TITLE

A Non-interventional Observational Longitudinal Post-Authorization Safety Study (PASS) of SIMPONI® in Treatment of Ulcerative Colitis using Nordic National Health Registries

PASS information

Title	**A Non-interventional Observational Longitudinal Post Authorization Safety Study (PASS) of SIMPONI® in Treatment of Ulcerative Colitis using Nordic National Health Registries**
Protocol Version identifier	Version 2.0
Date of last version of protocol	05-Sep-2018
EU PAS Register No:	ENCEPP/SDPP/11484
Active substance	ATC code L04AB06 – golimumab
Medicinal product:	SIMPONI® (golimumab) solution for injection
Product reference:	EU/1/09/546
Procedure number:	EMEA/H/C/992 MEA026
Marketing authorization holder(s) (MAH)	Janssen Biologics B.V.
Joint PASS	No
Research question and objectives	This study will address whether, in ulcerative colitis (UC) patients with similar baseline disease characteristics, the use of golimumab (GLM) is associated with the risk of colorectal cancer (CRC), all-cause total colectomy, and hepatosplenic T-cell lymphoma (HSTCL) as compared with alternative therapies. In this study, no prior research hypotheses have been formulated. **Primary objectives** 1. To characterize 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 alternative therapies 2. To describe the risk of the following endpoints in patients exposed to GLM and exposed to alternative therapies

	2a. Incident CRC 2b. All-cause total colectomy 3. If baseline characteristics suggest comparability, to estimate 3a. The risk of CRC associated with GLM use relative to that associated with alternative therapies 3b. The risk of all-cause total colectomy associated with GLM use relative to that associated with alternative therapies **Exploratory objective** To describe the risk of incident HSTCL associated with exposures of interest
Countries of study	Denmark; Sweden
Author	PPD Institute of Applied Economics and Health Research Ewaldsgade 3 DK-2200 Copenhagen N, Denmark Telephone number: PPD E-mail: PPD
MAH	Janssen Biologics B.V. Einsteinweg 101 2333 CB Leiden, The Netherlands
MAH Contact Person	PPD Janssen Biologics B.V. Einsteinweg 101, P.O. Box 251 2300 AG Leiden The Netherlands Telephone number: PPD Fax number: PPD E-mail: PPD

1. Table of contents

2. List of abbreviations

5-ASA	5-aminosalicylic acid
6-MP	6-mercaptopurine
ADA	Adalimumab
AER	Adverse Event Reporting
ApEHR	Institute of Applied Economics and Health Research
ATC	anatomical therapeutic chemical
AZA	azathioprine
CD	Crohn's disease
CRC	colorectal cancer
EMR	electronic medical record
GLM	golimumab
HSTCL	hepatosplenic T-cell lymphoma
IBD	inflammatory bowel disease
ICD	International Classification of Diseases
IFX	infliximab
MAH	marketing authorization holder
PASS	post-authorization safety study
PIN	personal identification number
PY	person-years
QBA	quantitative bias analysis
SAP	statistical analysis plan
TNF	tumor necrosis factor
UC	ulcerative colitis

Note: "register" and "registry" are synonymous terms that are used to refer to an official written record of names or events or transactions. In this protocol, "registry" will be used as a general term for such an entity, and "register" will be used when referring to a specific entity in which "register" is part of the official English translation of the name of that entity.

3. Responsible parties

Principal investigator	PPD Institute of Applied Economics and Health Research Ewaldsgade 3 DK-2200 Copenhagen N, Denmark Phone: PPD E-mail: PPD
Sponsor contacts	Merck Sharp & Dohme, Corp., a subsidiary of Merck & Co., Inc. One Merck Drive P.O. Box 100 Whitehouse Station, NJ, 08889-0100, U.S.A. Contacts: PPD Principal Scientist, Epidemiology Merck Sharp & Dohme Corp. PPD Executive Director, Epidemiology Merck Sharp & Dohme Corp.
Marketing authorization holder (MAH)	Janssen Biologics B.V. Einsteinweg 101 2333 CB Leiden, The Netherlands
Investigators	See stand-alone document Specification of Investigators

4. Abstract

Title	A Non-interventional Observational Longitudinal Post Authorization Safety Study (PASS) of SIMPONI® in Treatment of Ulcerative Colitis (UC) using Nordic National Health Registries
Protocol Number / Version	MK8259-013-01
Date	05-Sep-2018
Author	PPD
Rationale & Background	Simponi received European marketing authorization for treatment of moderately to severely active UC on 19-Sep-2013. This registry-based study is being established to provide additional information on colorectal cancer (CRC), colectomy, 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 nationwide health registries in Denmark and Sweden. To provide a context for interpreting the long-term safety data on UC patients treated with golimumab (GLM), this study will also follow patients with UC treated with alternative therapies for UC, including non-biologic and biologic therapies. UC is a severe, relatively frequent and life-long disease that may be associated with severe symptoms and impaired quality of life. In the case of intractable disease, colectomy may be indicated. Patients with UC are at increased risk of developing cancer of the colon and rectum compared to the general population.
Research Question and Objectives	This study will address whether, in UC patients with similar baseline characteristics, the use of GLM is associated with the risk of CRC, all-cause total colectomy, and HSTCL as compared with alternative therapies for moderate-to-severe UC. In this study, no prior research hypotheses have been formulated. **Primary objectives** 1. To characterize 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 alternative therapies 2. To describe the risk of the following endpoints in patients exposed to GLM and exposed to alternative therapies

	2a. Incident CRC 2b. All-cause total colectomy 3. If baseline characteristics suggest comparability, to estimate 3a. The risk of CRC associated with GLM use relative to that associated with alternative therapies 3b. The risk of all-cause total colectomy associated with GLM use relative to that associated with alternative therapies **Exploratory objective** To describe the risk of incident HSTCL associated with exposures of interest.
Study Design	This study is a long-term observational (non-interventional), prospective PASS. It will use a new user cohort design and will be based on data that were primarily collected for administrative reasons. One exposed and two comparator cohorts will be assembled: • An inception cohort of UC patients treated with GLM (GLM Cohort): this cohort will comprise all patients registered as new users of GLM. • An inception comparator cohort of UC patients treated with anti-tumor necrosis factor (TNF) agents other than GLM (Other anti-TNF cohort): this cohort will comprise patients initiating one of the anti-TNF agents approved for the treatment of UC (currently infliximab (IFX) and adalimumab (ADA)). • An inception cohort of UC patients treated with a thiopurine (azathioprine (AZA) or 6-mercaptopurine (6-MP)) (Thiopurine Cohort): this cohort will comprise patients initiating first time treatment with thiopurines (AZA or 6-MP), which are non-biologic immunomodulating drugs. Patients in each cohort will be followed for the outcomes of CRC, colectomy, and HSTCL [redacted]
Population	The study will include all patients with UC starting a cohort-defining treatment in Denmark and Sweden during the period [redacted] (when GLM was approved for UC) through [redacted].
Variables	***Exposures***

	The main exposure of interest is treatment with GLM in patients with moderate-to-severe UC, compared with exposure to alternative therapies for moderate-to-severe UC: (1) other anti-TNF agents and (2) thiopurines (AZA or 6-MP). Over the course of time, UC patients who initiate GLM may have received or may later receive comparator treatments in various sequences. ***Outcomes*** *The primary outcomes* comprise: • Incident CRC • All-cause total colectomy The *exploratory* outcome is: • HSTCL The list of *covariates* includes: • Calendar year of any outcome • Age • Gender • Country (Denmark or Sweden) • Hospital department • Time (calendar year) of enrolment • Disease duration • Disease activity (measured through proxies) • Extent of disease • Switching from another biologic agent • Co-morbidities • Colonoscopy and sigmoidoscopy
Data Sources	• Nationwide central hospital registration systems in Denmark and Sweden will be used to identify the study base population of patients with UC. • Nationwide central hospital registration systems and prescription registries will be used to identify the GLM inception cohort and the inception comparator cohorts. • Nationwide health registries (hospital activity registries, prescription registry, cancer registries) will be used to ascertain outcomes and changes in treatment. Cancer registries will be used to validate ascertainment of malignancies. • Deaths will be captured from the Civil Registration Services in Denmark and Sweden. Death will be treated as a censoring event. • Data from the various sources will be linked at individual patient level using the unique personal

	identification number (PIN) assigned to all citizens in Denmark and Sweden.
Study Size	The size of the study base (i.e., the current prevalent population of patients with UC) is estimated to be approximately 55,000 (20,000 patients in Denmark and 35,000 patients in Sweden). The sizes of the inception cohorts to be established and followed-up during the study are unknown. Based on the most recent sales projections, about 100 patients across both countries will initiate treatment with GLM each year leading to 2450 person-years (PY) of follow-up over the course of the study. Assuming a background incidence rate of 1/1000 PY for CRC in the unexposed study population, 2450 PY of GLM exposed person-time, α=0.05 and a 1:5 ratio of GLM to comparator group in follow-up time, this study is estimated have 80% power to exclude a relative risk of 4.0 for the CRC endpoint between GLM- and thiopurine- exposed UC patients. Under the same conditions but assuming an incidence rate of 2/1000 PY for the CRC endpoint in the unexposed study population, this study will be able to exclude a relative risk of 2.9.
Data Analysis	• Baseline analyses will describe each cohort in terms of patient characteristics that can be ascertained from automated registry data. For each of the three inception cohorts, annual enrolment will be described, along with cumulative PY accrued. • The risk of primary outcomes (incident CRC, all-cause total colectomy) will be estimated as cumulative incidence within each inception cohort in unadjusted and adjusted analyses. Censoring events will include death, emigration, and first occurrence of study outcomes. The cohort analyses will be conducted based on automated data only, using survival analysis techniques to estimate cumulative risks. • The incidence rate of HSTCL will be described for each inception cohort. • A range of sensitivity analyses has been specified for the final analyses that examine the robustness of results to alternate specifications of the study population, risk window, outcome definitions, and additional adjustment using quantitative bias analysis (QBA) to

	account for missing information on a confounder.
Milestones Start of data collection: End of data collection: Study progress report(s): Final report of study results:	[redacted]

5. Amendments and updates

This Section summarizes all key changes made. Amendments made as consequence of such key changes are not listed. All changes made are first amendments with effective date [redacted]

Section of Study Protocol	Amendment	Reason
7.2. Background	Addition of new references and adjustment of text	To update and align background with revised objectives
8. Research question and objectives *and* 9.3.3. Outcomes	Primary objectives: (1) Endpoint changed from "colectomy for intractable disease" to "all-cause total colectomy. (2) Endpoint changed from composite of incident colorectal cancer or high-grade colorectal dysplasia to incident colorectal cancer only (promoting the secondary objective to primary) Exploratory objective: (3) Removed evaluation of duration of remissions after golimumab discontinuation	(1) Automated data do not provide reason for colectomy (2) High-grade colorectal dysplasia could not be ascertained in automated data (3) Infeasible due to insufficient data in registries
9. Research methods 9.2.2. Inception cohorts: Ulcerative colitis patients treated with golimumab and alternative therapies *and* 9.3.2. Establishment of inception cohorts by treatment	Entry criteria revised: no longer require that study drugs were prescribed to treat ulcerative colitis	Cannot ascertain reason for treatment from available data
Original 9.2.3. Identification of cases and controls for nested case-control analyses *and* 9.4.1. Study Procedures	Removed plans for nested case-control analyses using chart review	Judged infeasible. Added instead plans for supportive study in Sweden comparing national registry with electronic medical record data to validate colectomy endpoint and ascertain proportion of unmeasured infliximab exposure, which will be used in a quantitative bias analysis
9.2.3. Registry sources (original 9.2.4)	Exclusion of mentioning of Pathology Registers	Pathology registers will not be used as data source
Original 9.2.5. Pilot	Deleted	Work was completed and used to

Study		inform amended protocol
9.3.4. Covariates	Update in list and description of covariates	Changes based on knowledge gathered from Pilot Study
9.3.5. Follow-up time, switches and definitions of risk windows	Text is updated. Original Table 3 moved from Section 9.3.2 and simplified and re-numbered to Table 2	Consideration on exposure concepts put in more appropriate context
9.7. Data analysis	Wording changed to align with updated outcomes, dropping the nested case control design, and revision of sensitivity analyses, including quantitative bias analysis	Changes based on knowledge gathered from Pilot Study
9.7.3. Sensitivity analyses	New section of Sensitivity analyses was added. (1) Evaluates alternative specification of colectomy outcome to approximate colectomy for intractable disease (2) Evaluates sensitivity of missing information on maximal disease extent as inferred from ICD10-codes (3) Repeats main analyses limited to patients with at least 2 ulcerative colitis diagnoses by time of cohort entry (4) Performs quantitative bias analysis (QBA) to adjust for unmeasured IFX exposure in Sweden	Provides clearer overview of the planned sensitivity analyses. (1) Uses clinical logic to approximate colectomies done for intractable disease (2) Optimizes assessment of validity and interpretation of results (3) Goal is to exclude from analysis patients who may not truly have ulcerative colitis (4) To adjust for residual confounding due to unmeasured prior IFX exposure
11. Management and reporting of adverse events	Updated text after dropping nested case control study design	Aligns with dropping nested case control design and no medical chart review in main study

6. Milestones

Milestone	Planned date
Start of data collection	[redacted]
End of data collection	[redacted]
Study progress report 1	[redacted]
Subsequent study progress reports	[redacted]
Registration in the EU PAS register	ENCEPP/SDPP/11484
Final report of study result	[redacted]

7. Rationale and background

7.1. Rationale

SIMPONI® (golimumab [GLM]) was approved on October 1, 2009 in the European Union for the indications rheumatoid arthritis, ankylosing spondylarthritis and psoriatic arthritis. GLM is a tumor necrosis factor (TNF) antagonist that received European marketing authorization for treatment of moderately-to-severely active ulcerative colitis (UC) on 19-Sep-2013. This registry-based study is being established to provide additional information on colorectal cancer (CRC), colectomy, and hepatosplenic T-cell lymphoma (HSTCL), as outlined in the Risk Management Plan for SIMPONI® that was approved with authorization of the UC indication.

Since 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, nor any of the other clinical trials in the development program for GLM, have demonstrated an association between GLM and CRC, HSTCL, or colectomy.[1,2] Use of data collected in population-based registries provides an additional source of safety data.

The study will use data from nationwide health registries in Denmark and Sweden. Because they are population-based, these registries are expected to capture all GLM treatment used in UC in these countries since the date of its market authorization. The national health registries in these countries also provide comprehensive, validated data on all dispensings of prescribed medicines, interventions performed in hospital for diagnostic assessment and treatment, which will be used to ascertain the study endpoints of relevance. Information from the national cancer registries will be used to validate cases of CRC captured from hospital activity registrations. 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 thiopurines and other TNF inhibitors.

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 (IFX), adalimumab (ADA) 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 (systemic steroids are not recommended for use as long-term therapy in UC). Anti-TNF therapies, then, are preferentially prescribed to patients with more severe, treatment-resistant disease. In addition to the preceding general treatment scheme, newer therapies have shown activity against UC, including tacrolimus (a calcineurin inhibitor that is primarily used as an anti-rejection drug in organ transplantation), tofacitinib (an oral Janus kinase inhibitor), and biologic agents natalizumab and vedolizumab (both anti-integrins). These drugs are not commonly used in current clinical practice either because they are not approved for UC indications or because their role in the clinical cascade is not well established.[3,4] In mid-2014, vedolizumab was approved for treatment of UC.

Colectomy. When medical therapy fails, total colectomy is the surgical therapy of choice. 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 population-based UC cohorts.[5,6] Risk for colectomy appears greatest in the first two years following diagnosis.[7] Since the introduction of anti-TNF agents, there appears to be temporally decreasing rates of colectomy correlating with anti-TNF use in North America and Europe.[8] Anti-TNF therapies are used for moderate-to-severe UC and often as second or higher line therapies when other therapies fail. Accordingly, the risk for total colectomy may depend on treatment history including short-term response to the TNF biologics, concomitant use of immune-suppressants, as well as treatment sequence. A recent retrospective cohort study that followed patients with UC reported that colectomy occurred in about 1.7%, 12%, and 19% of the patients treated with GLM as the first, second, and third anti-TNF therapy, respectively.[9] A similar study of UC patients commencing ADA also found a higher proportion (17.2%) of TNF-experienced patients underwent colectomy than TNF-naïve patients (2.9%).[10] The higher risk of colectomy observed in TNF-experienced patients compared to TNF-naïve patients may be associated with lower treatment response among patients who are more difficult to treat.[9]

The majority of total colectomies in UC are performed for intractable disease. Colectomy is also performed to treat CRC and depending on the clinical context, colorectal dysplasia. In a 2009 Norwegian population-based cohort study, CRC and dysplasia accounted for only about 10% of colectomies, while medically refractory UC accounted for 90% of the 49 colectomies.[7] Less common reasons for colectomy in UC include emergent complications such as toxic megacolon, colonic perforation, massive hemorrhage and colonic obstruction; all of these complications of UC can be considered as manifestations of intractable disease.

Neoplasia. Compared to the general population, patients with UC are at an increased risk of CRC and colorectal dysplasia.[11] Colorectal dysplasia carries an increased probability of progression to CRC and may be treated with colectomy or managed through surveillance

colonoscopy.[12] Professional society guidelines suggest enhanced surveillance with colonoscopy to monitor for CRC and dysplasia in patients with longstanding UC.[3,4]

The 10-year cumulative risk of CRC in UC is approximately 2%.[13] The cumulative probability of developing dysplasia or CRC was reported as 7.7% at 20 years and 15.8% at 30 years.[14] The risk 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; and a family history of CRC. More recent studies suggest that the risk of colonic neoplasia in UC has diminished over time.[15,16] A population-based study in the Danish UC population (1978-2008) found that the incidence rate of CRC was 1.05/1000 PY.[16]

A separate 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 CRC risk (adjusted relative risk 1.0 (0.48 – 2.08).[17]

In addition to CRC, HSTCL (a rare lymphoma that is often fatal) has been identified as a possible risk in patients with IBD treated with anti-TNF therapies. Almost all patients who develop HSTCL have also been previously exposed to thiopurines, which may also contribute to the risk of HSTCL.[18]

8. Research question and objectives

This study will address whether, in UC patients with similar baseline characteristics, the use of GLM is associated with the risk of CRC, all-cause total colectomy, and HSTCL as compared with alternative therapies for moderate-to-severe UC. In this study, no prior research hypotheses have been formulated.

8.1. Primary objectives

8.1.1. To characterize 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 alternative therapies

8.1.2. To describe the risk of the following endpoints in patients exposed to GLM and exposed to alternative therapies

8.1.2a. Incident CRC

8.1.2b. All-cause total colectomy

8.1.3. If baseline characteristics suggest comparability, to estimate

8.1.3a. The risk of CRC associated with GLM use relative to that associated with alternative therapies

8.1.3b. The risk of all-cause total colectomy associated with GLM use relative to that associated with alternative therapies

8.2. Exploratory objective

To describe the risk of incident HSTCL associated with exposures of interest.

9. Research methods

9.1. Study design

This is a long-term observational (non-interventional), prospective post-authorization safety study (PASS). It will be based on a new user cohort design and will use national health registry data that were primarily collected for administrative reasons.

In both Denmark and Sweden, the healthcare system is public and hospital services are free of charge for all citizens. In both countries it is possible to monitor virtually the entire population of patients with chronic diseases like UC by means of information that is continually reported to the national hospital registration systems in each country. These systems capture information about inpatient care, as well as care in emergency departments and hospital-based ambulatory specialty clinics. Each patient who receives hospital-based care will be registered and made identifiable by his/her unique personal identification number (PIN), and for each hospital activity all relevant diagnoses, surgical interventions and other treatments as well as procedures performed as part of diagnostic evaluation will be registered. The data in the hospital activity registries may be used for research purpose(s) if approval has been obtained from the authorities.

Using the PINs, it is possible to perform linkage at the individual patient level to other nationwide health registries including the cancer registries, as well as the registries of prescribed medicines in Denmark and Sweden. The possibilities of using data and record linkage in the nationwide health registries in Denmark and Sweden offer unique opportunities to address the research questions and corresponding objectives of the present study.

Data from these registries will be used to identify cohorts of patients with UC with first-time use of GLM. Registry data will also be used to identify the comparator cohorts, which will comprise patients with UC who are first-time users of comparator therapies (see `Exposures` below).

Patients who are in the GLM and comparator cohorts will be characterized at baseline and followed for up to 7 years from cohort entry to determine the incidence of CRC, all-cause total colectomy, 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. In view of this, analyses of risk of CRC and colectomy will use different definitions of exposure-risk windows.

9.2. Setting

9.2.1. Study base: Patients with UC and relevant treatment ascertained by the health registers

The identification of patients with UC in Denmark and Sweden will be based on the centralized hospital registration system by all inpatient and outpatient encounters registered

with UC (ICD10-code K51.0-K51.9 (see Annex 3: Classification Codes – part 1)) as the primary activity diagnosis. In operational terms, the date of diagnosis of UC will be taken as the earliest date of encounter where UC has been recorded as the primary activity diagnosis.

In usual clinical practice, all patients with UC in Denmark and Sweden will be registered as inpatients or outpatients for the diagnosis and management of UC. For each hospital contact, discharge diagnoses including the primary activity diagnosis as well as other relevant diagnoses (secondary diagnoses) will be recorded.

9.2.2. Inception cohorts: Ulcerative colitis patients treated with golimumab and alternative therapies

In Denmark and Sweden, patients with moderate to severe UC are managed exclusively by specialists with competence in gastroenterology in hospitals and hospital-based ambulatory clinics. Only in these settings are immunomodulatory therapies and anti-TNF agents prescribed. Registrations indicating new treatments with anti-TNF agents or thiopurines (AZA or 6-MP) will form the basis for membership in the inception cohorts of interest. Ascertainment of exposures is discussed in section 9.3. Thus, the currently defined inception cohorts will be defined as follows:

- The inception cohort of UC patients treated with GLM (the GLM Cohort): will be defined as all patients who have at least one UC diagnosis and who for the first time after 19-Sep-2013 are registered as treated with GLM. The enrolment date will be defined operationally as the date of initiation of this particular treatment. New members will be added to the cohort with last possible date of enrolment at18-Sep-2020. Patients with antecedent diagnoses of CRC, colectomy, or HSTCL will be excluded from the cohort.
- The inception cohort of UC patients treated with anti-TNF agents other than GLM (Other Anti-TNF Cohort): will be defined as all patients who have at least one UC diagnosis and who for the first time after 19-Sep-2013 are registered as treated with IFX or ADA. The enrolment date will be defined operationally as the date of initiation for this particular treatment. New members will be added to the cohort with last possible date of enrolment at 18-Sep-2020. Patients with antecedent diagnoses of CRC, colectomy, or HSTCL will be excluded from the cohort.
- The inception cohort of UC patients treated with a thiopurine (AZA or 6-MP) (Thiopurine Cohort): will be defined as all patients who have at least one UC diagnosis and who for the first time after 19-Sep-2013 are registered as treated with AZA or 6-MP. The enrolment date will be defined operationally as the date of initiation for this particular treatment. New members will be added to the cohort until last possible date of enrolment at 18-Sep-2020. Patients with antecedent diagnoses of CRC, colectomy, or HSTCL will be excluded from the cohort, as will patients with prior treatment with biologic agents.

9.2.3. Registry sources

All health registries in the Nordic countries can be linked to each other by means of the unique PIN, assigned to each inhabitant within the respective countries, and used throughout

the public sector. This unique setting permits the establishment of the GLM inception cohort and the comparator cohorts, as well as the long-term follow-up of all the study patients.

A comprehensive description of the Danish registries concerning their structure, organizational aspects and use for research purposes may be found in Thygesen et al 2011[19] and at the homepage of the Danish National Board of Health via the State Serum Institute http://www.ssi.dk/Sundhedsdataogit/Registre%20og%20kliniske%20databaser/De%20nationale%20sundhedsregistre.aspx (in Danish). An overview of the Swedish national health registries is available at the homepage of the Swedish National Board of Health and Welfare, http://www.socialstyrelsen.se/english.

The health registries in Denmark and Sweden of relevance for the present study are listed in Table 1.

Table 1. The central health registries of relevance for the present study

Health registry	Year starting automated operation		Role in the study
	Denmark	Sweden	
Hospital activity registration systems (Patient Registries)	1977 (1995)[1]	1987 (2005)[1]	Identification of - Patients with UC - Clinical units managing UC patients - Therapies administered by hospital departments[2] - Procedures (colonoscopies, sigmoidoscopies and colectomies) - Occurrences of CRC and HSTCL
Civil registration systems	1968	1965	Deaths and other major vital demographic events, incl. emigrations
Prescription registries	1995	1999/ mid-2005[3]	Information on therapies by drugs purchased via prescriptions to patients with UC
Cancer registry	1943	1958	Validation of cases of CRC and HSTCL

[1] Years in parentheses indicate when registration of outpatient activities commenced
[2] In Sweden, the coding of administering therapies at hospital departments and clinics is incomplete
[3]The Swedish National Prescription Registry commenced data collection in 1999 but contains patient identifiers only from mid-2005 and onwards

The sections below describe each of the relevant registries in further detail.

Hospital activity registration systems (patient registries)

Denmark: The Danish National Patient Register contains data on all hospital activities in Denmark since 1977 (since 1995 for encounters on an outpatient basis) with coded diagnostic procedures and treatments together with administrative data.[20] Link: http://www.ssi.dk/Sundhedsdataogit/Registre/Landspatientregisteret.aspx. Particularly concerning IBD and its components Crohn's disease (CD) and UC, a study has verified that the validity of the UC diagnosis is more than 90% in the registry, increasing to more than 94% for the patients registered with specialized departments.[21]

The register is dynamic and updated, in practical terms, on a monthly basis, including amendments to previous registrations. The delay from an event to public release of the

corresponding data is in general about 2-3 months for encounters on an inpatient basis and in general longer for encounters on an outpatient basis.

Sweden: From 1987 there is information on all completed in-patient admissions at publicly operated hospitals. Registration of encounters on an outpatient basis commenced in 2005. The key variables are diagnosis, surgery, external causes of injury (E-codes), gender, age, residence, hospital, specialty, and hospital admissions and discharges. The coding of therapies given to patients by the hospitals or hospital- based outpatient clinics is incomplete, but to an unknown extent. For the present study this will affect the identification of treatments with IFX since drugs administered by intravenous infusion will as a principle not be identified from the prescription registration system.

Currently, the registry is updated annually with the release of data taking place in August/September, corresponding with a lag-time at about 9-21 months. Link: http://www.socialstyrelsen.se/register/halsodataregister/patientregistret/inenglish.

Completeness of registrations of inpatient encounters is believed to be close to 100%. Reporting from private outpatient clinics is incomplete, which causes an estimated overall deficit at 20% for outpatient encounters (Jacobsson A, personal information). However, according to judgment from a clinical expert, activities for a disease like UC will, with very few exceptions, be managed in public clinics thus reducing reporting deficits to a minimum (Hertervig, personal communication).

Civil registration systems

Denmark: The Danish Civil Registration System was established on a digitalized platform in 1968 and operates the allocation of PINs (and, from the PIN, information on age and gender) to all Danish citizens and stores information on births, moves (including migrations) and deaths in the Danish population.[22] Link: http://www.ssi.dk/Sundhedsdataogit/Registre/CPR-registeret.aspx.

Sweden: The digitalized civil registration system in Sweden is similar to the Danish system. Information on deaths in the study population can be obtained from the Swedish Cause of Death Registry, link: http://www.socialstyrelsen.se/register/dodsorsaksregistret.

Prescription registries

Denmark: The Danish National Prescription Registry contains data on all redemptions of prescriptions with information including person, date, location, prescribing doctor, package and strength and, thereby, numbers of defined daily dosages (DDD) purchased.[23] The registry is accessible for research purposes via Statistics Denmark and is updated continually, i.e. on a monthly basis. Link:

http://www.ssi.dk/Sundhedsdataogit/Registre/Laegemiddelstatistikregisteret.aspx.

Sweden: The National Prescription Registry contains data on prescriptions redeemed in Sweden since 1999. Identity of the drug, prescribing doctor, amount, price and the date of prescription are registered in the database. From July 2005 the registry also contains data on the patient's PIN, gender, age and place of residence. The registry is updated monthly. Link:

http://www.socialstyrelsen.se/register/halsodataregister/lakemedelsregistret

Cancer registries

There is a longstanding tradition for operating cancer registries of high quality in both Denmark and Sweden, and the content is accessible for research subject to relevant permissions. However, because of the time-consuming process of case validation there is a substantial delay in the release of information from these registries. In this protocol cases of CRC and HSTCL will be ascertained primarily from diagnoses in the hospital registration described above. For the final study report, cases of CRC and HSTCL in the inception cohorts will be validated from the cancer registries. The registries are described in detail here:

Denmark: Since 1943 all cases of cancer have been reported to a nationwide registry, initially operated by the Danish Cancer Society. Since 2003, the Danish Cancer Registry has been operated by the Danish National Board of Health closely linked with the Danish National Patient Register and supplemented with data from the Danish Pathology Register. The registry contains, in addition to the PIN, cancer diagnosis information (ICD10-code system), date of diagnosis, information on location and morphology as well as data on cancer stage.

The cancer registration system has a very high validity concerning completeness of case registration as well as validity of the data reported for the cases. It is however a disadvantage that there is a delay of up to two years before new information in the registry is published. See Engholm et al for a description of the Nordic cancer registries and their potential for epidemiological research.[24] Link:

http://www.ssi.dk/Sundhedsdataogit/Registre/Cancerregisteret.aspx.

Sweden: Since 1958 all cases of cancer have been reported to a nationwide registry. Link:

http://www.socialstyrelsen.se/register/halsodataregister/cancerregistret/inenglish.

The registry contains, in addition to the PIN, cancer diagnosis information (ICD10-code system), date of diagnosis, information on pathology and morphology as well as data on cancer stage and other clinical data. The cancer registration system has a very high validity concerning completeness of case registration as well as validity of the data reported for the cases. See Engholm et al for a description of the Nordic cancer registries and their potential for epidemiological research.[24]

9.3. Variables

Conceptual definitions of exposure, outcomes, and covariates are presented here. All study variable definitions will be specified in detail in the statistical analysis plan (SAP) before any comparative analyses are conducted.

9.3.1. Basic demographic data

The basic demographic variables include age at the operationally established date of diagnosis of UC, gender, and country (Denmark or Sweden), as well as date of death and date of emigration.

9.3.2. Exposures to study drugs

The main exposure of interest is treatment with GLM for UC. The GLM cohort will be compared with cohorts of patients exposed to two alternative therapies for moderate-to-severe UC: (1) other anti-TNF agents and (2) thiopurines (AZA or 6-MP). Over the course of time, UC patients who initiate GLM may have received or may later receive comparator treatments in various sequences.

GLM cohort: patients with UC starting on GLM. They can receive thiopurines before, during, or after GLM treatment; can also have other anti-TNF before or after GLM treatment

Anti-TNF comparator cohort: patients with UC starting on anti-TNF drugs other than GLM. They can receive thiopurines before, during, or after GLM treatment; the anti-TNF agent that qualifies the patient for cohort entry may be the patient's first or subsequent anti-TNF agent. Besides GLM, the only anti-TNF agents currently indicated to treat UC are IFX and ADA.

Thiopurine comparator cohort: patients with UC initiating thiopurines (AZA, 6-MP), a non-biologic immunomodulator, and who are naïve to biologic agents.

Only new use of a study medication during the calendar time of the study will qualify for cohort entry. Prevalent use of a study medication will not qualify. A patient switching from one anti-TNF agent to another will, however, be considered a new user of that specific agent.

Information about the use of anti-TNF therapies differs between Sweden and Denmark. In Denmark, these therapies are dispensed at the hospital departments managing patients with UC, but will not be ascertainable from the prescription registry. Since the hospitals must register these therapies in order to obtain state refunds, therapies will be captured by means of the procedure codes applicable to this class of drugs. In Sweden, the anti-TNF therapies GLM and ADA will be ascertainable from the prescription registry by means of the anatomical therapeutic chemical (ATC) code. The ascertainment of IFX will occur both from the hospital registry and prescription registry. Information about hospital-based IFX use will be extracted from the hospital registration system, but capture of drugs administered in the hospital setting is expected to be incomplete (see section 9.2.3). Ordinarily one does not expect to see pharmacy dispensings of IFX because it is administered by IV infusion. However, some dispensing of IFX from pharmacies does occur, and it can be seen in the prescription register (described below). Clinical consultants have indicated that this may occur when patients receive IFX infusions through private clinics. In summary, in Sweden IFX use will be ascertained through both the hospital and pharmacy registries, and it is expected to be incomplete (addressed by sensitivity analysis in Section 9.7.3).

Information on non-biologic therapy in UC will be available from prescriptions registries in both countries. Concerning prescription data, information will be available on date and amount of the purchase of a given prescribed drug. Neither the indication for prescribing the drug nor the daily dosage prescribed will be available. Annex 3 includes classification codes for study exposures.

Conventions for ascertaining duration of exposure are described in section 9.3.5 (*Ending therapies*).

Treatment of UC often involves changes in therapy over time. Considerations about drug switching, as well as a conceptual framework for evaluating exposure categories in the context of prior and concomitant therapies, appears in Table 2 in section 9.3.5.

9.3.3. Outcomes

The *primary* outcomes are specified as follows:

Incident colorectal cancer: CRC is considered a relevant outcome if first diagnosed after cohort entry. (A history of CRC disqualifies for cohort entry). The date of diagnosis will be defined as the date of referral for the first encounter where CRC is registered as a primary activity diagnosis in the hospital registry. The outcome of CRC will be validated against registrations in the cancer registries at the end of the study, due to the delay in the release of updated information in these cancer registries. For CRCs diagnosed toward the end of the study period, when it is anticipated cancer registry data will not be available, diagnoses from the hospital register will be accepted as true.

All cause total colectomy: Total colectomy for ICD-10 codes (see Annex 3: Classification of codes) will be considered a relevant outcome if performed after cohort entry. A study reported that among patients with IBD in an acute hospital setting, codes in the Danish NPR for total colectomy had a positive predictive value (PPV) of 97% (156/161 confirmed; 95%, CI 93-99%), indicating they have high validity for purposes of this PASS.[25] No study has been identified describing the validity of total colectomy ascertainment in the Swedish hospital register. To address this question, the Sponsor is performing a separate methods study that evaluates the validity of total colectomy codes in Sweden, using a sample of electronic medical records (EMR) as the reference standard.

The *exploratory* outcome is:

Hepatosplenic T-cell lymphoma: HSTCL will be considered a relevant outcome if it is first diagnosed after cohort entry. The date of diagnosis will be defined as the date of referral for the first admission where HSTCL is registered as a primary activity diagnosis. The outcome of HSTCL will be validated retrospectively against registrations in the cancer registries. Because the background incidence of HSTCL is so low (in the order of 1 case per million person-years (PY)[26]), it is not likely that many outcomes will be encountered during this study, and as a consequence, confidence limits around this estimate will be wide.[26]

9.3.4. Covariates

Baseline patient characteristics and covariates will be ascertained to describe each cohort and evaluated as potential confounders in this study. Factors will include age, gender, disease extent, disease duration, history of UC treatments, hospitalization for UC, and selected comorbidities. Covariates will be ascertained from automated registry data, Please refer to Background (section 7.2) for a description of risk factors for CRC and colectomy

The concept of "disease activity" factors into the risk of each outcome, but it is important to recognize that the term refers to two related but distinct concepts. First, as is commonly used in clinical practice, current disease activity refers to current biological colonic inflammatory activity, and it can be measured in several ways.[3,27] 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

(fever and pulse), laboratory findings (haemoglobin 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 more closely associated with risk of 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 more closely associated with risk for advanced neoplasia.

In prospective clinical studies, acute disease activity is often measured using the Mayo Score, which reflects several clinical inputs that do not appear in automated data. However, several clinically reasonable proxies can be found in automated data, including disease extent and evidence of a recent increase in UC-related healthcare utilization (e.g., hospital encounters for UC and modifications in treatments 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.[12,16,17,28]

In the analysis, time-dependent covariates, such as disease extent and drug treatment history (a proxy for disease activity), will be updated just before any switch into a new exposure category. First-time treatment with other immunomodulatory therapies including ciclosporin, tacrolimus and vedolizumab will be registered and managed analytically as covariates.

The full list of *covariates* includes

- Calendar year of any outcome
- Age: Age will be defined as (a) age at enrolment in an inception cohort or at time of switch into a new exposure category, and (b) age at the time of an outcome
- Gender: Gender will be classified as males or females
- Country (applicable to comparative analysis of aggregated data, see 9.7): Country of residence will be classified as Danish or Swedish
- Time of enrolment: Time of enrolment will be classified by the calendar year of enrolment in one of the inception cohorts
- Disease duration: Disease duration will be measured as (a) the date at enrolment in an inception cohort minus the date of earliest diagnosis of UC; and (b) the date of registered outcome minus the date of earliest diagnosis of UC.
- Disease activity: Disease activity cannot be measured directly in automated data and will be described through proxies from the automated data to provide perspective on the comparability of cohorts at baseline and evaluated as potential confounders in the analysis. Potential proxies include
    - hospitalization for UC in the year prior to cohort entry
    - number of other anti-TNF-agents previously used
    - history of treatment with immunomodulatory or biologic treatments for UC not included as exposures, such as cyclosporine and vedolizumab

- Extent of disease: Based on the automated data, the extent of disease will be categorized according to the Montreal classification for the extent of UC[29] based on the maximal disease extent recorded previously by ICD10-codes as follows (see also Annex 3 Classification codes, section 1):

Extent	**ICD-10 codes**	
E1: Ulcerative proctitis	K51.2	Ulcerative proctitis
E2: Left sided UC (distal UC)	K51.3	Ulcerative proctosigmoiditis
	K51.5	Left sided colitis
	K51.5A	Left sided proctocolitis (Danish variant)
	K51.5B	Left sided hemicolitis (Danish variant)
E3: Extensive UC (pancolitis)	K51.0	Ulcerative pancolitis
Unclassifiable extent	K51	Ulcerative colitis
	K51.4	Inflammatory polyp
	K51.8	Other ulcerative colitis
	K51.8B	Ulcerative colitis with oral manifestations
	K521.9	Ulcerative colitis, unspecified

- Being a recent switcher from another biologic agent: the time period following a switch may represent effects of the acute illness that prompted the change in therapy, rather than effects of the new therapy.
- Co-morbidities: Co-morbidities will include primary sclerosing cholangitis, as well as conditions for which anti-TNF therapy is also indicated, including rheumatoid arthritis, psoriasis, psoriatic arthritis, and ankylosing spondylitis. CD will be included as a comorbidity if registered prior to a switch or an outcome (see Annex 3 Classification codes, section 1 for specification of diagnosis codes).
- Colonoscopy and sigmoidoscopy: In UC, colonoscopy is used for diagnosis, and both colonoscopy and sigmoidoscopy are used for evaluation of flares, and as a screening tool for malignancy. It is possible that the frequency of colonoscopy/sigmoidoscopy is a marker of chronic inflammatory load and possibly recent acute disease activity. In automated data, one cannot differentiate the indication for either procedure. The count of colonoscopies/sigmoidoscopies in the 12 months preceding cohort entry (or preceding switch in therapy) will be measured

Final specification of the covariates will appear in the SAP.

9.3.5. Follow-up time, switches and definitions of risk windows

Follow-up will commence at the time of cohort entry. In the separate analysis of each outcome, follow-up time will be censored at the earliest of: occurrence of the outcome of interest, death, emigration, or date of end of study. In addition, because patients whose colons have been removed are no longer at risk for colon cancer, colectomy will be a censoring event in the analysis of CRC.

Time at risk will be further categorized according to changes in treatment, with inference concerning discontinuing treatment and switching therapies made as follows.

Ending therapies. Discontinuation of treatment is not registered explicitly but must be inferred by absence of recurrent dosing (or dispensing) after a clinically expected interval. In both Denmark and Sweden, for medications obtained from a pharmacy, duration of treatment will be inferred from a period from the date of latest registered redemption until a date corresponding with expected days supplied based on the amount of drug dispensed plus 30 days as lag time ('grace period'). Regarding medications ascertained from the hospital registries, such as the anti-TNF agents in Denmark, discontinuation of therapy will be inferred from a period of three months after the last registered treatment of an anti-TNF agent (In DK, typical practice is to dispense a 3-month supply of anti-TNF agents from gastroenterology clinics.)

Switching therapies. If a patient's therapies subsequently change, from an analytic perspective, the patient can leave one exposure category and enter another. In particular, patients can enter the "other anti-TNF category" several times, depending on whether they have previously used none, one, or two or more anti-TNFs. Because the underlying severity of disease likely increases as one has moved beyond the initial anti-TNF, we judge that switches to subsequent anti-TNFs should constitute a different exposure category.

Exposure definitions have been informed by clinical knowledge about treatment pathways and knowledge that over the course of several years, a patient may be exposed to several agents that may have an effect on study endpoints. A theoretical solution to the problem would be to focus on patients with long-term monotherapy exposures, but based on current practice, such patients will be very rare (and also non-representative of the population of UC patients treated with study drugs). Rather, current clinical knowledge suggests several potential combinations of relevant treatments, shown in Table 2 and to be fully specified in the SAP.

Table 2. Outline of exposure categories by defining drug and history of relevant combinations of treatment at cohort entry

Cohort defining drug	TNFI history	Thiopurine history
GLM	TNFI naïve	Never, Prior and/or current
	Previous other TNFI	Never, Prior and/or current
Other TNFI	TNFI naîve	Never, Prior and/or current
	Previous other TNFI	Never, Prior and/or current
Thiopurine	TNFI naïve	Not applicable

Risk windows. The definition of risk window associated with each course of treatment is further explained below. Note that because of the different potential for latent drug effects after drug discontinuation, the risk window for CRC differs from the colectomy outcome.

For attribution of CRC. 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 conceptual risk window for this exposure is "once exposed, always at risk". Anticipating that treatments will be dynamic for many patients in this study, we proposed the following scheme for attribution of outcomes to specific exposures.

- Thiopurine group: the risk window begins one day after initiation of thiopurine until end of follow-up or switch to an anti-TNF, whichever occurs first. If patients switch from thiopurine to a biologic agent, the PY and events before the switch will attribute to the thiopurine group; after the switch to the anti-TNF exposure groups.
- GLM and other anti-TNFs group: The risk window begins one day after initiation of the agent and extends until end of follow up, regardless of drug discontinuation. If patients switch from one biologic to another biologic agent, subsequent PY and events will attribute to both agents.

Because the actual biologic risk window for neoplastic risk 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 exposure discontinuation:

- Alternative scenario 1: For all study exposures, the risk window begins 6 months after start of current exposure and ends 6 months after exposure discontinuation;
- Alternative scenario 2: For all study exposures, the risk window begins 6 months after start of current exposure and ends 2 years after exposure discontinuation

For attribution of all-cause total colectomy, in the base-case analysis, the risk window begins one day after exposure initiation and extends through 90 days after the last treatment, or until one of the general triggers for end of follow-up occurs, whichever comes first. In a sensitivity analysis, we would use an alternative risk window extending to 6 months after the last treatment, or until one of the general triggers for end of follow-up occurs, whichever comes first. Because the therapeutic effect of thiopurines is not expected until approximately 3 months into therapy, in any comparison of GLM with TPs for colectomy, the alternative risk window starts 90 days after the start of therapy and ends 90 days after the last treatment, or until one of the general triggers for end of follow-up occurs, whichever comes first.

9.4. Data Sources

The principal data sources for this protocol are central nationwide health registries providing automated data. Reference is made to the section 9.2.3 concerning details about data sources, the reasons underlying the choices of data sources and how data linkage between data sources will be used.

9.4.1. Study Procedures

The handling of data in the present study involves a series of steps and requires submissions of applications and approvals for access to data in Denmark and Sweden:

1. The Danish and Swedish national scientific coordinators will submit applications to the relevant authorities and agencies for permission to perform the study and to get access to data, including The Danish National Data Protection Board and the Danish National Board of Health and the Swedish National Board of Health and Welfare.
2. The Danish and Swedish national scientific coordinators will establish the national study populations and the ascertainment of all relevant events, after assembly of the inception cohort of new users of GLM and the comparator cohorts. Each national scientific coordinator is responsible for acquiring and validating the datasets.

9.5. Study Size

The study population comprises all patients with a diagnosis of UC in Denmark and Sweden. During the preparation of this protocol aggregated data on UC diagnoses were requested from the health authorities in Denmark and Sweden. It is estimated that the current prevalent population of patients with UC covers approximately 20,000 patients in Denmark and approximately 35,000 patients in Sweden, totalling approximately 55,000 patients.

The number of patients to be enrolled in the inception cohort of new users of GLM and the comparator cohorts over the full study period is unknown and depends on uptake of the new treatment, but will be estimated prospectively with each annual data extraction.

Power considerations, statistical precision

There are no prior hypotheses specified for testing as part of this study. While there are no formal calculations of power, it is possible to estimate the amount of statistical information to be obtained from the study.

A tentative assessment of detectable risk ratios under reasonable assumptions is provided here. Based on Merck's commercial forecasts, across Denmark and Sweden, an estimated 100 patients per year will initiate GLM for UC. Thus, after the first 5 years after marketing, about 500 UC patients are anticipated to have initiated GLM therapy, accruing approximately 1,250 PY of observation. This assumes that observation time continues to accrue for all initiators regardless of drug discontinuation, a reasonable assumption for a cancer outcome that can have a long latency. Given the comprehensive health registries in Denmark and Sweden, we can reasonably assume negligible loss to follow-up (only to death and emigration). The expected accrual of person-time at risk for GLM over different study periods appears in Table 3.

Table 3. Anticipated follow-up time for GLM cohort*

Study duration (years)	Cumulative number of patients exposed	Estimated cumulative follow-up time (PY)
5	500	1,250
7	700	2,450
8	800	3,200

*Follow-up time is counted regardless of drug discontinuation

The literature reports variable rates of CRC for UC cohorts. For this exercise, we consider scenarios over a range of incidence rates for the endpoint of CRC ranging from 1/1,000 to 4/1,000 PY in populations of UC patients unexposed to GLM; the alpha level is set at 5%; the power is set at 80%; the ratio of GLM to comparator group is estimated at 1:5.[11,13,28] Based on these assumptions, Table 4. shows the detectable relative risks over a range of scenarios, calculated using Dupont and Plummer's PS computer program.[30]

Table 4. Detectable Relative Risk versus an unexposed UC population based on various annual incidence rates of colorectal cancer

Follow-up time for GLM cohort (PY)	Annual incidence rates of colorectal cancer			
	1/1,000 PY	2/1,000 PY	3/1,000 PY	4/1,000 PY
1,250	5.9	4.0	3.3	2.9
2,450	4.0	2.9	2.5	2.2
3,200	3.5	2.6	2.2	2.0

9.6. Data management

Reference is made to Section 9.4.1. Study Procedures. Data will be managed at national level under the responsibility of the national coordinators in Denmark and Sweden, respectively, according to the requirements and conditions defined by the health authorities in each of the respective countries. This will be specified in connection with obtaining the relevant permissions.

Data management will be further described in the SAP.

9.7. Data analysis

The statistical analysis will address each of the study objectives based on a conceptual framework for exposures and outcomes described above. The final analytic approach will be described in a stand-alone SAP, which will be completed before any comparative analyses are undertaken. Basic descriptive tabulations and the statistical analyses will be performed at national level in Denmark and Sweden. Data management and analysis must be performed in the server environment chosen by the data provider, due to legal data protection requirements. The choice of software used for the statistical analysis will thus be determined by the data provider, but will certainly support reproducible data management procedures with full documentation of scripts and algorithms.

The analysis of combined data from Denmark and Sweden will be done on the basis of an aggregate data set, prepared using identical table shells. Prior to any combination of data, however, analyses will be stratified by country to evaluate any heterogeneity of findings. Moreover, the combined analysis will treat country as a stratification variable, and results will include findings stratified by country. Comparative analyses among study exposures will evaluate country by exposure interactions.

The main focus is to investigate the occurrence of outcomes as a function of exposures to treatment. In this pathway the study must take into consideration the potential role of confounding factors, as well as competing risks.

Confounding: As noted in section 7.2, the choice of medical treatment for UC depends on disease severity and extent. At the same time, these characteristics are also risk factors for the study endpoints, CRC and total colectomy. Thus, data for this study are at risk for confounding by indication. UC is a chronic disease, and over time, treatments may change, as may the course of disease.

Competing risks: During the clinical course of UC, the risk of developing CRC will compete with the risk of having a colectomy, and vice versa. Colectomy precludes (at least in theory) subsequent risk for CRC, and CRC nearly always leads to treatment by colectomy, which precludes the risk of colectomy.

The data analysis plan contains two main parts: (1) Descriptive analyses and (2) comparative analyses of the inception cohort of new users of GLM users and the comparator cohorts. Several sensitivity analyses have also been specified, including a quantitative bias analysis (QBA) that addresses the potential effect of missing confounder (history of IFX) information in Sweden.

9.7.1. Descriptive analyses of the inception cohorts

1. The assembly of the study population will be described, including attrition figures that indicate why potential subjects did not qualify for study. The number of patients enrolled in each of the three inception cohorts defined previously will be enumerated for each calendar year during the study period, together with the number of PY accrued within each cohort.
2. Demographic details and baseline clinical characteristics based on automated data will be tabulated for each cohort. Covariates will include demographic, temporal, and clinical characteristics, as described in section 9.3.4. 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.
3. Within each inception cohort, switches in treatment will be identified and described by category of switch, according to the exposure group categories outlined in Section 9.3.5. Potential determinants of switches will be explored.
4. Analyzing outcomes of CRC and total colectomy in each of the inception cohorts.
    a. The frequency of each study outcome will be described, stratified by inception cohort and country.
    b. Incidence rates of each outcome will be calculated along with 95% confidence intervals (CIs). Incidence rates will be defined as the number of events divided by the PY at risk

c. Cumulative incidence proportions (using standard survival analysis techniques including Kaplan-Meier estimations) of the occurrence of each primary outcome.

d. Steps 4a-c will be conducted first on the basis of 'a fixed cohort approach', i.e. the inception cohorts will be followed as defined at baseline without consideration of any subsequent changes in treatment.

e. Steps 4a-c will also be conducted according to an 'as treated' approach, which will include only exposed person-time at risk, according to the primary risk windows for each outcome (see section 9.3.5)

5. *Estimating rates of HSTCL*. Since HSTCL is expected to occur very rarely, only descriptive incidence rates will be calculated according to inception cohort; no comparative analyses will be conducted.

9.7.2. Comparative analyses

Assuming that inspection of baseline characteristics indicates the cohorts are reasonably comparable, time-to-event across inception cohorts will be compared using standard methods such as log rank test or Cox proportionate hazard methods, to be specified in the SAP.

Separate time-to-event analyses will be conducted to estimate the relative risk of CRC and total colectomy, comparing exposure to GLM to exposure to other anti-TNF agents (the primary comparator group) and to TP exposure (assuming cohort comparability). 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 9.3.5. Time-dependent covariates will be updated based on most current available data prior to switch to a new exposure group (e.g., from IFX to GLM). Hazard ratios will be used to estimate relative risk and will be adjusted for confounding factors.

Candidate variables to evaluate as potential confounders are listed in section 9.3.4. Adjustment for confounding will occur through stratification and multivariate adjustment. A change of 10% or more between the baseline and the adjusted point estimate will be considered evidence of confounding. Variables that qualify as confounders will be selected for the multivariate models; however, 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. A priori, it is expected that country and time period immediately following a switch in treatment (for colectomy outcome) will function as 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.

9.7.3. Sensitivity analyses

Several sensitivity analyses are planned to examine the robustness of results to alternate specifications of the study population, risk window, outcome definitions, competing risk, and

additional adjustment using QBA to account for missing information on a confounder. The SAP will contain the detailed descriptions of these analyses which are outlined as follows to address the potential influence of:

Including patients who did not truly have UC in the study population. The comparative analyses will be rerun on subsets of the original study population (1) excluding patients who ever carried a diagnosis of Crohn's disease and (2) limited to patients who had at least 2 recorded diagnoses of UC by the time of cohort entry.

Alternative specifications of the risk window. The comparative analyses will be rerun using the alternative definitions of the risk windows for CRC and total colectomy, as stated in section 9.3.5.

Imputing classification of maximal extent of UC for patients where data are missing. Patients without a registration of maximal extent of disease will be assigned minimal extent (E1, see 9.3.4) and maximal extent (E3, see 9.3.4), respectively, to assess the robustness of using extent of disease as a covariate.

Alternative definition of the total colectomy outcome. Instead of using the all-cause total colectomy definition, the main comparative analyses will be repeated to estimate the comparative risk of "colectomy for intractable disease". Although the indication for performing a colectomy is not available in the hospital activity registration system or other sources of automated data, the alternative outcome can be estimated using the following clinical logic. A colectomy performed during hospitalization where the principal activity diagnosis is CRC (ICD10-codes C18. – C20.) will be considered as surgery to treat CRC. All other total colectomies will be assumed performed for intractable disease.

Distortion of risk estimates because of competing risks. Two approaches will be used. First, cumulative incidence analyses will be repeated using censoring criteria modified from the base case. Instead of censoring follow-up for individuals who experience a competing risk in the analysis of each specific risk, these patients will be retained in the risk pool. Thus, for the outcome of CRC, the cumulative risk will be estimated *without* censoring at colectomy. For the outcome of total colectomy, the cumulative risk will be estimated *without* censoring at development of CRC. The second approach to address the competing risk issue will involve a sensitivity analysis that describes cumulative event-free incidence of survival, using a composite definition of outcome that includes occurrence of CRC, colectomy (for any cause), or death.

Managing new treatments. New biologic therapies (including vedolizumab) will be approved with UC as an indication after 13-Sept-2013, when the study period begins. Treatment with such drugs will be managed as covariates, but will also be ignored to assess their impact of the effects of the study drugs.

Distortion of risk estimates because of residual confounding. In this study, IFX use is a potential exposure, but history of its use is also a potential confounder. Patients initiating ADA or GLM who are bio-naïve have a lower risk of colectomy than those who have previously received biologic therapy.[9,10] IFX is systematically under-ascertained using Swedish registry data, and if history of its use differs between GLM and ADA initiators, confounding is likely. A QBA is proposed to address this issue. The QBA would start with the main comparative analysis in Sweden for the outcome of total colectomy, limiting the

contrast to GLM vs ADA exposures. The adjusted HR (HR_{adj}) from that comparison would be further adjusted in the QBA using two external data inputs, using the approach of Lash and colleagues.[31] The first input, the prevalence of prior IFX use in UC patients initiating ADA and GLM, will come from an independent bias-estimation study that the Sponsor is conducting in Sweden. The population for that project will be assembled by applying the entry criteria for the current study to the national registry data for a subset of Sweden for whom there is access to a database of EMRs, which is expected to include complete information on use of IFX.

The second input for the QBA is an estimate of the association between the unmeasured confounder (prior IFX exposure) and the colectomy outcome. This estimate (and its distribution) will be synthesized from literature, supplemented with queries to authors if publications do not include sufficient detail. For example, a recent retrospective cohort study in patients with UC reported that when GLM was used as second line and third line therapy, the risk of colectomy was substantially increased, but published results were expressed in terms of cumulative incidence, rather than relative association.[9] Correspondence with the author clarified that the HR for colectomy was 7.2 (95% CI, 1.0-64.5) and 7.8 (95% CI, 1.1-62.4), respectively, compared to first-line use (C.Taxonera, personal communication, 19 July 2018).

Based on the two inputs (which represent point estimates with probability distributions), multiple simulations are performed on the HR_{adj} from the automated data analysis comparing GLM with ADA. The output thus represents a further adjustment of original $HR_{adj,}$ which now factors in information about the unmeasured prior IFX exposure that was obtained from sources external to the main study. Further details will be provided in the SAP.

9.8. Quality Control

By signing this protocol, the Sponsor agrees to be responsible for implementing and maintaining a quality management system with written development procedures and functional area standard operating procedures to ensure that studies are conducted and data are generated, documented, and reported in compliance with the protocol, accepted standards of Good Clinical and Pharmacoepidemiology Practice, and all applicable federal, state, and local laws, rules and regulations relating to the conduct of the study.

The study makes, to a large extent, use of existing automated data, contained in central health registries established and operated for administrative purposes not related to the present study. The quality of the data in these registries is generally considered to be high, but data errors may inevitably occur. See also Section 9.9 (Limitations of the research methods).

Quality control of data includes, but is not limited to, a check for legitimate values for each categorical variable, and check for consistency between dates (e.g. date of birth precedes all other dates).

9.9. Limitations of the research methods

9.9.1. General assumptions and limitations:

- All necessary permissions will be obtained in order to obtain access to the relevant national health registries.
- Comparative analyses between GLM and alternative exposures may be 'confounded by indication' if GLM is systematically channelled to sicker patients.
- It is possible that not all relevant confounders will be captured, and it is always possible that not all confounders will be perfectly measured.

9.9.2. Specific limitations

A number of specific limitations apply to the study:

Ascertainment of patients with UC

- It is assumed that the patients with UC requiring the therapies of relevance for this study will be ascertained with a very high level of completeness from the hospital activity and prescription registrations systems. However, up to 20% of the patients registered with a diagnosis of UC may also be registered with encounters where CD has been recorded as diagnosis, as IBD phenotype may change, especially shortly after diagnosis.[32] This may happen before, as well as after, the first registered activity with UC as the primary diagnosis. Patients who have received both diagnoses will be included in the study base only to the extent that UC is the main diagnosis registered prior to enrolment in the inception cohorts. To address this concern, a sensitivity analysis is planned that excludes patients who also carry a CD diagnosis from the study population.

Capturing information on treatment, changes in treatment and exposure windows

- Treatments for UC are rapidly evolving. This protocol is based on the current therapeutic landscape. However, as exemplified with vedolizumab, new agents may be introduced over the course of the study. Such drugs will in the first place be dealt with as covariates in the statistical analysis, but it may become necessary to add an additional comparator group to the study if there is substantial exposure to such drugs.
- In Denmark, errors or nonspecific coding of relevant therapies may cause exposure misclassification and inappropriate exclusion from the study if treatment with biologic agents is coded at high-order level rather than the detailed level where the specific drug is identified. According to experiences from the Pilot Study, less than 2% of the treatment codes registered from 2012 and onwards are high-level, unspecific treatment codes.
- Ascertainment of hospital-based IFX infusions are expected to be under-ascertained in automated Swedish registry data. To the extent that prior use of IFX is a marker of a higher risk of colectomy, missing information on IFX could confound the relationship between GLM vs ADA if prevalence of prior IFX varied by exposure. A probabilistic QBA is planned to address this issue, as explained in section 9.7.3.

- The treatment with study drugs may be indicated by CD or for dermatological and rheumatologic indications. In some instances patients may be exposed to these therapies before the first registration of UC (or CD), and the same treatment may be administered under different indications over time for the same patient. According to experiences from the Pilot Study it is not possible from the automated data to obtain information on the precise indication for any given treatment. For this protocol, study drugs are included in the analysis regardless of underlying indication. However, patients need to have at least one main diagnosis of UC prior to study drug initiations.
- Whereas the commencement of a new therapy may be established as the first time ever, the event of discontinuing a therapy will not be explicitly registered in the automated data sources, but will need to be established indirectly.
- Treatment crossover may limit ability to tease out independent drug effects, particularly if a large number of patients will be registered with multiple exposures and crossovers. One solution would be to confine analysis to those who never switch, but given current knowledge of dynamic UC treatment, such patients will likely be rare, and certainly not representative of all patients newly started on study drugs.
- Assuming there is an effect of study drugs on the outcomes of interest, the true biologic risk windows are not known for the outcome of CRC. Our base case analyses for this outcome assume that for biologic agents, the risk window is "once exposed, always at risk"; secondary analyses are planned to explore the effect of alternative specification of risk windows.

Capturing information on outcomes

- The diagnoses of CRC made by colonoscopy may be subject to detection bias; patients perceived to be at highest risk may undergo more frequent colonoscopic and sigmoidoscopic surveillance. Analyses are planned to adjust for frequency of colonoscopy and sigmoidoscopy, as measured in the periods prior the start of the patient's current treatment.
- In general, errors and inappropriate coding of surgical procedures in the central health registries may cause deficits or misclassifications in the identification of outcomes. In Denmark, however, the PPV of total colectomy using national registry data, validated against clinical records, was shown to be high.[25] The validity of total colectomy codes in the Swedish register is being investigated by the Sponsor in a separate methods project.
- Given the underlying rarity of the condition, estimates of HSTCL incidence are expected to be imprecise.

Information on covariates

- Severity and activity of UC and variations over time for the individual patients will, to the extent possible, be ascertained by proxies from the automated data, to be used as covariates in the analysis of outcomes in the inception cohort data.
- Information on extent of UC is inferred from the ICD10-codes registered as activity diagnosis for hospital encounters. Accordingly, the validity of this information

depends on the level of details used in daily coding practice. From the first annual captures of data information on extent of UC has been specified this way for 80% or more of the study patients.

- The time window concerning availability on information on history of therapies varies between registers and between Denmark and Sweden. However, from year 2005 and onwards drug use as captured from prescription registers and hospital administered therapies as captured from hospital activity registrations are available on equal terms in both Denmark and Sweden. This is considered satisfactory for the purpose of the present study.

9.10. Other Aspects

None.

10. Protection of human subjects

This is a non-interventional observational study with no administration of any therapeutic or prophylactic agent according to the study protocol. Patients observed in this study will continue with the standard of care as provided by their physician. National registries of cancer, death, hospital contacts and socio-economic factors will be the main data source.

10.1. Informed Consent

According to Danish and Swedish law, registry-based studies can be carried out without consent from the study subjects where the processing takes place for the sole purpose of carrying out statistical or scientific studies of significant public importance and where such processing is necessary in order to carry out these studies. It is an absolute requirement that the publication of statistical or scientific results may never reveal the identity of individuals or otherwise compromise study subjects. Approval will be obtained by the data agencies in the two countries before data management and data analyses will be performed.

10.1.1. Consent and Collection of Specimens for Future Biomedical Research

Not applicable.

11. Management and reporting of adverse events

Adverse Event Reporting

This is a non-interventional database study based on secondary use of data collected for other purposes. No administration of any therapeutic or prophylactic agent is required in this protocol. No reporting of individual adverse events to regulatory agencies is planned for this database study because there is no access to individual patient/subject records and it is not possible to assess the causality of individual cases. Pre-specified health outcomes of interest, including any that qualify as adverse events, will be summarized as part of any interim analysis (including safety analysis, if required) and in the final study report, which 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/Periodic Benefit Risk Evaluation Report and/or Development Safety Update Reports if required.

If an investigator elects to spontaneously report any suspected adverse reactions, they should be reported via fax to Merck AER Mailbox FAX #PPD (US), or toll-free fax PPD (ex-US and US availability), in English using an AE form (Annex 1: stand-alone document no 2) for reporting to worldwide regulatory agencies as appropriate.

12. Plans for disseminating and communicating study results

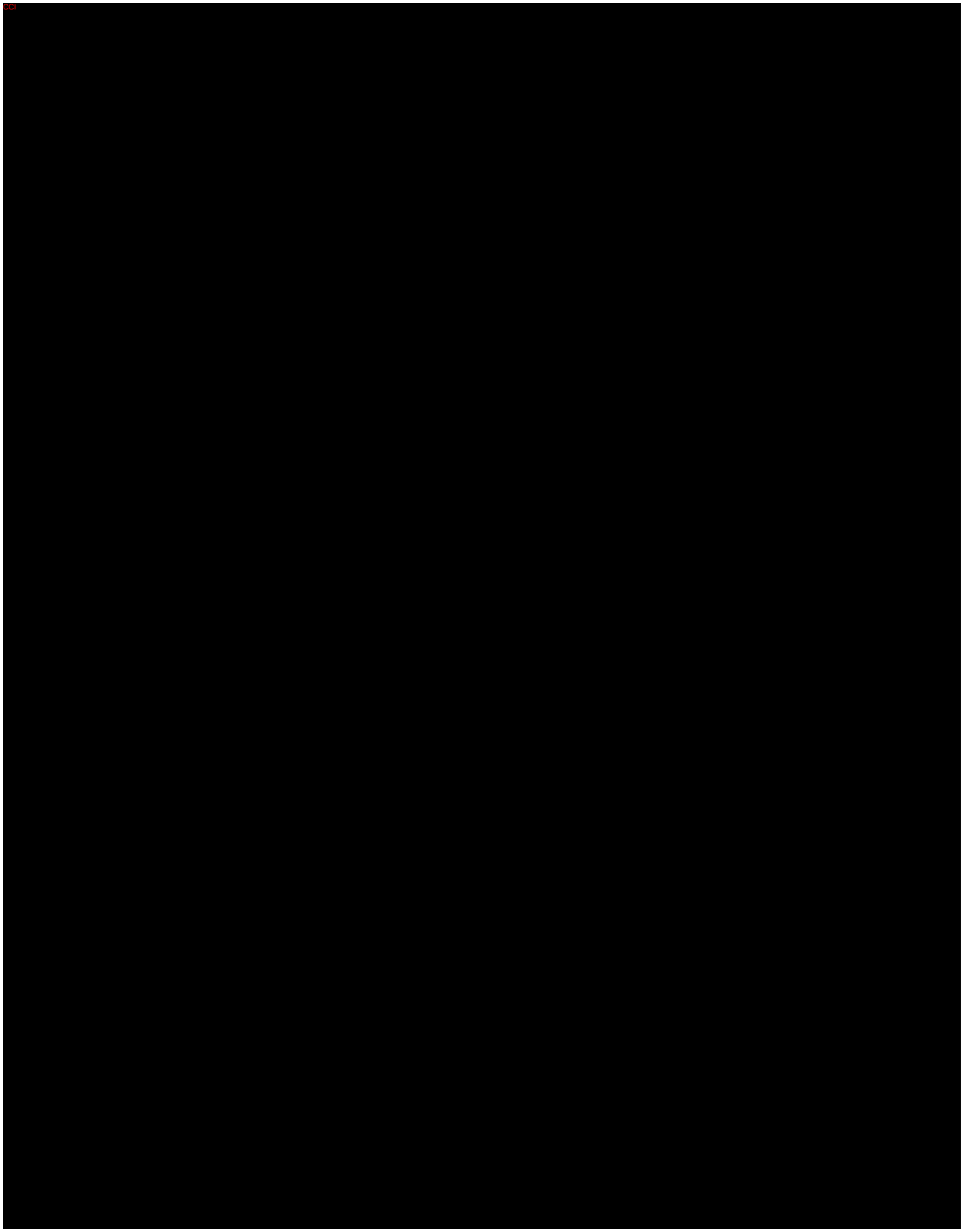
Progress reports will be submitted to the European Medicines Agency at several points during the study, according to the schedule under Milestones (Section 6). Each progress report will include information about the number of patients who have entered the study in

each of the treatment cohorts (the GLM cohort, the 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. The final study report 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. Authorship will be determined for each manuscript in accordance with the Vancouver rules. 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 comment prior to submission of the manuscript for publication.

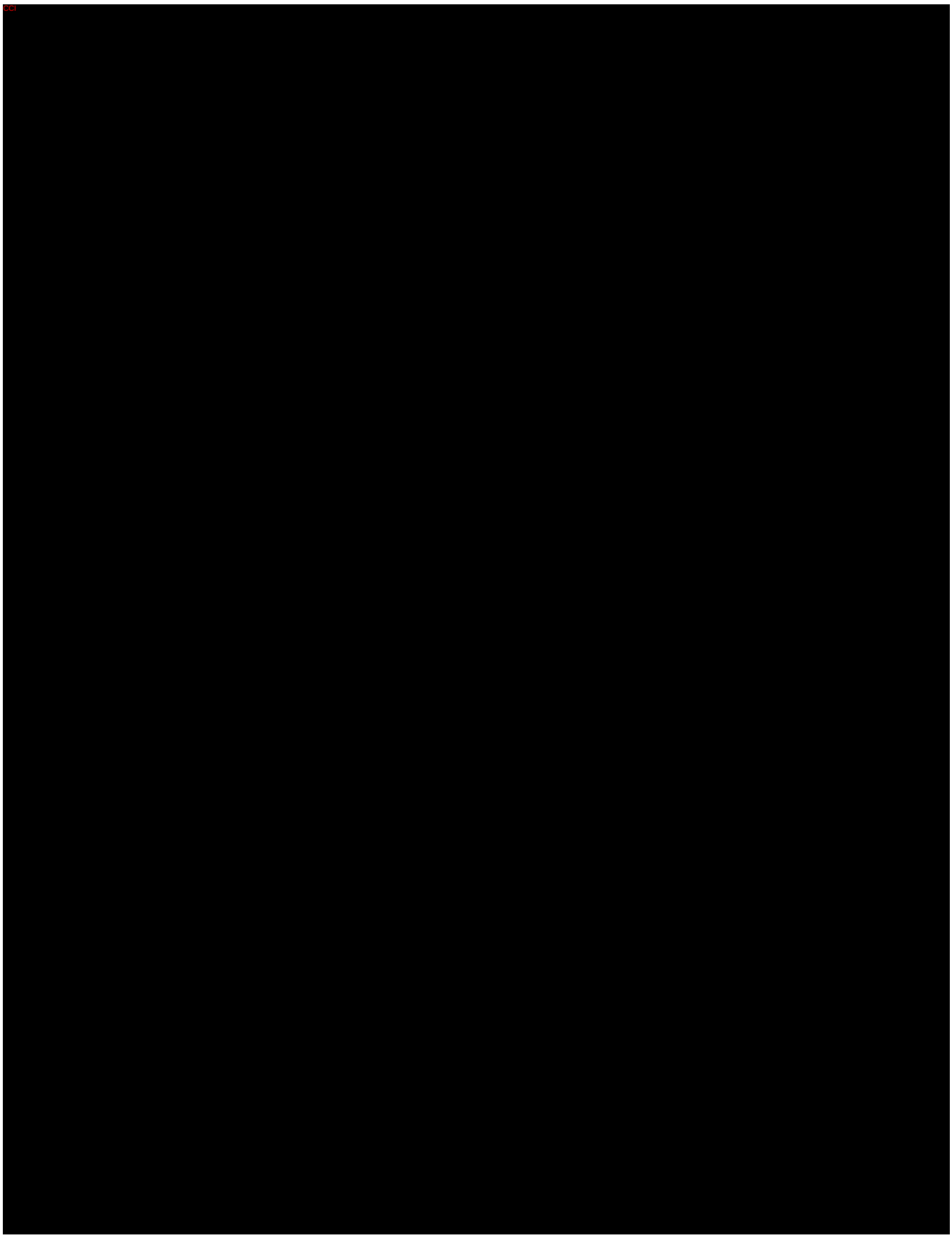
13. References

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; quiz e14-85.
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.e101.
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. Meier J, Sturm A. Current treatment of ulcerative colitis. *World J Gastroenterol.* 2011;17(27):3204-3212.
5. 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.
6. Canas-Ventura A, Marquez L, Ricart E, et al. Risk of colectomy in patients with ulcerative colitis under thiopurine treatment. *J Crohns Colitis.* 2014;8(10):1287-1293.
7. 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.
8. 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.
9. Taxonera C, Rodriguez C, Bertoletti F, et al. Clinical Outcomes of Golimumab as First, Second or Third Anti-TNF Agent in Patients with Moderate-to-Severe Ulcerative Colitis. *Inflamm Bowel Dis.* 2017;23(8):1394-1402.
10. Taxonera C, Iglesias E, Munoz F, et al. Adalimumab Maintenance Treatment in Ulcerative Colitis: Outcomes by Prior Anti-TNF Use and Efficacy of Dose Escalation. *Dig Dis Sci.* 2017;62(2):481-490.
11. Loftus EV, Jr. Epidemiology and risk factors for colorectal dysplasia and cancer in ulcerative colitis. *Gastroenterol Clin North Am.* 2006;35(3):517-531.
12. Van Assche G, Dignass A, Bokemeyer 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.
13. Eaden JA, Abrams KR, Mayberry JF. The risk of colorectal cancer in ulcerative colitis: a meta-analysis. *Gut.* 2001;48(4):526-535.
14. Rutter MD, Saunders BP, Wilkinson KH, et al. Thirty-year analysis of a colonoscopic surveillance program for neoplasia in ulcerative colitis. *Gastroenterology.* 2006;130(4):1030-1038.
15. Castano-Milla C, Chaparro M, Gisbert JP. Systematic review with meta-analysis: the declining risk of colorectal cancer in ulcerative colitis. *Aliment Pharmacol Ther.* 2014;39(7):645-659.
16. 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.

17. Nyboe Andersen N, 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.
18. 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.
19. Thygesen LC, Daasnes C, Thaulow I, Bronnum-Hansen H. Introduction to Danish (nationwide) registers on health and social issues: structure, access, legislation, and archiving. *Scand J Public Health.* 2011;39(7 Suppl):12-16.
20. Lynge E, Sandegaard JL, Rebolj M. The Danish National Patient Register. *Scand J Public Health.* 2011;39(7 Suppl):30-33.
21. Fonager K, Sorensen HT, Rasmussen SN, Moller-Petersen J, Vyberg M. Assessment of the diagnoses of Crohn's disease and ulcerative colitis in a Danish hospital information system. *Scand J Gastroenterol.* 1996;31(2):154-159.
22. Pedersen CB. The Danish Civil Registration System. *Scand J Public Health.* 2011;39(7 Suppl):22-25.
23. Kildemoes HW, Sorensen HT, Hallas J. The Danish National Prescription Registry. *Scand J Public Health.* 2011;39(7 Suppl):38-41.
24. Engholm G, Ferlay J, Christensen N, et al. NORDCAN--a Nordic tool for cancer information, planning, quality control and research. *Acta Oncol.* 2010;49(5):725-736.
25. Erichsen R, Sværke C, Tøttrup A. The coding quality of acute total colectomy in inflammatory bowel disease patients in a nationwide hospital registry. *Pharmacoepidemiol Drug Saf.* 2011;20:S165-S166.
26. 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.
27. 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.
28. 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.e168.
29. Satsangi J, Silverberg MS, Vermeire S, Colombel JF. The Montreal classification of inflammatory bowel disease: controversies, consensus, and implications. *Gut.* 2006;55(6):749-753.
30. Dupont WD, Plummer WD, Jr. Power and sample size calculations. A review and computer program. *Control Clin Trials.* 1990;11(2):116-128.
31. Lash TL, Fox MP, MacLehose RF, Maldonado G, McCandless LC, Greenland S. Good practices for quantitative bias analysis. *Int J Epidemiol.* 2014;43(6):1969-1985.
32. Henriksen M, Jahnsen J, Lygren I, et al. Change of diagnosis during the first five years after onset of inflammatory bowel disease: results of a prospective follow-up study (the IBSEN Study). *Scand J Gastroenterol.* 2006;41(9):1037-1043.



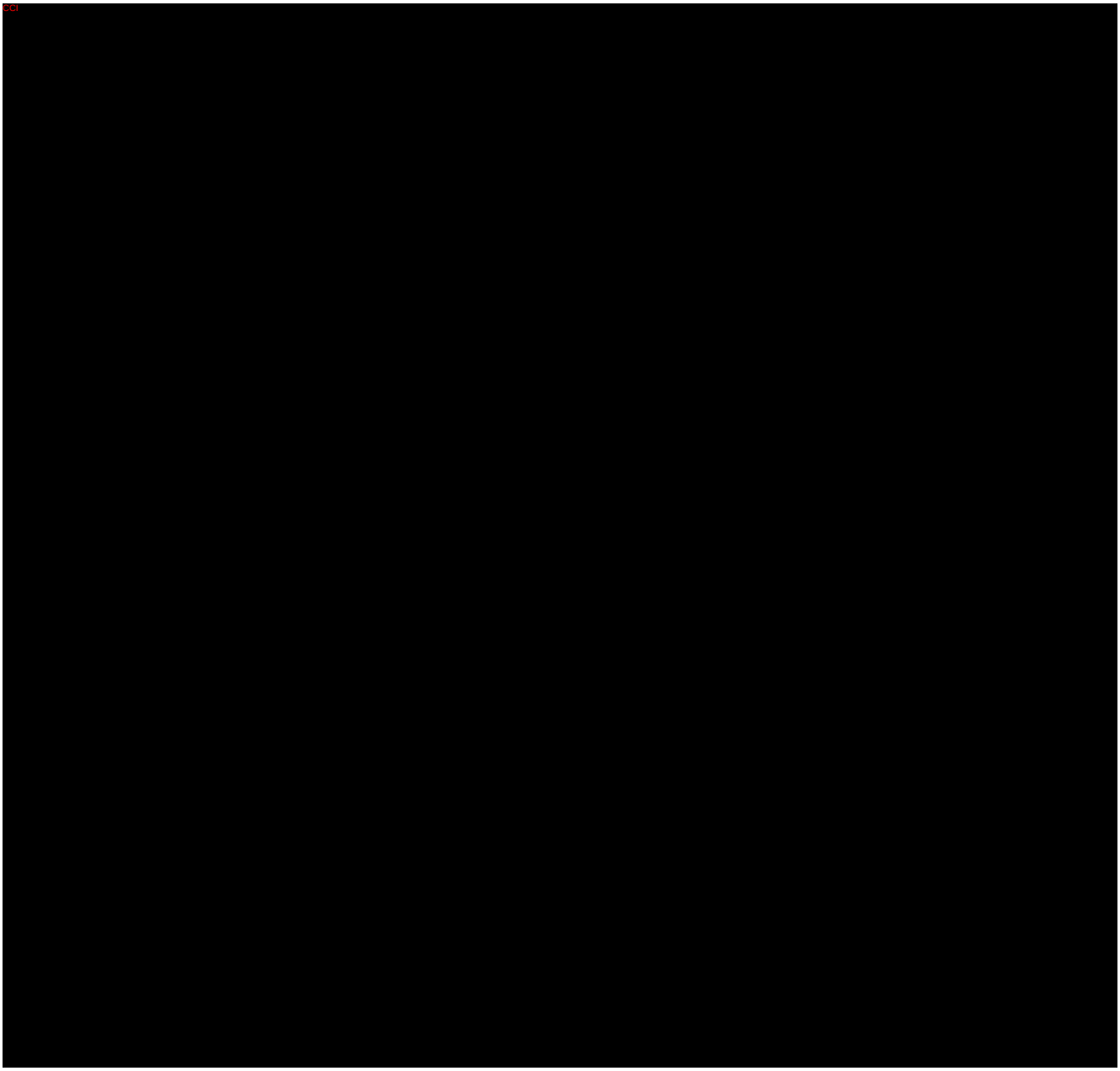
[redacted]



[redacted]

CCI

[redacted]



Name of the main author of the protocol: PPD

Date: 10/AUGUST/2018

Signature: _ PPD