Title	A Post-Authorization Safety Study of Golimumab in UC Using the Spanish ENEIDA Registry
Protocol Version identifier	Version 1.1
Date of last version of protocol	17 Nov 2022
EU PAS Register No:	EUPAS15752
Active substance	ATC code L04AB06 – golimumab
Medicinal product:	SIMPONI® (golimumab) solution for injection
Product reference:	[redacted]
Procedure number:	[redacted]
Marketing authorisation holder(s) (MAH)	Janssen Biologics B.V.
Joint PASS	No
Research question and objectives	This study seeks to evaluate whether the use of GLM is associated with risk of colectomy for intractable disease, advanced neoplasia (colorectal cancer (CRC) or high grade dysplasia (HGD)), and hepatosplenic T cell lymphoma (HSTCL) in patients with UC as compared with alternative therapies for similar severity of disease. No a priori hypotheses will be evaluated. <u>Primary objectives</u>: 1. To describe the clinical and demographic profile of first-time users of GLM in the treatment of UC compared with the corresponding profile of first-time users of comparator therapies (other anti-TNF agents or thiopurines) 2. For patients with UC initiating GLM or other anti-TNF agents, describe the risk of incident colectomy for intractable disease 3. For patients with UC initiating GLM, other anti-TNF agent, or a thiopurine, describe the risk of the composite endpoint of incident CRC or HGD [hereafter 'advanced colonic neoplasia' (ACN)] 4. Compare risk of incident colectomy for intractable disease

	between GLM and other anti-TNF agents 5. Compare risk of incident ACN between GLM and other anti-TNF agents <u>Secondary objectives</u> 1. For patients with UC initiating GLM, other anti-TNF agent, or thiopurines, describe the risk of incident CRC 2. Compare risk of incident CRC between GLM and other anti-TNF agents 3. If baseline characteristics suggest comparability between cohorts of patients receiving GLM and thiopurines, the following risks will be compared between the two cohorts: a. risk of incident ACN b. risk of incident CRC <u>Exploratory objective</u> - To describe the incidence of HSTCL in each of the study cohorts
Country(-ies) of study	Spain
Authors	[redacted]

Marketing authorisation holder(s)	Janssen Biologics B.V. [redacted]
MAH Contact Person	[redacted]

1 TABLE OF CONTENTS

2 LIST OF ABBREVIATIONS

ACN	advanced colonic neoplasia
AD	automated data
AE	adverse event
5-ASA	5-aminosalicylic acid
AZA	Azathiopurine
CD	Crohn's disease
CI	cumulative incidence
CRC	colorectal cancer
CRF	case report form
ENEIDA	Estudio Nacional en Enfermedad Inflamatoria intestinal sobre Determinantes genéticos y Ambientales
EU-RMP	European Union Risk Management Plan
GETECCU	Grupo Español de Trabajo en Enfermedad de Crohn y Colitis Ulcerosa
GLM	golimumab
HGD	high-grade colorectal dysplasia
HOI	health outcome of interest
HSTCL	hepatosplenic T cell lymphoma
IBD	inflammatory bowel disease
IRB	institutional review board
LGD	low-grade dysplasia
LTFU	lost to follow-up
6-MP	6-mercaptopurine
MTX	methotrexate
NCC	nested-case control
NSAR	non-serious adverse reaction
PSC	primary sclerosing cholangitis
PY	person-years
SAP	statistical analysis plan
SAR	serious adverse reaction
TNF	tumor necrosis factor
TP	thiopurine
UC	ulcerative colitis

3 RESPONSIBLE PARTIES

Principal investigators	[redacted]
Sponsor contacts	Merck Sharp & Dohme, LLC a subsidiary of Merck & Co., Inc. [redacted]
Marketing authorization holder	Janssen Biologics B.V. Einsteinweg [redacted]
Investigators	See stand-alone document in Annex 1.

4 ABSTRACT

Title	A Post-Authorization Safety study of Golimumab in UC Using the Spanish ENEIDA Registry
Protocol Number / Version	MK-8259-042/Version 1.1
Date	17-November-2022
Authors	[redacted]
Rationale & Background	Ulcerative colitis (UC) is a chronic inflammatory bowel disease (IBD) of unknown etiology that may be
	associated with severe symptoms and impaired quality of life. UC may affect all ages and medical treatment depends on disease severity and extent. In the case of intractable disease, colectomy may be indicated. Patients with UC are at increased risk of developing colorectal cancer (CRC) and dysplasia compared to the general population. Simponi (golimumab; GLM) received European marketing authorization for treatment of moderately to severely active ulcerative colitis on 19-Sep-2013. This registry-based study will provide additional information on colectomy, colorectal cancer and dysplasia, and hepatosplenic T-cell lymphoma (HSTCL), as outlined in the Risk Management Plan for SIMPONI® that was approved with authorization of the UC indication. This study will use data from ENEIDA, a large prospectively maintained registry of patients with IBD in Spain.
Research Question(s) & Objective(s)	This study seeks to evaluate whether the use of GLM is associated with a risk of colectomy for intractable disease, advanced neoplasia (colorectal cancer or high grade dysplasia), and HSTCL in patients with UC as compared with alternative therapies for similar severity of disease. No a priori hypotheses will be evaluated. <u>Primary objectives</u>: 1. To describe the clinical and demographic profile of first-time users of GLM in the treatment of UC compared with the corresponding profile of first-time users of comparator therapies (other anti-TNF agents or thiopurines)

	2. For patients with UC initiating GLM or other anti-TNF agents, describe the risk of incident colectomy for intractable disease 3. For patients with UC initiating GLM, other anti-TNF agent, or a thiopurine, describe the risk of composite endpoint of incident CRC or high-grade dysplasia (HGD) [hereafter 'advanced colonic neoplasia' (ACN)] 4. Compare risk of incident colectomy for intractable disease between GLM and other anti-TNF agents
	5. Compare risk of incident ACN between GLM and other anti- TNF agents <u>Secondary objectives</u> 1. For patients with UC initiating GLM, other anti-TNF agent, or thiopurine, describe the risk of incident CRC 2. Compare risk of incident CRC between GLM and other anti- TNF agents 3. If baseline characteristics suggest comparability between cohorts of patients receiving GLM and thiopurines, the following risks will be compared between the two cohorts: a. risk of incident ACN b. risk of incident CRC <u>Exploratory objective</u> The exploratory objective will be to describe the incidence of HSTCL in each of the study cohorts
Study Design	Long-term, non-interventional observational study design. It will use a new user bi-directional cohort design with the option for a nested case-control (NCC) analysis. The cohort study will use data that are primarily collected for the Spanish ENEIDA IBD registry, and the NCC analysis will also use data from retrospective review of selected medical charts.

Population	The study population will be drawn from the ENEIDA centers judged to have research quality data. Patients will be selected into the study if they are older than 18 years, diagnosed with UC, have not experienced a study outcome, and initiate therapy with GLM, an anti-TNF agent other than GLM, or thiopurine between 19 September 2013 (date of GLM EU approval for UC) through 31 December 2021. Patients enrolled during this interval will be followed-up through 30 March 2022. Patients will be excluded from the study if they were previously exposed to vedolizumab or any novel immunomodulatory agent that is newly marketed during the study period.
Variables	*Exposures* The main exposure of interest is treatment with GLM prescribed for UC. The primary comparator will be other anti-TNF agents prescribed for UC (i.e., infliximab, adalimumab, including biosimilars). A secondary comparator group will comprise the thiopurine analogs, azathioprine and 6-mercaptopurine. The two anti- TNF agents that comprise the primary comparator group will be evaluated in subgroup analyses if sample size permits. Over time, UC patients who initiate GLM may have received or may later receive comparator treatments in various sequences. Analyses will consider the time- dependent nature of these exposures. *Outcomes* The primary outcomes comprise: • Colectomy due to intractable disease • Composite ACN (CRC or HGD) The secondary outcome is CRC The exploratory outcome is HSTCL The *covariates* include (but may not be limited to): • Age • Gender • Year of cohort entry • Disease duration • Extent of disease

	• Treatments for UC that are not study exposures (e.g., systemic steroids and cyclosporine), which may be proxies for disease activity • History of previous treatments with other anti-TNF agents • History of hospitalization for UC • History of primary sclerosing cholangitis (PSC) • Other covariates may be included after analysis of the descriptive data
Data Sources	This study will use automated data that have been collected as part of ENEIDA, a large prospectively maintained registry of patients with IBD in Spain. With a current census of more than 15,000 UC patients, ENEIDA contains information about patients' disease history, medical and surgical treatments, and potential complications including colectomy and CRC. In addition to automated registry data, additional clinical data will be obtained from study sites that agree to participate in a chart review substudy; medical records will be obtained for all cases of study outcomes and a sample of control patients (those who did not experience these outcomes).
Study Size	This study will use all available data from qualified ENEIDA sites for patients with UC who initiate GLM, other anti-TNFα agents, and thiopurines during the study period. The number of patients who will qualify for this study is not known in advance and will depend on prescribing practices in ENEIDA participating centers over the course of the study. Results of a pilot study provide some information to estimate the thresholds of detectable relative risk of the primary study outcomes. For the outcome of colectomy for intractable disease, this study is estimated to have 80% power to detect a relative risk (RR) of 2.0 for the comparison of GLM exposure to other anti-TNF agents, assuming 1700 PY at risk across all anti-TNF exposures, an estimated 25% of which would be for GLM; an incidence rate of colectomy of 33.8/1000 PY; and alpha set at 5%. For the outcome of ACN, which is much rarer, the detectable RR is estimated to be 6.1, under identical assumptions except for an incidence rate of 1.8/1000 PY and a cumulative 3150 PY at risk to all anti-TNF agents (the risk window is longer for neoplasia outcomes).

Data Analysis	All analyses will be conducted based on automated registry data, except for the nested-case control analysis, which will rely on automated data supplemented with information from chart reviews. Baseline analyses will describe each cohort in terms of patient characteristics. For each of the three cohorts, annual enrolment will be described, along with the frequency of study outcomes and cumulative person- years of follow-up accrued. The incidence rate of primary and secondary outcomes (colectomy due to intractable disease, ACN, and CRC) will be estimated for each cohort. Next, the cumulative incidence of primary and secondary outcomes will be estimated using time-to-event analyses, overall by cohorts and then in stratified analyses. Stratification factors (as measured at study entry) will be evaluated one at a time and will include gender, time since initial UC diagnosis, history of primary sclerosing cholangitis (PSC), UC hospitalization, and previous use of systemic steroids; and for anti-TNF cohorts, concurrent use of thiopurines and history of previous anti-TNF use. For each primary and secondary outcome, risk will be compared between GLM and comparators using survival analysis (Kaplan Meier plots and Cox proportional hazards regression models). Study exposures will be treated as time-dependent variables. Hazard ratios will be used to estimate relative risk. Candidate variables to evaluate as potential confounders include the list of covariates under Variables, listed above. At the conclusion of the study, nested case-control analyses will be conducted for the outcomes of colectomy for intractable disease and, if sufficient events are observed, ACN. For each case of an outcome, up to 2 controls will be sampled at random from risk sets matched on calendar time of outcome and duration of UC on the index date. Exposure will be based on the history of treatment with GLM and comparator treatments, ascertained from automated data and supplemented with information from chart review. Conditional logistic regression will be used to calculate odds ratios as an estimate of relative risk. Potential confounders include the list of covariates under Variables, listed above. The incidence rate of HSTCL will be described for each cohort.

Milestones	
Start of data collection:	[redacted]
End of data collection:	[redacted]
Study progress report(s):	[redacted]
Final report of study results:	[redacted]

5 AMENDMENTS AND UPDATES

This table summarizes all key changes made to the protocol.

Number	Date	Section of Study Protocol	Amendment or Update	Reason
1	24 Oct 2019	9.2.3.1	Update	To reflect changes in methods as described in 4th progress report.
2	23 Nov 2021	9.2.2, 9.2.3.1, 9.3.2, 9.7.2.1,	Update	To reflect changes in study design, methods, and outcomes as described in 5th progress report.
3	17 Nov 2022	Title Page, 3, 4, 6, 9.2.2, 9.2.3.2, 9.3.3, 9.7, 9.8, 12, Annex 1	Amendment	To reflect administrative and milestone changes, and protocol changes inclusive of study design, methods, data analysis, outcomes, and quality control.

6 MILESTONES

Milestone	Planned date
Start of data collection	[redacted]
End of data collection	[redacted]
Study progress report 1	[redacted]
Study progress reports	[redacted]
Final report of study results	[redacted]

Note that the study period begins on [redacted] and the end of the follow-up period is [redacted].

7 RATIONALE AND BACKGROUND

7.1 Rationale

SIMPONI® (golimumab [GLM]), a TNF antagonist, was approved on October 1, 2009 in the European Union for the indications RA, AS and PsA. GLM subsequently received European marketing authorization for the treatment of moderately-to-severely active ulcerative colitis (UC) on September 19, 2013. In connection with the approval in this indication, it was agreed that the EU-RMP should include post-marketing follow-up activities for collection of additional information such as colorectal cancer (CRC), colorectal dysplasia, hepatosplenic T cell lymphoma (HSTCL) and colectomy in the patient population with moderately to severely active UC. This registry-based study provides additional information on CRC and dysplasia, colectomy, and HSTCL, as outlined in the EU-RMP for SIMPONI® that was approved with authorization of the UC indication.

Since tumor necrosis factor (TNF) mediates inflammation and modulates cellular immune responses, the possibility exists for TNF antagonists, including GLM, to cause immune suppression affecting host defenses against infections and malignancies. To date, neither the pivotal registration trials evaluating GLM for the induction and maintenance of remission in patients with moderate to severe UC,[1, 2] nor any of the other clinical trials in the development program for GLM, have demonstrated an association between GLM and increased risk of CRC, dysplasia, HSTCL, or colectomy. Use of data collected in large observational registries provides an additional source of safety data.

This study will use data collected as part of an ongoing registry of inflammatory bowel diseases in Spain, ENEIDA (Estudio Nacional en Enfermedad Inflamatoria intestinal sobre Determinantes genéticos y Ambientales). This registry is expected to capture the majority of GLM treatments used for UC in Spain since the date of its market authorization.

To provide a context for interpreting the long-term safety data on UC patients treated with

GLM, this study will also follow similar patients with UC treated with alternative therapies for UC, including other TNF inhibitors and thiopurines.

7.2 Background

Disease background. UC is a chronic inflammatory bowel disease (IBD) of unknown etiology characterized by inflammation primarily involving the colonic mucosa. The diagnosis of UC peaks between the ages of 15 and 35 years, but UC may affect all ages. The symptoms of UC depend on the extent and severity of disease and may include bloody diarrhea and rectal bleeding, along with systemic symptoms of fever and weight loss. The clinical course is typically relapsing and remitting, although occasionally it may take an unremitting, continuous course. Anatomically, the inflammation in UC is uniform and continuous, with no intervening areas of normal mucosa. In nearly all cases, inflammation involves the rectum, and it extends proximally for a variable distance.[3]

Medical therapy. Medical treatment for UC depends on disease severity and extent. Patients with mild-to-moderate UC are treated initially with oral 5-aminosalicyclic acid (5-ASA) medications combined with topical treatment with 5-ASA or steroid suppositories, plus 5- ASA enema or steroid foam preparations. The oral 5-ASA preparations should then be continued for maintenance of remission. For patients with moderate-to-severe UC, oral steroids are used to achieve initial disease control, followed by thiopurines (6- mercaptopurine (6-MP) or azathioprine (AZA)) for maintenance in a step-up therapeutic strategy. For patients who do not respond to or are intolerant to thiopurines, the anti-tumor necrosis factor (anti-TNF) agents infliximab, adalimumab or GLM may be introduced. For patients hospitalized with acute severe disease or moderate-to-severe persistent UC who have not responded to corticosteroids after five-to-seven days of treatment, therapeutic choices are either intravenous cyclosporine, an anti-TNF agent or colectomy. As compared with first- and second-line therapies, anti-TNF agents tend to be preferentially prescribed to patients with more severe, treatment-resistant disease.

In addition to the preceding general treatment scheme, vedolizumab, a monoclonal antibody against alfa-4-beta-7 integrin, was approved in 2014 to treat moderate to severe UC when conventional therapy or TNF-alpha antagonists are ineffective, no longer effective, or cannot be tolerated by the patient.

Colectomy. When medical therapy fails, complete colectomy is the surgical therapy of choice. The vast majority of colectomies in UC are performed for intractable disease, which is the one of the primary endpoints of this study. Colectomy is also performed to treat CRC. High-grade colorectal dysplasia (HGD) is also an indication for colectomy, while controversy exists on how to manage UC patients with low-grade dysplasia (LGD). In a recent Norwegian population-based cohort study, colonic neoplasia accounted for only about 10% of colectomies.[4] Less common reasons for colectomy in UC include emergent complications such as toxic megacolon, colonic perforation, massive hemorrhage and colonic obstruction. The reported rates of colectomy in patients with UC vary widely depending on severity and extent of disease, clinical practice in different countries, and other factors. The 10-year cumulative colectomy rate ranged from 9% to 28% in large UC cohort studies.[5, 6] An ENEIDA study of UC patients who had used thiopurines for at least 3 months reported a colectomy rate of

9.4/1000 person-years (PY).[5] Risk for colectomy appears greatest in the first two years following diagnosis[4] and both in North America and Europe appears to be diminishing over time, in parallel with increased use of potent immunomodulatory therapies for UC.[7] In randomized clinical trials, patients with moderately to severely active UC treated with infliximab were less likely to undergo colectomy through 54 weeks than those receiving placebo.[8]

Neoplasia. Compared to the general population, patients with UC are at an increased risk of colorectal cancer and colorectal dysplasia.[9] Colorectal dysplasia is subcategorized into HGD and LGD. Because HGD carries a high probability of progression to CRC, it is typically treated with colectomy unless it arises from a sporadic polyp, in which case polypectomy is the treatment of choice. LGD is associated with a smaller risk of progression to cancer and only rarely leads to colectomy; it can usually be managed with endoscopic removal of affected lesions and enhanced surveillance.[10] Professional society guidelines suggest enhanced surveillance with colonoscopy to monitor for CRC and dysplasia in patients with longstanding UC.[3, 11] Screening colonoscopy is generally recommended to start eight years after initial diagnosis.

The incidence of CRC varies within the UC population. Risk factors for cancer in persons with UC include a long duration of disease regardless of clinical activity; extensive involvement of the colon; a young age at onset of UC; severe persistent inflammation; the presence of primary sclerosing cholangitis (PSC); and a family history of CRC. In a recent meta-analysis, the incidence rate of CRC was 1.24/1000 PY in population-based studies and 4.02/1000 PY in studies that only included patients with extensive colitis.[12]

The association between anti-TNF agents and risk of CRC has not been widely studied in the UC population. However, a nationwide register-based study in Denmark was conducted among patients with IBD from 1999-2012. This study found no association between anti- TNF therapy and colorectal cancer risk (adjusted relative risk 1.0 (95% CI, 0.48-2.08).[13]

In addition to CRC and dysplasia, HSTCL has been identified as a possible risk in patients with IBD treated with anti-TNF therapies. In the IBD population, the disease affects primarily males younger than 35 year, and almost all patients who develop HSTCL have also been previously exposed to thiopurines.[14] The incidence of HSTCL in the UC population is extremely rare, although it has not been formally quantified. In a population-based study in a managed care population in the US (2000-2006), the standardized incidence rate of HSTCL was 0.3 (95% CI, 0.11-0.65) per million PY.[15]

8 RESEARCH QUESTION AND OBJECTIVES

This study will address whether, in patients with UC, the use of GLM is associated with risk of colectomy for intractable disease, advanced neoplasia (CRC or high grade dysplasia), and HSTCL as compared with alternative therapies for similar severity of disease. In this study, no a priori research hypotheses have been formulated.

Primary objectives:

1. To describe the clinical and demographic profile of first-time users of GLM in the treatment

of UC compared with the corresponding profile of first-time users of comparator therapies (other anti-TNF agents or thiopurines)

2. For patients with UC initiating GLM or other anti-TNF agents, describe the risk of incident colectomy for intractable disease

3. For patients with UC initiating GLM, other anti-TNF agent, or a thiopurine, describe the risk of the composite endpoint of incident CRC or HGD [hereafter 'advanced colonic neoplasia' (ACN)]

4. Compare risk of incident colectomy for intractable disease between GLM and other anti-TNF agents

5. Compare risk of incident ACN between GLM and other anti-TNF agents

Secondary objectives

1. For patients with UC initiating GLM, other anti-TNF agent, or thiopurine, describe the risk of incident CRC

2. Compare risk of incident CRC between GLM and other anti-TNF agents

3. If baseline characteristics suggest comparability between cohorts of patients receiving GLM and thiopurines, the following risks will be compared between the two cohorts:

    a. risk of incident ACN

    b. risk of incident CRC

Exploratory objective

Describe the incidence of HSTCL in each of the study cohorts

9 RESEARCH METHODS

9.1 Study design

This is a long-term observational (non-interventional), post-authorization safety study. It will use a new user bi-directional cohort design with the option for a nested case-control (NCC) analysis. The cohort study will use data that are primarily collected for the Spanish ENEIDA IBD registry, while the NCC analysis will also use data from retrospective review of selected medical charts, which were primarily recorded to document clinical care.

Data from the ENEIDA registry will be used to identify cohorts of patients with UC who are new users of GLM (the main exposure of interest) or new users of comparator therapies (other

anti-TNF agents or thiopurines).

Study patients will be characterized at baseline and followed for up to 8 years from cohort entry to determine the incidence of colectomy for intractable disease, ACN, and HSTCL.

Over the course of the study period, it is possible that patients may add, discontinue, or switch therapies. Because the potential effect of biologic agents on subsequent cancer risk may extend beyond the actual period of exposure, patients who develop cancer outcomes may have been exposed to more than one study drug. The planned NCC analyses provide the flexibility of evaluating several definitions of exposure with separate analyses for colectomy and ACN. These analyses will also incorporate supplemental information from clinical records to verify outcomes and support more complete adjustment of covariates that are incompletely captured in the automated data.

9.2 Setting

9.2.1 The ENEIDA registry

ENEIDA is a large prospectively maintained registry of patients with IBD in Spain. It is conducted under the auspices of a national scientific society devoted to the study of IBD (GETECCU: Grupo Español de Trabajo en Enfermedad de Crohn y Colitis Ulcerosa). Data in the ENEIDA registry come from a network of more than 50 academic and community gastroenterology practices across Spain that have an interest in IBD. Its census of patients with UC is more than 15,000. Participation in the registry is voluntary for both physicians and patients, and physicians who contribute data to ENEIDA receive no payment for this registry work. ENEIDA does, however, employ its own data manager.

Participating practices enter the data onto an electronic case report form (CRF), which are immediately updated in the central ENEIDA database. Automated data in ENEIDA comprise a pre-defined set of clinical variables. The registry mandates completion of a limited number of fields (e.g., age, IBD diagnosis, immunomodulator and biologic therapy start/stop dates, any bowel surgery) on the CRF, while completion of other fields is optional. ENEIDA's data coordinating center monitors each site's data monthly with respect to the completeness of capture in the CRF of the "mandatory variables". Sites with low data completion are requested but not compelled to be more thorough. Mandatory variables are described in section 9.4, Data sources.

Some studies using the ENEIDA registry include data only from selected practices that are judged to be research quality, based on the percentage of the practice's IBD patients who have been enrolled in ENEIDA and the percentage completion of mandatory data fields.

9.2.2 Study population and cohort identification

The study population for this study will be drawn from the ENEIDA centers judged to have research quality data. Sites that satisfy the following two criteria at the start of the study will qualify: (a) enroll at least 75% of their IBD patients into the registry (a self-reported variable) and (b) enter at least 75% of mandatory variables on the CRFs, based on monthly audit of

cumulative practice data. These sites will remain in the study for the duration of the study period. At the end of the study, data quality will be reassessed retrospectively, and a secondary analysis will exclude patients from any sites that no longer meet the initial qualification criteria.

Patients will be selected into the study population if they are older than 18 years, diagnosed with UC, have not experienced a study outcome, and initiate therapy with GLM, an anti-TNF agent other than GLM, or thiopurine starting 19 September 2013 (the date of GLM EU approval for UC) through 31 December 2021 (an 8-year enrollment period). Follow-up will continue through 30 March 2022. Follow-up duration for the study population is thus expected to range from 2 months to 8 years, depending on the date of cohort entry. Specifications of each of the 3 cohorts are as follows:

- The GLM cohort: Patients with UC will qualify for entry into the GLM cohort if he/she newly initiates GLM for the treatment of UC for the first time after 19 September 2013; prior use of a different anti-TNF agent is permitted, as is prior or concurrent use of thiopurines.

- The other anti-TNF cohort: Patients with UC will qualify for entry into the anti-TNF cohort if he/she initiates a new anti-TNF agent (infliximab, adalimumab including biosimilars) for the treatment of UC for the first time after 19 September 2013; prior use of a different anti-TNF agent is permitted, as is prior or concurrent use of thiopurines. This cohort will serve as the primary reference group for all comparative evaluations.

- The thiopurine cohort: Patients with UC who newly initiate a thiopurine (AZA or 6-MP) for the treatment of UC for the first time after 19 September 2013 will qualify for entry into the thiopurine cohort if they are naïve to both thiopurines and anti-TNF agents. This cohort will serve as the secondary comparator for advanced colonic neoplasia outcome only. In practice, an estimated 10% of patients starting AZA develop early intolerance of the drug, and many are switched to its metabolite 6-MP. Patients who switch from AZA to 6-MP will be considered to have continuous exposure to thiopurines, provided that the gap between end of AZA and start of 6-MP is less than 4 weeks.

Given how thiopurines are currently used for treatment of UC, specifying thiopurine exposure as a comparator for the colectomy outcome would not be appropriate. Nearly all colectomies for intractable disease in Spain are preceded by a trial of anti-TNF agents, and nearly all users of anti-TNF agents will have first received a thiopurine. In a separate ENEIDA study, it was shown that 95.8% of patients starting adalimumab for UC had received thiopurines at some point in the past.[16]

Specifically, patients are required to meet the following selection criteria to be included for the main cohort study:

Inclusion Criteria:

- Patient with UC in a research-quality site.
- Aged 18 years or older at the date of study drug initiation.

- Qualified for one of the cohorts between 19 September 2013 and 31 December 2021.
- Date of first prescription of cohort-defining drug (index date) occurred within a clinically credible period (< 6 months) after the last recorded clinic visit in ENEIDA. Index dates beyond this range raise concerns that the clinical record for this patient may be incomplete.

Exclusion criteria:

- Initiation of study drugs for indications other than UC, such as rheumatoid arthritis or psoriasis.
- Evidence of any of the study outcomes before cohort entry:
  - Complete or partial colectomy
  - ACN
  - HSTCL
- For each of the 3 cohorts (ie, GLM, other anti-TNF, and thiopurine), if the patient initiated the drug defining the corresponding cohort for the first time before 19 September 2013. In other words, patients did not qualify for entry into a cohort if they were prevalent users of that drug before the study start date. However, patients could enter the study later based on subsequent initiation of other cohort-defining drugs.
- If, prior to cohort entry, patients had initiated a novel biological or immunomodulator agent, for example:
  - Vedolizumab: Entyvio®
  - Natalizumab: Tysabri®, Antegren®
  - Denosumab: Prolia®, Xgeva®
  - Etrolizumab: Raptiva®
  - Tocilizumab: Actemra®, RoActemra®
  - Ustekinumab: Stelara®
  - Certolizumab: Cimzia®
  - Tofacitinib: Xeljanz®

This list defines the drugs referred to as “vedolizumab and other novel immunomodulators” mentioned in the protocol and covers drugs that would likely be prescribed to patients with more severe UC and that are potentially related to the study outcomes

9.2.3 Patient follow-up

9.2.3.1 Definitions of follow-up time

The patient follow-up time is the time from date of cohort entry until the date of end of follow-up. Date of cohort entry corresponds to the date of first treatment with one of the cohort-defining medications after other inclusion criteria have been met (age of 18 years and UC diagnosis).

In the separate analysis of each outcome, the date of end of follow-up will be defined as the earliest of the following events: occurrence of study outcome of interest (colectomy for intractable disease, CRC, HGD or HSTCL), withdrawal from the registry, death, end of study period, or other censoring events. These include:

- Total or partial colectomy for any cause other than intractable UC or ACN are censoring criteria for both colectomy for intractable disease and ACN. Patients whose colon has been removed or partially removed are no longer at risk, or are at greatly reduced risk, for colectomy due to intractable disease or ACN. Based on consultation with the clinical specialists in UC, the amount of residual colon left after a partial colectomy is considered small enough so that the risk of a second colectomy for intractable UC or the risk to develop an ACN is greatly reduced. Therefore, the risk in patients who have had a partial colectomy is no longer comparable to the risk among patients with a complete colon.
- Apparent loss to follow-up. Clinical expectations are that patients receiving anti-TNF or immunomodulatory therapies will have regular follow-up visits with ENEIDA physicians at least every 6 months. Patients who appear to be receiving these agents but who have no recorded follow up visit for at least 13 months after last clinical contact will be deemed "lost to follow-up" (LTFU).
- Initiation of vedolizumab or any novel immunomodulatory agent that is newly marketed during the study period. Because these agents may have direct effect on study outcomes and are likely to be preferentially prescribed to UC patients with more active or severe disease. Including person-time exposed to these agents will very likely lead to intractable confounding and uninterpretable results. At the same time, because uncontrolled disease activity may prompt therapeutic changes, it is possible that censoring follow-up immediately after switching may actually mask effects of the drug that preceded the switch. For this reason, the date of censoring will be set 90 days after the start of vedolizumab or a novel immunomodulatory agent.

In the clinical care of patients with UC, therapies often change over time. Changes in therapy will not terminate follow up (except as noted immediately above), but patients' exposure status will change; exposure will be treated as a time-dependent variable.

9.2.3.2 Definition of risk windows

The definition of risk window associated with each course of treatment is described below. Because of possible different drug effects after drug discontinuation, the risk window for neoplasia outcomes differs from the colectomy outcome.

For outcome of colectomy for intractable disease: The risk window for all exposures of interest (including GLM, other anti-TNF) begins 1 day after exposure initiation through 90 days after the last treatment, or until one of the general triggers for end of follow-up occurs, whichever comes first. Because a treatment can theoretically affect risk of colectomy even long after its discontinuation, an alternative risk window will also be considered.

Alternative risk window: begins 1 day after exposure initiation through the end of follow-up, regardless of any subsequent switch in therapy.

For ACN outcome: For anti-TNF agents, the potential effect on risk of neoplastic outcomes may persist after discontinuation of the anti-TNF agent. As has been the practice in other studies of risk of cancer associated with the anti-TNF agents, the primary risk window for these exposures is "once exposed, always at risk".[13, 17, 18] Anticipating that treatments will be dynamic for many patients in this study, the following scheme will be used for attribution of CRC/HGD outcomes to specific exposures:

- Thiopurine use: the risk window begins at the start of thiopurine until end of follow-up or 6 months after drug discontinuation, whichever occurs first.
- For GLM and other anti-TNFs use, the risk window begins with the initiation of the agent and extends until end of follow up, regardless of drug discontinuation. If patients switch from one anti-TNF to another anti-TNF agent, subsequent person- years and events will be attributed to both agents.
- Patients exposed to both thiopurines and anti-TNF agents will accrue person-time to each exposure as noted above; a patient may thus contribute time at risk to more than one agent simultaneously.

Because the actual biologic risk window for neoplastic outcomes is not known, we will also evaluate alternative definitions of the risk window that lag the start of time at risk and vary the potential period at risk after discontinuation of exposure:

- Alternative scenario 1: For all study exposures, the risk window begins 6 months after start of current exposure and ends 6 months after discontinuation of exposure or at the end of follow-up whichever occurs first;
- Alternative scenario 2: For all study exposures, the risk window begins 6 months after start of current exposure and ends 2 years after discontinuation of exposure or at the end of follow-up whichever occurs first.

9.2.4 Identification of cases and controls for nested case-control analyses

In addition to the main cohort study, nested case-control analyses are planned to assess the associations between GLM and the two main outcomes of interest, colectomy for intractable disease and ACN. The motivation for this additional analytic approach is to reduce any residual confounding in analyses conducted exclusively on automated data. Some potential confounding variables are not well captured in the automated data, such as systemic steroid use and hospitalization for UC, both of which are potential proxies for disease activity. The nested-case control analyses will be limited to patients from ENEIDA sites that agree to participate in this chart review substudy.

From this subpopulation of the main study, automated data will be used to identify cases of colectomy for intractable disease and ACN arising from the pooled cohorts. For each case, up to two controls will be randomly selected from risk sets assembled using incidence density sampling, matching on calendar time of outcome and time since initial UC diagnosis. This means that controls will need to be alive and at risk for the outcome on the calendar date of the

case's diagnosis, which will serve as an index date for matched controls. The control's exposure status will be determined on that index date. Additional operational details concerning sampling of controls will be available in the statistical analytic plan (SAP), which will be finalized before any comparative analyses are undertaken.

9.3 Variables

Automated ENEIDA data will be used for initial variable ascertainment which will include UC disease diagnosis, treatment exposure, patient demographics, study outcomes and some covariates. Chart review will be performed on all cases and a sample of controls to review data about other drug exposures, confirm outcomes, and gather information about potential confounders incompletely recorded in the automated data. The definitions of exposure, outcomes, and covariates are presented below.

9.3.1 Exposure

The main exposures of interest in this study are treatment with GLM, other anti-TNF agents, and thiopurines. Data about treatment with anti-TNF therapies and thiopurines are considered mandatory variables on the ENEIDA CRF. Drug data also include information on start date and discontinuation date of each treatment course of these drugs. This permits capture of time-dependent information about key exposures.

The validity of ENEIDA data on exposure to study agents has never been formally validated against medical records. It should be emphasized, however, that these data are abstracted from the medical records directly by ENEIDA site staff. Indirect evidence of the validity of exposure classification is provided by the pilot study performed in preparation of this protocol. In that study, the profile of new users of anti-TNF agents was consistent with clinical expectations. Patients in the anti-TNF cohort appeared to be sicker than those in the thiopurine cohort, with a higher prevalence of extensive disease, use of prior methotrexate or cyclosporine, and longer disease duration. Most (57.3%) anti-TNF agent initiations occurred in the context of prevalent use of thiopurines. In a separate ENEIDA study, 95.8% of patients starting adalimumab for UC had received thiopurines at some point in the past.[16]

9.3.2 Outcomes

The study endpoints for this post-authorization safety study are colectomy for intractable disease, ACN (a composite of CRC or HGD), CRC and HSTCL. Only incident diagnoses (those with an onset after cohort entry) will qualify as outcomes.

- Colectomy due to intractable disease: Information about bowel surgery is considered mandatory data in ENEIDA. In the field for colectomy, the CRF specifically asks for indication. Categories include intractable disease ("refractoriness to medical treatment"), stenosis, perforation, hemorrhage, dysplasia or cancer, and "other". The date of colectomy is also available in the automated database. In addition, based on consultation with clinical specialists in UC, and in line with European guidelines[19], the following 2 scenarios are considered to meet the definition of colectomy due to intractable disease if no other reasons for colectomy were mentioned:

- Subtotal colectomy with terminal ileostomy among patients aged 65 years or older
- Subtotal colectomy with ileorectal anastomosis among women aged up to 50 years

In these 2 specific subpopulations (elderly people aged ≥65 years and women aged ≤50 years), subtotal (ie, partial) colectomy is the currently preferred surgical approach to treat resistant UC (more conservative approach).

- Advanced colorectal neoplasia (a composite endpoint that includes both colorectal cancer and high-grade dysplasia): The rationale for using a composite endpoint is that both constituents represent different phases of the same disease pathway. Current understanding of CRC is that all such cancers go through a dysplastic phase before malignant transformation. Patients with HGD have a high probability of progressing to CRC, and like patients with CRC, they are typically treated with colectomy. [10] A recent study of the natural history of IBD in France used the same composite outcome of CRC and HGD.[20] In the current study, diagnoses of CRC and HGD will be ascertained from several fields in the automated data.

Variable	Comments
Colectomy, with indication	Mandatory variable, date available
Neoplasia AE	Mandatory to report during treatment with biologic or immunomodulatory therapy. Type of tumor may not be specified
Colonoscopy findings	Helpful if present, but not mandatory
General neoplasia data	Helpful if present, but not mandatory

- Hepatosplenic T cell lymphoma: Information about HSTCL will be ascertained using two approaches, first from the comorbidity section of the CRF (a non-mandatory variable) and second, through CRF AE reports of neoplasia (a mandatory variable, but specification of tumor type is not mandatory in ENEIDA) that specify HSTCL. In addition, for all AE reports that mention lymphoma, sites will be queried to determine if it is HSTCL. The date of HSTCL will also be determined during the process of HSTCL identification, along with relevant clinical details. Cases of HSTCL may be difficult to capture because these rare lymphoid malignancies can conceivably be classified as lymphomas without further specification in the ENEIDA registry. To ensure that all reported cases of HSTCL in the ENEIDA registry are captured, registry data managers run a sensitive search for lymphoma-related diagnoses in the registry free-text fields on a yearly basis. The ENEIDA registry data managers then retrieve only the identification codes for the identified patients and provide them to the research team at RTI Health Solutions (RTI-HS). The RTI-HS study analysts check for the

presence of these identification codes in the GLM ENEIDA PASS database stored at RTI-HS. Any matches for patients in the ENEIDA PASS are followed up with the investigators at the corresponding hospitals for further case evaluation and chart abstraction, if applicable.

Although colectomy, CRC, and HGD are each captured on the ENEIDA CRF, it is possible that some occurrences of these clinical endpoints may escape detection in automated ENEIDA data. There is no mechanism to determine the extent to which this occurs. However, results from the ENEIDA pilot data analysis indicated that these outcomes are captured with reasonable completeness. In the pilot project, which was based on a retrospective cohort analysis of UC patients (2010-2013) who initiated a non-GLM anti-TNF agent or a thiopurine, the incidence rate of all-cause colectomy was 36.5/1000 PY. This value falls within in the range of corresponding rates reported in the literature (14.1/1000 PY in a population-based setting to 127/1000 PY in clinical trials evaluating the efficacy of infliximab in induction and maintenance of UC. [21, 22]). Consistent with the described epidemiology of UC, risk of colectomy was greatest in the first two years following diagnosis in this pilot data analysis.

The incidence rate of CRC from the pilot study, 1.8/1000 PY (95% CI 0.2 –6.4/1000 PY) was also consistent with rates of 1.24 – 4.02/1000 PY reported in the literature.[12]

9.3.3 Covariates

Baseline patient characteristics and covariates will be ascertained and evaluated as potential confounders in this study. This will include age, gender, disease extent, disease duration, medications use (e.g., systemic steroid and cyclosporine), hospitalization for UC, and selected comorbidities. Some covariates will be ascertained from automated registry data, while other elements will only be available from medical records, which will be reviewed as part of the NCC for a subset of study patients.

The concept of "disease activity" factors into the risk of each of the main outcomes, but it is important to recognize that the term refers to two related but distinct concepts. First, as is commonly used in clinical practice, disease activity refers to current biological colonic inflammatory activity, and it can be measured in several ways.[3, 23] In clinical practice, it is commonly categorized into remission, mild, moderate and severe UC, and depending on the measurement scheme, assessment may factor in stool frequency, rectal bleeding, vital signs, laboratory findings (hemoglobin concentration and acute phase reactants), bowel endoscopic appearance, and patient and physician global assessments. Disease activity is dynamic; UC is characterized by flares, remission, and relapses. Disease activity is a major determinant in choice of UC therapy, and is a major risk factor for colectomy for intractable disease. The second sense of disease activity refers to disease activity integrated over time, which can be conceptualized as "chronic inflammatory load", and is a major risk factor for advanced neoplasia.

In prospective clinical studies, current disease activity is often measured using the Mayo Score, which reflects several clinical inputs that do not appear in automated data and may not be systematically documented in usual clinical practice. However, several clinically reasonable proxies are available in automated data, including disease extent and evidence of recent

modifications in medications for UC. There is also no direct measurement of chronic inflammatory load in automated data or clinical records, but markers that have been used in prior research include disease duration, disease extent, and history of UC-specific treatments [10, 13, 24, 25]

Table 1 lists the covariates for this study.

Table 1. Covariates

Variable	Description	Source Automated data (AD)/ Medical records (MR)	Time-varying *	How treated in analysis	Comment
Patient identification number	Unique patient identification number assigned when patient is registered	AD	No		Patient identifier and ENEIDA site identifier from automated data will be used to link to patient's record for review
ENEIDA site identifier	Unique identifier for clinical center providing UC care for study patient	AD	No	Data quality from each center will be evaluated as a condition of study participation. Will not be considered as a confounder.	
Age	Depending on the analysis, age may defined as -Age at diagnosis (< 35 years, yes/no) -Age at study entry -Age at time switch into a new exposure category -Age at the time of an outcome (used	AD	Yes	Used to describe study cohorts at baseline. Potential confounder, but will likely be very collinear with disease duration	-UC onset at a young age has been described as a risk factor for colectomy -Older age is a risk factor for CRC

Variable	Description	Source Automated data (AD)/ Medical records (MR)	Time-varying *	How treated in analysis	Comment
	in analyses based on risk-sets)				
Gender	Male/female	AD	No	Stratification factor; potential confounder	
Date of study entry	Date when all cohort criteria have been met	AD	No	As an anchor for longitudinal analyses. Year of cohort entry may be used to evaluate any temporal trends during study.	
Date of first UC diagnosis		AD	No	Used to calculate disease duration	UC diagnosis may be entered retrospectively in registry data; the registry may enroll patients with prevalent disease
Disease duration	Depending on analysis, may be described at -time of cohort entry -time of switch into new exposure	AD	Yes	Used to describe cohorts at baseline. Potential confounder. In cohort analyses will be categorized depending on outcome (refer to section 9.7.1)	

Variable	Description	Source Automated data (AD)/ Medical records (MR)	Time-varying *	How treated in analysis	Comment
	group -time of an outcome (used in analyses based on risk-sets)				
Duration of exposure to study drugs	Defined as time from date of first dose to the date of its discontinuation if the drug is discontinued or end of study period for continuous users. Duration of interrupted courses of treatment will be summed for each exposure.	AD	Yes	To evaluate duration response, not for confounder adjustment	Duration of use of any other anti-TNFs will be summed across all treatment episodes for all non-GLM anti-TNF agents.
Extent of disease		AD; MR Automated data reflects maximal extent ever	In AD, no; in MR, yes	Categorized as (a) localized distal disease; (b) left sided only disease, (c) pancolitis; (d)	Risk factor for colectomy and ACN

Variable	Description	Source Automated data (AD)/ Medical records (MR)	Time-varying *	How treated in analysis	Comment
		recorded-values are not date stamped		unclassifiable/unknown. Used to describe cohorts at baseline and as potential confounder, dichotomized into "extensive" vs "not extensive"	
Treatments for UC not included as exposures: systemic steroids and cyclosporine		AD; supplemented with MR for NCC analysis	Yes	Prior treatment (Yes/no) with steroids and cyclosporine, as separate variables. Steroid refractory/dependent status will be obtained during chart review	Steroids and other immunosuppressants are markers of more active disease. In particular, cyclosporine is a marker for steroid-refractory disease
Hospitalized for UC	Notation of previous hospitalization a) Within 12 months after initial diagnosis b) Within previous 12 months	AD (incomplete recording); supplemented with MR for NCC analysis	a) No b) Yes	Dichotomous yes/no;	Hospitalization early in disease course is a sign of poor UC prognosis. Recent hospitalization is also a marker of active disease and worse prognosis for colectomy.

Variable	Description	Source Automated data (AD)/ Medical records (MR)	Time-varying *	How treated in analysis	Comment
	before entering study or switching to a new study exposure				
Primary sclerosing cholangitis (PSC)		AD; supplemented with MR for NCC analysis	Yes	Ever prior diagnosis yes/no at time of cohort entry or switch to new exposure group.	PSC is a risk factor for CRC
Number of previous anti-TNF agents	Measures history of use of other anti-TNF agents prior to current exposure	AD	Yes	Stratification factor and potential confounder.	
Recent switcher after short term use of another anti-TNF agent	Refers to the 90-day period after starting a new anti-TNF agent following discontinuation of another anti-TNF agent that was used for 3 months or less.	AD	Yes	Evaluate as a potential confounder or effect modifier for colectomy outcome	The period following a switch may represent effects of the acute illness that prompted the change in therapy, rather than effects of the new therapy. Patients with primary non-response have a higher risk of colectomy than those with

Variable	Description	Source Automated data (AD)/ Medical records (MR)	Time-varying *	How treated in analysis	Comment
					secondary loss of response. Short-term use of the first anti-TNF agent suggests primary non-response is the reason for discontinuation.
Recent switcher after longer- term use of another anti-TNF agent	Refers to the 90-day period after starting a new anti-TNF agent following discontinuation of another anti-TNF agent that was used for more than 3 months	AD	Yes	Evaluate as a potential time-dependent confounder or effect modifier for colectomy outcome	The period following a switch may represent effects of the acute illness that prompted the change in therapy, rather than effects of the new therapy. Patients with primary non-response have a higher risk of colectomy than those with secondary loss of response. Longer-term use of the first anti-TNF agent suggests secondary loss of response was the reason for discontinuation.

Variable	Description	Source Automated data (AD)/ Medical records (MR)	Time-varying *	How treated in analysis	Comment
Screening colonoscopy		AD; supplemented with MR	Yes	Will not use colonoscopy in confounder adjustment, as it may be on causal pathway	Will be used to characterize cohorts at baseline

AD, automated data; MR, medical records
Notes
*In the analysis, time-dependent covariates, such as drug treatment will be updated just before any switch into a new exposure category.
Following analysis of descriptive data (but before any comparative analyses), specification of current covariates may change, and other covariates may be added

9.4 Data Sources

The two data sources used in this study include: (1) the automated data from the ENEIDA registry; (2) clinical data abstracted from medical records. Data from the ENEIDA registry, which was founded in 2006, have been used in several studies published in the peer-reviewed literature on the epidemiology and treatment of IBD.[5, 16] Please refer to section 9.2.1 for a discussion of the general organization and process of data collection in ENEIDA. Additional information below describes the "mandatory" variables that ENEIDA sites are expected to record, as well as the system for collecting adverse event reports as part of the registry's normal operations (i.e., not for the purposes of the current study).

List of "mandatory" variables and calculation of variable completeness

ENEIDA sites are expected to complete a set of mandatory variables for each registered patient. On a monthly basis, ENEIDA assesses the completeness of data for each ENEIDA site that includes its census of registry patients, not just the patients seen in clinic that month. The "mandatory" variables are grouped into the following nine categories, or sections:

1. Demographic data: name, family name, date of birth, gender, date of inclusion in the registry, date of the last visit.

2. Clinical data: date of the first diagnosis (UC), current diagnosis (in case IBD diagnosis changed since presentation), disease location.

3. Immunosuppressive therapy (yes or no).

4. Immunosuppressive therapy details: drug type, date of drug initiation, continuation of treatment (yes/no), drug indication, efficacy adverse events (including type and if it was necessary to stop the drug). (It is important to note that the ENEIDA steering committee has indicated that efficacy data will not be available for the current company-sponsored study).

5. Biologic therapy (yes or no).

6. Biologic therapy details: drug name, date of drug initiation, continuation of treatment (yes/no), drug indication, efficacy (which ENEIDA will not make available for a company-sponsored analysis), secondary loss of response, adverse events (including type and if it was necessary to stop the drug).

7. Surgical therapy (yes or no).

8. Details of surgical therapy: date of surgery, type (urgent/programmed), type of surgery (abdominal or perianal), indication for surgery (list).

9. Risk factors: family history, smoking habit (at present and at the moment of diagnosis).

Calculations of percent complete include all patients registered within a site, except for sections 4, 6, and 8, which include only patients who received the relevant immunosuppressive, biologic, or surgical therapies. Note that if a single variable in a specific section is missing, the entire section is considered to be incomplete. Thus each site receives scores for completeness of nine separate sections of mandatory data.

Adverse event reports in ENEIDA

The ENEIDA CRF includes fields to indicate AEs as a limited, fixed selection of safety outcomes, and this list differs between biologic agents (including anti-TNF agents) and immunomodulators (including thiopurines) (Table 2). While reporting these safety events for patients receiving biologic agents or immunomodulators is a requirement for practices that wish to participate in ENEIDA studies, the extent of underreporting of AEs has not been evaluated. It should be noted that the AE report form does not include any clinical details. The requested information is in the form of yes/no responses and whether the AE prompted therapy discontinuation; there is also a free-text field for events that are not specified in the AE list ("Other ____"). AEs reported in association with immunomodulators are date-stamped. In contrast, AEs reported in association with biologic agents do not indicate a specific date of onset, but presumably began after initiation of the biologic agent and before any decision to discontinue therapy.

AEs collected on ENEIDA's CRF are entered into the ENEIDA database during the routine course of the registry. AE collection and reporting for the current study are described in Section 11.

Table 2 List of AEs solicited in ENEIDA by drug class

Biologics (including anti-TNF agents)	Immunomodulators (including thiopurines)
• Infection • Infusion reaction • Heart failure • Late hypersensitivity reaction • Neurologic disease • Edema • Anaphylaxis • Auto-antibodies • Neoplasm • Other _____	• Alopecia • Anemia • Aplasia • Hepatotoxicity • Hyperglycemia • Hypertricosis • Infection • Renal insufficiency • Leucopenia • Nausea/ vomiting • Neoplasia • Neuropathy • Pancreatitis • Other _____

9.4.1 Study Procedures

This study involves a secondary analysis of the ENEIDA registry database, which documents patient care in a usual care setting. Accordingly, there are no study-related procedures required for this PASS.

This study has been approved by the ENEIDA Steering Committee. The ongoing ENEIDA registry has been approved by participating institutional review boards (IRBs), and patients whose data comprise ENEIDA have all provided individual consent.

In addition to the analysis of automated data, this study involves a substudy that involves a targeted chart review. The chart review portion of the protocol may require additional IRB reviews at the level of the Spanish autonomous health units and by the hospitals associated with participating sites, commensurate with local law.

Individual patient-level data will remain entirely within the ENEIDA environment, and only ENEIDA study personnel will have access to abstracted data from medical charts and will conduct data analyses. Thus, only aggregated analysis results and study reports will be shared with the Sponsor.

9.5 Study Size

This study will use all available data from qualified ENEIDA sites for patients with UC who initiate GLM, other anti-TNFα agents, and thiopurines during the study period. The number of patients who will qualify for this study is not known in advance and will depend on prescribing practices in ENEIDA practices over the course of the study.

No hypotheses have been specified for testing as part of this study, and thus no formal power calculations have been conducted.

Results of the pilot study conducted in 2014 provide some information for estimating the size of the actual PASS and the thresholds of detectable relative risk of the main study outcomes. Estimation of the detectable relative risk for colectomy for intractable disease involved several assumptions.

In the pilot study, approximately 100 new treatment courses of a non-GLM anti-TNF were initiated each year during the period 2010-2013, with a slight upward trend during the observation period. Extrapolating this trend, we project 996 new courses of anti-TNF therapy over the 8-year study period, which will include GLM as well as comparator anti-TNF therapies. The base-case risk window for the colectomy outcome is tied to treatment duration plus 90 days; for this exercise we assumed each treatment course contributed an average 1.7 PY, for a cumulative 1700 PY at risk across all anti-TNF use.

Several scenarios were considered that varied the proportion of total anti-TNF exposure that due to GLM (Table 3). Based on the pilot study, the expected incidence rate of colectomy was 33.8/1000 PY. The alpha level was set at 5%, and power was set at 80%. Detectable relative risk estimates range from 1.9 through 2.2.

Table 3

Percent of anti-TNF exposure from GLM	GLM time at risk (PY)	Other anti-TNF time at risk (PY)	Detectable relative risk
15	255	1443	2.2
20	340	1358	2.1
25	424	1273	2.0
30	509	1188	1.9
35	594	1103	1.9
40	679	1019	1.9

A corresponding exercise was performed to estimate the detectable relative risk for the ACN outcome. Calculations for this outcome were based on 3150 PY at risk to all anti-TNF therapies, assuming that time at risk extends until the end of follow-up, regardless of drug discontinuation (see section 9.2.3.2), and that 10% of the study population will be lost to follow up each year. Based on an ACN incidence rate of 1.8/1000 PY observed in the pilot study, detectable relative risks ranged from 5.6 to 7.4 as the proportion of follow-up time exposed to GLM varies from 40 to 15%.

9.6 Data management

This study involves a secondary analysis of the ENEIDA registry data, previously described in sections 9.2.1 and 9.4. The automated data are maintained and managed by the ENEIDA registry in a password-protected mySQL database, which is maintained on a server in a secure facility. The investigator or qualified designee is responsible for verifying the accuracy of patient data as it is entered into the ENEIDA database, Analyses will be conducted using SPSS software.

This study involves chart review for all cases with a primary outcome, as identified in the automated data, and a sample of controls. Linkage to between the automated record and the patient record will be based on the patient identification number and clinical site identifier.

A separate database will be created for this study, details of which will appear in a stand- alone Data Management Procedures manual. This manual will also include procedures for contacting sites for chart reviews and describe flow of abstracted clinical data into the study database. It will also include the chart abstraction instrument, which will be piloted during the first phase of the study.

9.7 Data analysis

This section outlines the analytic plans to address the study objectives. The final analytic approach will be described in a stand-alone statistical analytic plan (SAP), which will be completed before any comparative analyses are undertaken. All analyses other than those for the NCC will rely exclusively on automated data. With each update of analysis and report, ENEIDA refreshes the entire study data set; counts of cohort entries reported for earlier years may change due to retrospectively updated data.

9.7.1 Descriptive analyses

1. Describe assembly of study population, including attrition figures that indicate why potential subjects did not qualify for study.

2. Describe baseline patient baseline characteristics of each cohort. Analyses will use standard descriptive statistics including relevant measures of central tendency and variability, such as mean, standard deviation, median, interquartile range, and range for continuous variables. Number and proportions will be presented for categorical variables. *This step addresses Primary objective 1.*

3. Describe duration of follow-up (as defined in section 9.2.3) for the study population, stratified by cohort and year of study entry.

4. Tabulate reasons for premature end of follow-up, stratified by cohort and year of cohort entry.

5. Describe patterns of persistence and changing of study medications over time.

6. Describe the frequency of each study outcome, stratified by cohort (does not factor potential changes in drug exposure over time; these results will be reported in each progress report).

7. Describe incidence rates of study outcomes (colectomy for intractable disease, ACN, CRC, and HSTCL) along with 95% confidence intervals (CIs). Incidence rates will be defined as the number of events divided by the person-years at risk *This step addresses Primary objectives 1 and 2, secondary objective 1, and the exploratory objective.*
    a. For the GLM and other anti-TNF cohorts, calculate incidence rates for colectomy for intractable disease *(primary objective 2),* ACN *(primary objective 3)* and CRC *(secondary objective 1).*
    b. For the thiopurine cohort, calculate incidence rates for ACN *(primary objective 3)* and CRC *(secondary objective 1)*.
    c. For each cohort, calculate incidence rates of HSTCL *(exploratory objective).*
    d. Calculate the incidence rates of confounding for intractable disease and ACN for the following subgroups:
        i. GLM and other anti-TNF initiators with and without concurrent thiopurine use as measured at baseline.
        ii. Initiators of infliximab (or biosimilar infliximab) and adalimumab (or biosimilar adalimumab).
    e. For annual reports, incidence rates will be calculated based on the exposures that defined cohort entry (time-fixed analysis) regardless of subsequent discontinuation or switching.

8. Describe cumulative incidence of study outcomes using survival analysis (time-to- event) by cohort, including Kaplan-Meier plots with time since cohort entry as primary time axis. Estimate incidence of all primary and secondary study endpoints for the GLM and other anti-TNF cohorts. Estimate only incidence of ACN and CRC for thiopurine cohort.
    a. For the primary outcomes, (colectomy for intractable disease and ACN), repeat analyses stratified according to the following factors (as measured at study entry) one at a time:
        i. gender
        ii. time since initial UC diagnosis (cut points described below)
        iii. disease extent (extensive vs other)
        iv. history of UC hospitalization or cyclosporine use (yes/no)

v. previous use of systemic steroids;

b. Cut points for disease duration stratified analysis will depend on endpoint.
   i. Given a priori expectation that colectomy risk will vary inversely with time since initial UC diagnosis and be most pronounced shortly after diagnosis, stratify incidence of colectomy for intractable disease according to time since UC diagnosis (<2 y, ≥2 y).
   ii. Given a priori expectation that ACN will vary since time of UC diagnosis, stratify incidence rate of ACN into 3 categories of disease duration (<10 y; 10 to <20 y; 20+ years)

c. For anti-TNF cohorts only, repeat analyses for primary outcomes for the following subgroups:
   i. concurrent use of thiopurines (yes/no)
   ii. number of previous anti-TNF agents (0, 1, >1)
   iii. prior exposure to infliximab and adalimumab.

9.7.2 Comparative analyses

9.7.2.1 Comparative cohort analyses

GLM compared to other anti-TNF exposure
We will use survival analysis (Kaplan Meier plots and Cox proportional hazards regression models) to estimate the relative risk of primary and secondary outcomes comparing exposure to GLM to exposure to other anti-TNF agents. Analyses will use time since cohort entry as the primary time axis. Study exposures will be treated as time-dependent variables, following base-case risk windows described in section 9.2.3.2. Time-dependent covariates will be updated based on most current available data prior to switch to a new exposure group (e.g., from infliximab to GLM). Hazard ratios will be used to estimate relative risk and will be adjusted for confounding factors. In addition, Poisson regression will also be used for analysis of colectomy due to intractable disease (see details in SAP). *These analyses address primary objectives 4 and 5 and secondary objective 2.*

Candidate variables to evaluate as potential confounders are listed in Table 1. Adjustment for confounding will occur through stratification and multivariate adjustment. For each outcome, a separate Cox regression model will be generated for each candidate variable of interest where time from date of cohort entry to each outcome or censoring will be modeled as a function of the candidate variable only. The resulting hazard ratio and 95% CI for the candidate variable will be reported. Variables will ultimately be selected if they exhibit a hazard ratio greater than a threshold (i.e., 1.25) or less than its inverse (e.g., 0.80). Additionally, other variables may be forced into the model if deemed clinically important. For the colectomy due to intractable disease outcome only, the same series of univariable models will also be generated using the Poisson regression approach (see details in SAP). The ability to build a multivariate model will likely be limited by the number of outcomes observed. Priority for inclusion in the final model will be based on the magnitude of change in estimate between crude and adjusted models that evaluate each potential confounder one at a time. We will use results from stratified descriptive analyses to suggest potential effect modifiers. Effect modification will be evaluated based on the p-value of the interaction coefficient, with p-value of 0.05 as the threshold of statistical significance.

Subgroup analyses
If a sufficient number of exposures observed, subgroup analyses will be conducted to evaluate anti-TNF agent exposures grouped according to concomitant thiopurine use, history of prior anti-TNF agent use, and specific comparator anti-TNF-agent infliximab and adalimumab.

Sensitivity analyses

1. Comparative cohort analyses will be repeated to evaluate any effect of alternative specification of risk windows, as noted in section 9.2.3.2.
2. Comparative cohort analyses for the primary outcomes will be repeated with a different anchor for the primary time axis, using calendar date (starting with 19 September 2013) instead of time since cohort entry.
3. Sensitivity analyses will be conducted to evaluate the potential effect of competing risks*:* During the clinical course of UC, the risk of developing ACN will compete with the risk of having a colectomy, and vice versa. Colectomy largely precludes subsequent risk for ACN, and ACN very often leads to treatment by colectomy, which precludes the risk of colectomy for intractable disease. Competing risks can bias estimates of cumulative incidence. Competing risk models have been described to address this concern, but they will likely not be feasible in the current study given the limited number of outcomes expected. Instead, we will address the potential role of competing risk indirectly, in a sensitivity analysis that describes overall event-free survival, using a composite definition of outcome that includes ACN, colectomy (for any cause), or death.

GLM compared to thiopurine exposure (secondary comparator group)
A secondary cohort analysis may be conducted contrasting the risk of ACN (and CRC) between GLM and thiopurine use, if inspection of baseline characteristics suggests that the groups are reasonably comparable (*secondary objective 3*). A priori there is concern that the distribution of prognostic factors for ACN may differ substantially between GLM and thiopurines initiators, a situation that could give rise to intractable confounding. In the treatment of UC, anti-TNF agents are usually introduced after thiopurines, which suggests that thiopurine users may have a lower burden of chronic bowel inflammation, a major determinant of ACN risk. As noted in section 9.2.2, a study conducted using ENEIDA data found that 95.8% of patients starting adalimumab for UC had received thiopurines at some point in the past. The specific approach to gauge comparability between cohorts will be included in the SAP. If they are deemed comparable, comparative analyses will proceed as described for the GLM/other anti-TNF contrast.

9.7.2.2 Comparative analyses using the nested case-control approach

Separate nested case-control analyses are planned to evaluate the association between study exposures and the two primary outcomes, colectomy for intractable disease and ACN. Control selection will follow the process described in section 9.2.4. Additional details on selection of controls will appear in the SAP.

Cases and controls will first be described according to prevalence of exposures and covariates

at the time of cohort entry. In separate analyses, we will use conditional logistic regression to calculate the odds ratios of colectomy for intractable disease and ACN. Confounding adjustment will occur through both matching and adjustment. Matching factors include calendar time and duration of UC. Potential confounders are listed in Table 2. Separate models will be developed for each primary outcome. Given the limited number of outcomes anticipated, the ability to build a multivariate model will likely be constrained. Candidate variables to evaluate as potential confounders include those mentioned in Table 1. Priority for inclusion in the final model will be based on percent change in estimate between crude and adjusted models that evaluate each potential confounder one at a time. A NCC analysis will not be conducted for outcomes with fewer than 10 events, as it would not support adjustment for even one covariate.

Sensitivity analyses. To the extent that data permit, alternative analyses will explore the potential effect of treatment duration, alternative specifications of treatment combinations over time, and the sensitivity of results to alternative specifications of the risk window.

9.8 Quality Control

ENEIDA's system to assess data completeness was described in section 9.4. Data entered remotely at ENEIDA clinical sites are transmitted to the ENEIDA server using a secure connection (VPN, SSL).

Several variables in ENEIDA are subject to logical checks at the time of data entry, and forms cannot be saved if they include implausible values. For example, all dates must fall between 1 January 1900 and the current date. Several range checks exist for variables that that are represented as integers; acceptable values depend on the clinical context. During the preparation of the analytic file for this study, additional quality control measures will include, but will not be limited to, a check for legitimate values for each categorical variable, and logical checks for dates (e.g. date of birth precedes all other dates).

By signing this protocol, all parties agree to following applicable standard operating procedures (SOPs). All parties also agree to ensuring all existing and new study personnel are appropriately trained to ensure the study is conducted and data are generated, documented, and reported in compliance with the protocol, Good Pharmacoepidemiology Practice (GPP), and all applicable federal, state, and local laws, rules and regulations. All parties should maintain transparency and open communication in order to effectively manage the study and proactively mitigate any risks.

The Sponsor may conduct routine or for-cause audits to ensure oversight and conduct of the study are completed in accordance with the protocol, quality standards (e.g. GPP), and applicable laws and regulations. If a significant quality issue (SQI) is identified at any time during the conduct of the study, it must be escalated to the Sponsor immediately. A SQI is any issue with the potential to negatively impact, either directly or indirectly, the rights, safety and well-being of patients or study participants and/or the integrity of the data. In the event an audit or SQI results in corrective or preventive actions, all parties are expected to appropriately implement the action plan in a timely manner.

9.9 Limitations of the research methods

- Potential for information bias. The automated ENEIDA data may not contain complete information on all relevant study variables. Many fields are not mandatory for ENEIDA sites to complete, and based on pilot work, one expects that the true frequencies of certain medical events (e.g., 5-ASA use and hospitalizations) will be underestimated. However, the pilot study also indicated that capture of the main study outcomes was reasonably complete. Chart review is planned for all cases and a sample of controls to augment capture of information on potential confounding variables.

- Confounding by severity of disease is a concern in any observational safety study. Disease activity cannot be directly ascertained from the automated data, and it is not systematically recorded in medical records. However, proxies for chronic inflammatory burden, such as disease duration, disease extent, and history of prior UC therapies, are systematically recorded in the automated data. Moreover, other markers of disease activity, such as steroid treatment history and hospitalization for UC will be sought from medical records for the nested case-control analysis.
- Residual confounding may occur not just because of incomplete information on potential confounding variables but also because of the rarity of some study outcomes. The study outcomes are expected to occur rarely—especially ACN—and sparse outcomes limits the ability to perform multivariate statistical adjustment
- The structure of the database may also limit our ability to address certain study questions. Disease extent is not stored as a time-dependent variable; rather it is continually updated and at any point reflects the maximal disease extent ever observed. As a consequence, analyses based on automated data may exaggerate the maximal disease extent at the time of cohort entry. As part of the case-control study, the maximal disease extent at time of cohort entry will be ascertained from medical records. Similarly, dates of clinic visits are not stored in a longitudinal fashion. The database tracks only the date of the latest clinic visit, and as a result, our ability to address the completeness of follow-up care is constrained.
- Loss to follow up is anticipated to be approximately 10% per year based on the pilot study. If loss to follow up differs by exposure status, it could introduce selection bias. Baseline characteristics of patients who are lost to follow up will be compared to evaluate this possibility.
- The outcomes of interest, especially advanced colonic neoplasia, occur rarely, with an expected incidence rate in the order of 1/1000 PY. We therefore anticipate that observed incidence rates of these events will have limited precision. As a result, there is limited precision to estimate the relative risk of advanced colorectal neoplasia.
- Because the background incidence of HSTCL is so low (in the order of 1 case per million PY), it is likely that few if any HSTCL cases will be encountered during this study. As a consequence, confidence limits around this estimate will be wide.

9.10 Other Aspects

None.

10 PROTECTION OF HUMAN SUBJECTS

10.1 Informed Consent

This non-interventional observational study involved no administration of any therapeutic agent according to the study protocol; patients in ENEIDA are treated in the setting of usual clinical care. Patients whose data are included in ENEIDA previously consented to have their health information included in this database. No new consent will need to be obtained for this PASS.

11 MANAGEMENT AND REPORTING OF ADVERSE EVENTS/ADVERSE REACTIONS

This is a non-interventional study within routine medical practice. All decisions about medical therapy are made by patients' physicians in accordance with their usual practice. No individual administration of any therapeutic or prophylactic agent is required in this protocol, and there are no procedures required as part of this protocol.

This section describes procedures for the collection, management and reporting of individual cases of adverse events/adverse reactions and of any new information that might influence the evaluation of the benefit-risk balance of the product while the study is being conducted.

11.1 Definitions

11.1.1 Adverse Event (AE)

Any untoward medical occurrence in a patient or clinical investigation subject administered sponsor's product and which does not necessarily have to have a causal relationship with this product. An adverse event can therefore be any unfavorable and unintended sign (including an abnormal laboratory finding, for example), symptom, or disease temporally associated with the use of the product, whether or not considered related to the product. Any worsening (i.e., any clinically significant adverse change in frequency and/or intensity) of a preexisting condition that is temporally associated with the use of the product, is also an adverse event.

11.1.2 Adverse Reaction (AR)

An AE which has a causal relationship with the product, that is, a causal relationship between the product and the adverse event is at least a reasonable possibility.

11.1.3 Serious Adverse Event (SAE)/Serious Adverse Reaction (SAR)

An adverse event or adverse reaction that results in death, is life threatening, results in persistent or significant disability/incapacity, requires inpatient hospitalization, prolongation of existing inpatient hospitalization, is a congenital anomaly/birth defect, or is another important medical event. Other important medical events that may not result in death, may not be life-threatening, or may not require hospitalization may be considered an SAE/SAR

when, based upon appropriate medical judgment, they may jeopardize the patient or subject and may require medical or surgical intervention to prevent one of the other outcomes listed previously. Examples of such medical events include allergic bronchospasm requiring intensive treatment in an emergency room or at home and blood dyscrasias or convulsions that do not result in inpatient hospitalization.

11.1.4 Non-serious Adverse Reaction (NSAR)

An adverse reaction that does not meet any of the serious criteria in 6.1.3.

11.1.5 Special Situations

The following special situations are considered important safety information and must be reported, regardless of causality:

- Overdose
- Exposure to product during pregnancy or lactation
- Off-label use, medication error, misuse, abuse, or occupational exposure
- Suspected transmission via a medicinal product of an infectious agent

Although "lack of therapeutic effect" is considered a special situation according the GVP Module VI, the ENEIDA steering committee has stipulated that data on drug effectiveness will not be available for this study.

11.1.6 Health Outcome of Interest (HOI)

Health Outcomes of Interest (HOIs) are clinical events or outcomes that are collected according to the protocol as the focus of the study. The HOIs for this study are colectomy for intractable disease, ACN, and HSTCL. During this study, medical record review is planned for all cases of HOIs identified through automated data. HOIs must be assessed as part of AE collection and may meet criteria for AE reporting. Specifically, the investigator must assess each HOI for serious criteria and causality. If the HOI meets criteria specified in the protocol for AE reporting, then it must be reported as such.

11.1.7 Sponsor's product

For the purposes of this study, Sponsor's product includes Remicade (infliximab) and Simponi (golimumab).

11.1.8 Causality Assessment

A causality assessment is the determination of whether or not there is at least a reasonable possibility that a product caused the adverse event. Causality is determined by an explicit notation of a causal relationship between an AE and a product in the medical records or other secondary data being reviewed. Causality should not be inferred by a temporal relationship

between the product administration and an AE, but must be based on a definite statement of causality by a healthcare provider linking the product to the AE.

11.2 Adverse Event Reporting

Although adverse events are not actively solicited in this study, there are certain circumstances in which individual adverse events will be reported. Specifically, during review of medical records to collect data as required by the protocol, if a notation of a serious adverse reaction (SAR), including death, or a non-serious adverse reaction (NSAR) to Sponsor's products is identified, the event must be reported.

HOIs that meet criteria for adverse reactions (SAR or NSAR) must also be reported. Finally, all occurrences of special situations about which the investigator becomes aware should be reported. For each reportable event, the investigator will complete an Adverse Event Report Form in English and submit SARs within 24 hours and NSARs within 10 calendar days to the Sponsor or designee, as specified in the Safety Management Plan (Stand Alone Document), which will be distributed to all investigators before the start of data collection. The sponsor and the MAH will submit qualifying reports to Regulatory Agencies worldwide, as appropriate.

The data captured in the medical record will constitute all clinical information known regarding these adverse events. No follow-up on these adverse events will be conducted. Research staff is not required to search sections of patient medical records not related to study objectives for the purpose of identifying AEs.

All SARs and NSARs for Sponsor's products will be entered in the study database for tabulation in the study report. The final study report, and any planned interim analysis, will include aggregate listings of all SARs and NSARs collected for Sponsor's products and will be provided to regulatory agencies by the sponsor as required.

Any relevant safety information will be summarized in the appropriate Periodic Safety Update Report (PSUR)/Periodic Benefit Risk Evaluation Report (PBRER) and/or Development Safety Update Reports (DSUR) as required.

12 PLANS FOR DISSEMINATING AND COMMUNICATING STUDY RESULTS

Progress reports will be submitted to the EMA at several points during the study, according to the schedule under Milestones (Section 6). Each progress report will include information about the number of patients who have entered the study in each of the treatment cohorts (the GLM cohort, the other anti-TNF cohort, and the thiopurine cohort), cumulative follow-up time accrued in each cohort, and counts of each study outcome of interest tabulated by exposure status at cohort entry. All other analyses will only appear in the final study report which will be submitted within 12 months of the end of data collection.

Results of the study will be submitted to a peer-reviewed journal for publication. The principal investigator has the right to independently prepare that publication. The Sponsor is entitled to

review the results and interpretations included in the manuscript and provide comments prior to submission of the manuscript for publication.

REFERENCES

1. Sandborn WJ, Feagan BG, Marano C et al. Subcutaneous golimumab induces clinical response and remission in patients with moderate-to-severe ulcerative colitis. Gastroenterology 2014;146(1):85-95.

2. Sandborn WJ, Feagan BG, Marano C et al. Subcutaneous golimumab maintains clinical response in patients with moderate-to-severe ulcerative colitis. Gastroenterology 2014;146(1):96-109.

3. Dignass A., Eliakim R, Magro F, et al. Second european evidence-based consensus on the diagnosis and management of ulcerative colitis part 1: definitions and diagnosis. J crohns colitis 2012;6(10):965-990.

4. Solberg IC, Lygren I, Jahnsen J et al. Clinical course during the first 10 years of ulcerative colitis: results from a population-based inception cohort (ibsen study). Scand j gastroenterol 2009;44(4):431-440.

5. Canas-Ventura A, Marques L, Ricard E, et ca. Risk of colectomy in patients with ulcerative colitis under thiopurine treatment. J crohns colitis 2014;8(10):1287-1293.

6. Filippi J, Allen Pb, Hebuterne X, Peyrin-Biroulet L. Does anti-tnf therapy reduce the requirement for surgery in ulcerative colitis? A systematic review. Curr drug targets 2011;12(10):1440-1447.

7. Reich Km, Chang HJ, Rezaie A et al. The incidence rate of colectomy for medically refractory ulcerative colitis has declined in parallel with increasing anti-tnf use: a time-trend study. Aliment pharmacol ther 2014;40(6):629-638.

8. Sandborn WJ, Rutgeerts P, Feagan BG et al. Colectomy rate comparison after treatment of ulcerative colitis with placebo or infliximab. Gastroenterology 2009;137(4):1250-1260.

9. Loftus EV, Jr. Epidemiology and risk factors for colorectal dysplasia and cancer in ulcerative colitis. Gastroenterol clin north am 2006;35(3):517-531.

10. Van Assche G, Dignass A, Bokemyer B, et al. Second european evidence-based consensus on the diagnosis and management of ulcerative colitis part 3: special situations. J crohns colitis 2013;7(1):1-33.

11. Meier J, Sturm A. Current treatment of ulcerative colitis. World j gastroenterol 2011;17(27):3204-3212.

12. Castano-Milla C, Chaparro M, Gisbert JP. Systematic review with meta-analysis: the declining risk of colorectal cancer inulcerative colitis. Aliment pharmacol ther 2014;39(7):645-659.

13. Andersen NN, Pasternak B, Basit S et al. Association between tumor necrosis factor-(alpha) antagonists and risk of cancer in patients with inflammatory bowel disease. Jama 2014;311(23):2406-2413.

14. Deepak P, Sifuentes H, Sherid M, Stobaugh D, Sadozai Y, Ehrenpreis ED. T-cell non-hodgkin's lymphomas reported to the fda aers with tumor necrosis factor-alpha (tnf-alpha) inhibitors: results of the refurbish study. Am j gastroenterol 2013;108(1):99-105.

15. Herrinton LJ, Liu L, Abramson O, Jaffe ES. The incidence of hepatosplenic t-cell lymphoma in a large managed care organization, with reference to anti-tumor necrosis factor therapy, northern california, 2000-2006. Pharmacoepidemiol drug saf 2012;21(1):49-52.

16. Garcia-Bosch O, Gisbert Jp, Canas-Ventura A, et al. Observational study on the efficacy of adalimumab for the treatment of ulcerative colitis and predictors of outcome. J crohns colitis 2013;7(9):717-722.

17. Askling J, Baecklund E, Granath F et al. Anti-tumour necrosis factor therapy in rheumatoid arthritis and risk of malignant lymphomas: relative risks and time trends in the swedish biologics register. Ann rheum dis 2009;68(5):648-653.

18. Strangfeld A, Hierse F, Rau R et al. Risk of incident or recurrent malignancies among patients with rheumatoid arthritis exposed to biologic therapy in the german biologics register rabbit. Arthritis res ther 2010;12(1):r5.

19. Magro F, Gionchetti P, Eliakim R, et al. Third European Evidence-based Consensus on Diagnosis and Management of Ulcerative Colitis. Part 1: Definitions, Diagnosis, Extra-intestinal Manifestations, Pregnancy, Cancer Surveillance, Surgery, and Ileo-anal Pouch Disorders. J Crohns Colitis. 2017;11(6):649-670

20. Beaugerie L, Svrcek M, Seksik P et al. Risk of colorectal high-grade dysplasia and cancer in a prospective observational cohort of patients with inflammatory bowel disease. Gastroenterology 2013;145(1):166-175.

21. Samuel S, Ingle SB, Dhillon S et al. Cumulative incidence and risk factors for hospitalization and surgery in a population-based cohort of ulcerative colitis. Inflamm bowel dis 2013;19(9):1858-1866.

22. D'haens G, Sandborn WJ, Feagan BG et al. A review of activity indices and efficacy end points for clinical trials of medical therapy in adults with ulcerative colitis. Gastroenterology 2007;132(2):763-786.

23. Jess T, Rungoe C, Peyrin-Biroulet L. Risk of colorectal cancer in patients with ulcerative colitis: a meta-analysis of population-based cohort studies. Clin gastroenterol hepatol 2012;10(6):639-645.

24. Chaparro M, Ordas I, Cabre E et al. Safety of thiopurine therapy in inflammatory bowel disease: long-term follow-up study of 3931 patients. Inflamm bowel dis 2013;19(7):1404-1410.
25. Julia A, Domenech E, Chaparro M et al. A genome-wide association study identifies a novel locus at 6q22.1 associated with ulcerative colitis. Hum mol genet 2014;23(25):6927-6934.

ANNEX 1. LIST OF STAND-ALONE DOCUMENTS

The documents listed are available upon request.

No.	Date	Title
1.	5 January 2015	Pilot Data Analysis report
2.	13 June 2017	Data management plan
3.	25 June 2020	Statistical analysis plan
4.	20 July 2017	Safety reporting management plan
5.	4 February 2022	AE reporting form
6	31 May 2022	List of investigators
7	15 May 2015	Administrative and regulatory details: additional information

ANNEX 2. ENCEPP CHECKLIST FOR STUDY PROTOCOLS

Study title:

A Post-Authorization Safety Study of Golimumab in UC Using the Spanish ENEIDA Registry

Study reference number:

MK8259-042-01

Section 1: Milestones	**Yes**	**No**	**N/A**	**Page Number(s)**
1.1 Does the protocol specify timelines for				
1.1.1 Start of data collection[1]	☒	☐	☐	14
1.1.2 End of data collection[2]	☒	☐	☐	14
1.1.3 Study progress report(s)	☒	☐	☐	14
1.1.4 Interim progress report(s)	☒	☐	☐	14
1.1.5 Registration in the EU PAS register	☒	☐	☐	
1.1.6 Final report of study results.	☒	☐	☐	14

Comments:

Study was registered in the EU PAS register

Section 2: Research question	**Yes**	**No**	**N/A**	**Page Number(s)**
2.1 Does the formulation of the research question and objectives clearly explain:				
2.1.1 Why the study is conducted? (e.g. to address an important public health concern, a risk identified in the risk management plan, an emerging safety issue)	☒	☐	☐	8,14
2.1.2 The objective(s) of the study?	☒	☐	☐	8,16
2.1.3 The target population? (i.e. population or subgroup to whom the study results are intended to be generalised)	☒	☐	☐	18
2.1.4 Which formal hypothesis(-es) is (are) to be tested?	☐	☐	☒	
2.1.5 If applicable, that there is no *a priori* hypothesis?	☒	☐	☐	1,8,16

Comments:

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

[2] Date from which the analytical dataset is completely available.

Section 3: Study design	Yes	No	N/A	Page Number(s)
3.1 Is the study design described? (e.g. cohort, case-control, randomized controlled trial, new or alternative design)	☒	☐	☐	9,17,18,22
3.2 Does the protocol specify the primary and secondary (if applicable) endpoint(s) to be investigated?	☒	☐	☐	16,17
3.3 Does the protocol describe the measure(s) of effect? (e.g. relative risk, odds ratio, deaths per 1000 person-years, absolute risk, excess risk, incidence rate ratio, hazard ratio, number needed to harm (NNH) per year)	☒	☐	☐	12,39-42

Comments:

Section 4: Source and study populations	Yes	No	N/A	Page Number(s)
4.1 Is the source population described?	☒	☐	☐	18-19
4.2 Is the planned study population defined in terms of:	☒	☐	☐	10,18,19
4.2.1 Study time period?	☒	☐	☐	10,18,19
4.2.2 Age and sex?	☒	☐	☐	18,19
4.2.3 Country of origin?	☒	☐	☐	10,18,19
4.2.4 Disease/indication?	☐	☒	☐	
4.2.5 Co-morbidity?	☐	☒	☐	
4.2.6 Seasonality?				
4.3 Does the protocol define how the study population will be sampled from the source population? (e.g. event or inclusion/exclusion criteria)	☒	☐	☐	18-20

Comments:

Section 5: Exposure definition and measurement	Yes	No	N/A	Page Number(s)
5.1 Does the protocol describe how exposure is defined and measured? (e.g. operational details for defining and categorising exposure)	☒	☐	☐	10,23
5.2 Does the protocol discuss the validity of exposure measurement? (e.g. precision, accuracy, prospective ascertainment, exposure information recorded before the outcome occurred, use of validation sub-study)	☒	☐	☐	23
5.3 Is exposure classified according to time windows? (e.g. current user, former user, non-use)	☒	☐	☐	19

Section 5: Exposure definition and measurement	Yes	No	N/A	Page Number(s)
5.4 Is exposure classified based on biological mechanism of action and taking into account the pharmacokinetics and pharmacodynamics of the drug?	☒	☐	☐	19, 23
5.5 Does the protocol specify whether a dose-dependent or duration-dependent response is measured?	☒	☐	☐	21-23

Comments:

Section 6: Endpoint definition and measurement	Yes	No	N/A	Page Number(s)
6.1 Does the protocol describe how the endpoints are defined and measured?	☒	☐	☐	22,23-25
6.2 Does the protocol discuss the validity of endpoint measurement? (e.g. precision, accuracy, sensitivity, specificity, positive predictive value, prospective or retrospective ascertainment, use of validation sub-study)	☒	☐	☐	25

Comments:

Section 7: Confounders and effect modifiers	Yes	No	N/A	Page Number(s)
7.1 Does the protocol address known confounders? (e.g. collection of data on known confounders, methods of controlling for known confounders)	☒	☐	☐	25-33,40-42
7.2 Does the protocol address known effect modifiers? (e.g. collection of data on known effect modifiers, anticipated direction of effect)	☒	☐	☐	40

Comments:

Section 8: Data sources	Yes	No	N/A	Page Number(s)
8.1 Does the protocol describe the data source(s) used in the study for the ascertainment of:				
8.1.1 Exposure? (e.g. pharmacy dispensing, general practice prescribing, claims data, self-report, face-to-face interview, etc.)	☒	☐	☐	18,34-36
8.1.2 Endpoints? (e.g. clinical records, laboratory markers or values, claims data, self-report, patient interview including scales and questionnaires, vital statistics, etc.)	☒	☐	☐	18,34-36
8.1.3 Covariates?	☒	☐	☐	18,34-36

Section 8: Data sources	**Yes**	**No**	**N/A**	**Page Number(s)**
8.2 Does the protocol describe the information available from the data source(s) on:				
8.2.1 Exposure? (e.g. date of dispensing, drug quantity, dose, number of days of supply prescription, daily dosage, prescriber)	☒	☐	☐	34-36
8.2.2 Endpoints? (e.g. date of occurrence, multiple event, severity measures related to event)	☒	☐	☐	34-36
8.2.3 Covariates? (e.g. age, sex, clinical and drug use history, co-morbidity, co-medications, life style, etc.)	☒	☐	☐	34-36
8.3 Is a coding system described for:				
8.3.1 Diseases? (e.g. International Classification of Diseases (ICD)-10)	☐	☒	☐	
8.3.2 Endpoints? (e.g. Medical Dictionary for Regulatory Activities (MedDRA) for adverse events)	☐	☒	☐	
8.3.3 Exposure? (e.g. WHO Drug Dictionary, Anatomical Therapeutic Chemical (ATC)Classification System)	☐	☒	☐	
8.4 Is the linkage method between data sources described? (e.g. based on a unique identifier or other)	☒	☐	☐	22,38

Comments:

Section 9: Study size and power	**Yes**	**No**	**N/A**	**Page Number(s)**
9.1 Is sample size and/or statistical power calculated?	☒	☐	☐	37-38

Comments:
Detectable relative risks for main endpoints were estimated

Section 10: Analysis plan	**Yes**	**No**	**N/A**	**Page Number(s)**
10.1 Does the plan include measurement of excess risks?	☐	☒	☐	
10.2 Is the choice of statistical techniques described?	☒	☐	☐	40-41
10.3 Are descriptive analyses included?	☒	☐	☐	38-40
10.4 Are stratified analyses included?	☒	☐	☐	39-40
10.5 Does the plan describe methods for adjusting for confounding?	☒	☐	☐	40
10.6 Does the plan describe methods addressing effect modification?	☒	☐	☐	40

Comments:

Section 11: Data management and quality control	**Yes**	**No**	**N/A**	**Page Number(s)**
11.1 Is information provided on the management of missing data?	☐	☒	☐	
11.2 Does the protocol provide information on data storage? (e.g. software and IT environment, database maintenance and anti-fraud protection, archiving)	☒	☐	☐	38
11.3 Are methods of quality assurance described?	☒	☐	☐	42
11.4 Does the protocol describe possible quality issues related to the data source(s)?	☒	☐	☐	43
11.5 Is there a system in place for independent review of study results?	☒	☐	☐	36,42

Comments:

Section 12: Limitations	**Yes**	**No**	**N/A**	**Page Number(s)**
12.1 Does the protocol discuss:				
12.1.1 Selection biases?	☒	☐	☐	43
12.1.2 Information biases? (e.g. anticipated direction and magnitude of such biases, validation sub-study, use of validation and external data, analytical methods)	☒	☐	☐	43
12.2 Does the protocol discuss study feasibility? (e.g. sample size, anticipated exposure, duration of follow-up in a cohort study, patient recruitment)	☒	☐	☐	37-38
12.3 Does the protocol address other limitations?	☒	☐	☐	43

Comments:

Section 13: Ethical issues	**Yes**	**No**	**N/A**	**Page Number(s)**
13.1 Have requirements of Ethics Committee/Institutional Review Board approval been described?	☒	☐	☐	36
13.2 Has any outcome of an ethical review procedure been addressed?	☒	☐	☐	36
13.3 Have data protection requirements been described?	☒	☐	☐	36

Comments:

Section 14: Amendments and deviations	Yes	No	N/A	Page Number(s)
14.1 Does the protocol include a section to document future amendments and deviations?	☒	☐	☐	13

Comments:

Section 15: Plans for communication of study results	Yes	No	N/A	Page Number(s)
15.1 Are plans described for communicating study results (e.g. to regulatory authorities)?	☒	☐	☐	46
15.2 Are plans described for disseminating study results externally, including publication?	☒	☐	☐	46

Comments:

Name of the main author(s) of the protocol: [redacted]

Date: 23 November 2022

Signature: [redacted]

Name of the main author(s) of the protocol: [redacted]

Date: 29 November 2022

Signature: [redacted]

Name of the main author(s) of the protocol: [redacted]

Date: 29 November 2022

Signature: [redacted]