tofacitinib, and Vedolizumab UCD13 and CIP13 Codes

Medication	UCD13 code	CIP13 code
AMGEVITA 20MG/0,4ML SOL INJ SER	3400894388210	3400930142288
AMGEVITA 40MG/0,8ML SOL INJ SER	3400894388388	3400930141724
AMGEVITA 40MG/0,8ML SOL INJ SER	3400894388388	3400930141731
AMGEVITA 40MG/0,8ML SOL INJ SER	3400894388388	3400930141755
AMGEVITA 40MG/0,8ML SOL INJ STYLO	3400894388449	3400930141762
AMGEVITA 40MG/0,8ML SOL INJ STYLO	3400894388449	3400930141779
AMGEVITA 40MG/0,8ML SOL INJ STYLO	3400894388449	3400930141809
BENEPALI 25MG/0,51ML SOL INJ SER	3400894260318	3400930092675
BENEPALI 50MG SOL INJ STYLO	3400894166764	3400930044605
BENEPALI 50MG/0,98ML SOL INJ SER	3400894166535	3400930044599
CIMZIA 200MG PDR ET SOL INJ NSFP	3400892992594	3400959059291
CIMZIA 200MG/ML SOL INJ SER 1ML	3400893426012	3400939732008
CIMZIA 200MG/ML SOL INJ STYLO 1ML	3400894212232	3400930076286
ENBREL 10MG PDR ET SOL INJ PEDIAT	3400893910481	3400921676327
ENBREL 25MG PDR ET SOL INJ	3400892426150	3400936064997
ENBREL 25MG PDR INJ NSFP	3400892365954	3400935851710
ENBREL 25MG/0,5ML SOL INJ SER	3400892958538	3400937719100
ENBREL 25MG/0,5ML SOL INJ STYLO	3400894344674	3400930111017
ENBREL 25MG/ML INJ PEDIAT NSFP	3400892958767	3400937684118
ENBREL 50MG PDR ET SOL INJ NSFP	3400892707877	3400936586222
ENBREL 50MG PDR INJ NSFP	3400892800240	3400936585850
ENBREL 50MG/1ML SOL INJ SER	3400892958828	3400937719568
ENBREL 50MG/1ML SOL INJ STYLO	3400893417256	3400939605227
ERELZI 25MG/0,5ML SOL INJ SER	3400894275275	3400930098943
ERELZI 50MG/1ML SOL INJ SER	3400894275336	3400930098950
ERELZI 50MG/1ML SOL INJ STYLO	3400894275565	3400930098967
FLIXABI 100MG PDR INJ FL	3400894187875	3400955023012
HULIO 40MG/0,8ML SOL INJ FL	3400894418238	3400930157428
HULIO 40MG/0,8ML SOL INJ SER	3400894418009	3400930157374
HULIO 40MG/0,8ML SOL INJ STYLO	3400894418177	3400930157404
HUMIRA 20MG/0,2ML SOL INJ SER	3400894362425	3400930126165
HUMIRA 40MG/0,4ML SOL INJ SER	3400894166184	3400930042489
HUMIRA 40MG/0,4ML SOL INJ STYLO	3400894166306	3400930042687
HUMIRA 40MG/0,8ML SOL INJ FL NSFP	3400893686645	3400941851728
HUMIRA 40MG/0,8ML SOL INJ SER NSFP	3400892510927	3400936223059
HUMIRA 40MG/0,8ML SOL INJ STYLO NSFP	3400893075623	3400937801454
HUMIRA 80MG/0,8ML SOL INJ SER	3400894362593	3400930116494
HUMIRA 80MG/0,8ML SOL INJ STYLO	3400894362654	3400930116500
HUMIRA 80MG/0,8ML SOL INJ STYLO	3400894362654	3400930178805
HYRIMOZ 40MG/0,8ML SOL INJ SER	3400894403098	3400930155080
HYRIMOZ 40MG/0,8ML SOL INJ STYLO	3400894402909	3400930155097
IDACIO 40MG/0,8ML SOL INJ FL	3400894501534	3400930172766
IDACIO 40MG/0,8ML SOL INJ SER	3400894501305	3400930172780
IDACIO 40MG/0,8ML SOL INJ STYLO	3400894501473	3400930172797
IMRALDI 40MG/0,8ML SOL INJ SER	3400894400196	3400930143711
IMRALDI 40MG/0,8ML SOL INJ SER	3400894400196	3400930143728
IMRALDI 40MG/0,8ML SOL INJ STYLO	3400894400257	3400930144114
IMRALDI 40MG/0,8ML SOL INJ STYLO	3400894400257	3400930144121
INFLECTRA 100MG PDR INJ FL	3400893983522	3400958598043
REMICADE 100MG PDR INJ FL	3400892137377	3400956207015
REMSIMA 100MG PDR INJ FL	3400894023036	3400958676208
SIMPONI 100MG/1ML SOL INJ SER	3400893987483	3400927568480
SIMPONI 100MG/1ML SOL INJ STYLO	3400893987544	3400927568312

Medication	UCD13 code	CIP13 code
SIMPONI 50MG/0,5ML SOL INJ SER	3400893771204	3400939730974
SIMPONI 50MG/0,5ML SOL INJ STYLO	3400893771372	3400939730745
XELJANZ 10MG CPR	3400894501824	3400930159453
XELJANZ 5MG CPR	3400894277347	3400930088111
ENTYVIO 300MG PDR INJ FL	3400894020592	3400958672873

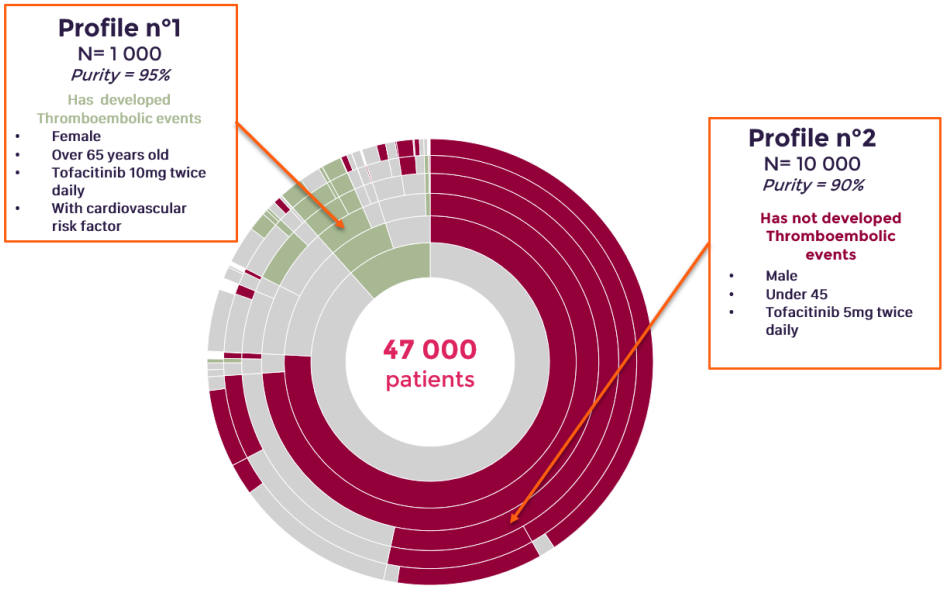
ANNEX 12. COHORT PROFILING ANALYSIS WITH MACHINE LEARNING

A cohort profiling analysis could be carried out using the characteristics of patients treated by tofacitinib to **identify homogeneous patient groups** who are more at risk of having each of the three events of interest (serious infections, major cardiovascular and thromboembolic events).

For that, we propose to use **supervised-learning analytic methods.** There are multiple technical methods of supervised learning (decision trees, linear statistical models, SVM, neural networks, etc.). Algorithms based on decision tree(s) are widely used on structured data for their performance and clarity. They consist in finding the best successive separation thresholds to isolate homogeneous groups of patients. The accumulation of these separations then forms a comprehensible and relevant profile (for example: women over 65 years old, treated by tofacitinib 10 mg twice daily, with one cardiovascular risk factor). We propose a visualization method to highlight these separations: the sunburst (example in (Figure 8).

The whole cohort of patients treated by tofacitinib starts in the centre of the figure (Ex.: 47 000 patients), then each patient follows an outward radius according to a decision criterion specific to each level, then the colour indicates the belonging to a profile with regard to the outcome (eg, "has developed" thromboembolic events in green and "has not" developed thromboembolic events in red). When the homogeneity of a band is judged insufficient, it is greyed out.

Figure 8. Example of "Sunburst" Graph for Subgroups Identification Regarding Thromboembolic Events.



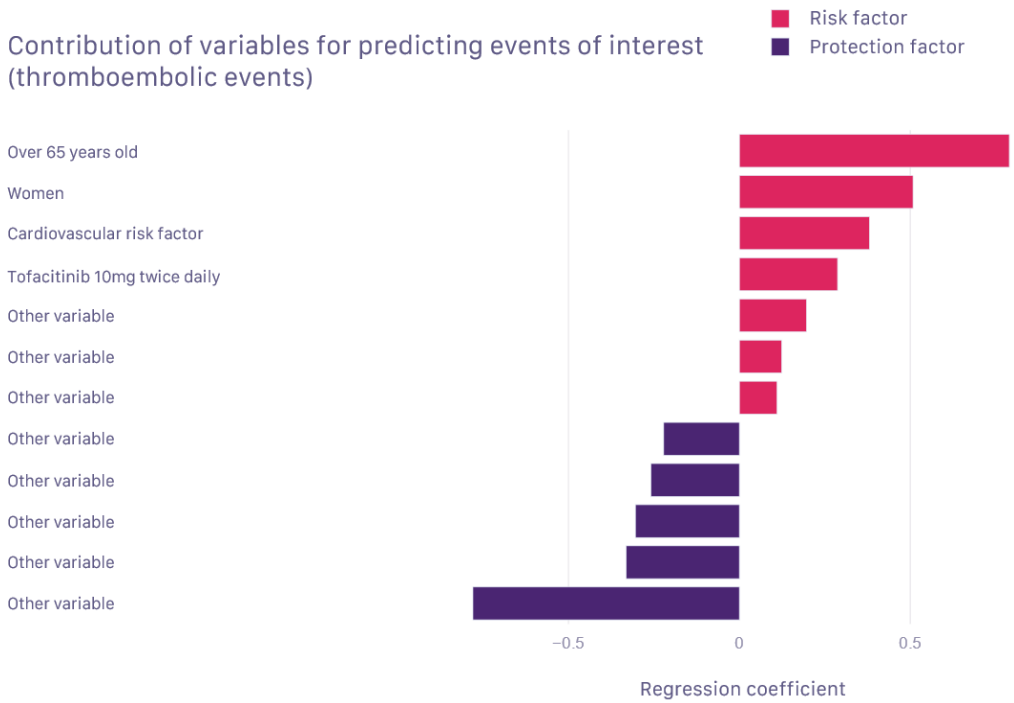
The **value of identifying subgroups** will be to identify precise trends within the cohort. This provides a finer knowledge than the overall view of the whole cohort. Moreover, if no subgroup is statistically identified, this will confirm the situation of a very random effect or a very homogeneous cohort.

Several axes can be discussed to obtain the most relevant profiles, especially the comorbidities or the treatment dosage.

The validity of the model will be ensured algorithmically by cross-validating on several metrics relevant to the cost analysis (Accuracy & Recall, F1 score, CUA, IEM, MAE). In addition, performance will be compared to two simplistic reference models: modal model (mean prediction) and random model.

The machine learning algorithm will also provide a **ranking of the importance of the explanatory variables** in the constructed model. We will have two results: (1) the result where this importance is only positive [0-100%] (the case of decision trees), and (2) it is negative and positive (such as the linear model shown in Figure 9). In this second case we can then differentiate factors correlated with increased risk from those correlated with less risk.

Figure 9. Example of a Graph Showing the Most Important Variable Contributions in a Model



Once the **relevant profiles** have been discovered, we propose to statistically compare each of them to the rest of the cohort. This will allow us to quantify the effect of belonging to a profile relative to the rest of the cohort. Thus, for each highlighted profile its odds ratio and relative risk are provided, with the bilateral 95% confidence interval, to better appreciate the measurement of the effect and its significance or not. Table 16 gives an example of how these measures can be reconstructed.

Table 16. Example of Quantification of the Effects of Belonging to a Profile with Odds Ratios

Event	Profile 1	Cohort except P1	Significant difference	Profile 2	Cohort except P2	Significant difference
Thromboembolic event	OR **2,25** [2,15 - 2,35]	1 (Baseline value)	☑	OR **0,67** [0,55- 0,83]	1 (Baseline value)	☑

For each profile identified, we produce an interpretable measure (eg, a patient corresponding to profile 1 is 2.25 times more at risk of developing a Thromboembolic event compared to the rest of the cohort).

In conclusion, the profiling analysis by Machine Learning will produce the following 3 results:

- The list of profiles identified (number of patients and their characteristics);
- A histogram of the contribution of the explanatory variables to the target variable;
- The comparative table for each subgroup vs. the rest of the cohort with respect to outcomes.

Document Approval Record

Document Name:	A3921399 Non-Interventional Protocol 30 April 2021
Document Title:	A3921399 Non-Interventional Protocol 30 April 2021

Signed By:	Date(GMT)	Signing Capacity
Mammar, Nadir	01-Jul-2021 13:06:48	Manager Approval
De Bernardi, Barbara	05-Jul-2021 13:47:23	EUQPPV Approval