Retrospective Study Protocol / EPD217

COMORBIDITIES AND DISEASE MANIFESTATIONS IN ANKYLOSING SPONDYLITIS: AN ANALYSIS OF US CLAIMS DATABASES

SPONSOR ADDRESS

Final Retrospective Study Protocol	**2 MAY 2017**
Retrospective Study Protocol Amendment 1	<dd MMM yyyy>
Retrospective Study Protocol Amendment 2	<dd MMM yyyy>

STUDY CONTACT INFORMATION

Marketing Authorization holder/Sponsor

Name:	Robert Suruki, Sc.D.
Affiliation:	UCB Biosciences Inc.
Address:	8010 Arco Corporate Drive, Raleigh, NC 27617
Phone:	919-767-1117
Email:	robert.suruki@ucb.com

UAB Site PI

Name:	Jeffrey Curtis, MD MS MPH
Affiliation:	University of Alabama
Address:	510 20th Street South, Birmingham AL 35294
Phone:	205-975-2176
Email:	jcurtis@uab.edu

Additional UAB Co-investigators and Personnel

Name:	Dr. Huifeng Yun (Epidemiologist) Dr. Lang Chen (Sr. Statistician) Fenglong Xie (Statistician) Shuo Yang (Programmer) Kathy Parham (Program Coordinator)
Address:	As above
Phone:	
Fax:	

OHSU Site PI

Name:	Atul Deodhar, MD
Address:	Oregon Health & Science University, 3181 SW Sam Jackson Park Road, Portland, OR. 97239
Phone:	503 494 8963
Email:	deodhara@ohsu.edu

Additional OHSU Coinvestigators and Personnel

Name:	Kevin Winthrop, MD MPH Ben Chan, MS (statistician/programmer) Sarah Siegel, MPH (epidemiologist)
Address:	Oregon Health & Science University, 3181 SW Sam Jackson Park Road, Mail code: GH 104, Portland, OR. 97239
Phone:	503 494 5496
Fax:	503-346-8277

DECLARATIONS AND SIGNATURES /SPONSORS DECLARATION

I confirm that I have carefully read and understand this Retrospective Study protocol and agree to conduct this retrospective study as outlined in this protocol and according to current Good Pharmacoepidemiology Practice.

Study Owner	Date/Signature
Clinical Project Manager	Date/Signature
Global Statistical Sciences Representative	Date/Signature
Lead Clinical Development Representative	Date/Signature
Medical Affairs Representative	Date/Signature

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AE	Adverse Event
AS	Ankylosing Spondylitis
ATC	Anatomical Therapeutic Chemical classification
CMS	Center for Medicare and Medicaid Services
DMARDs	Disease Modifying Anti-Rheumatic Drugs
DUA	Data Use Agreement
EU	Europe
EMR	Electronic Medical Record
ER	Emergency Room
FDA	US Food and Drug Administration
FFS	Fee for Service
HCPCS	Healthcare Common Procedure Coding System
HMO	Health Maintenance Organization
ICD-10	International Classification of Diseases – 10th Revision
ICD-9-CM	International Classification of Diseases – 9th Revision, Clinical Modification
IRB	Institutional Review Board
N/A	Not Applicable
NDC	National Drug Code
PS	Propensity Score
SAP	Statistical Analysis Plan
SOP	Standard Operating Procedures
TNFi	Tumor Necrosis Factor Inhibitor
US	United States

1 EXECUTIVE SUMMARY

Rationale and Background:

Ankylosing Spondylitis (AS) is a multi-system, immune-mediated, chronic inflammatory disease that predominantly affects the axial skeleton, and commonly affects the peripheral skeleton as well as non-articular organ systems. The extra-articular manifestations of the disease and comorbidities (henceforth combined to call 'comorbidities') lead to increased morbidity and mortality in AS patients compared to the general population. With the advent of novel therapies, especially tumor necrosis factor inhibitors (TNFi), AS patients have experienced reduction in signs and symptoms, improvement in physical function and quality of life. Whether these new modalities of treatment have altered the incidence of comorbidities is not known.

Research Question and Objectives:

The main research question is whether TNFi have changed the natural history of AS by reducing the incidence of AS-related comorbidities. The study objectives are (1) to investigate the prevalence of AS in insurance claims databases, (2) to identify the prevalence of comorbidities in AS patients compared to a non-AS general population sample, and (3) to compare the incidence of AS-related comorbidities across the following groups of AS patients: those managed with either no therapy or prescription non-steroidal anti-inflammatory drugs (NSAIDs), those given traditional disease modifying anti-rheumatic drugs (DMARDS), and those using TNFi.

Study Design (including data sources):

This will be a retrospective cohort study of AS patients from three commercial insurance claims databases: United Healthcare, Truven Marketscan, and the U.S. Medicare Fee-for-Service Claims data.

Study Population:

All AS patients represented in the Multi-Payer Claims Database (MPCD) (2007-2010) which includes information from United Healthcare (Optum Insight), Truven Marketscan (2010-2014), and U.S. Medicare claims data (2006-2014) will be included. Samples of the general population in United Healthcare and Medicare will be used for the non-AS comparators. Entry criteria will include a rheumatologist's diagnosis of AS, six-months

of pre-diagnosis insurance coverage and data availability, and (for drug-specific exposures) administration of AS exposures of interest after the AS diagnosis. For the estimation of the incidence of comorbidities, the data collection will end at the earliest of date of death, lost medical or pharmacy coverage, end of study period, first outcome occurrence, or treatment discontinuation.

Variables:

All variables available in the databases (e.g. age, gender, median household income (as available), geographic location, pre-existing comorbidities etc.) will be collected. The outcomes of interest include disease manifestations and comorbidities (total 13 categories: including cardiac, neurological, kidney, lung diseases, fracture, spondyloarthritis manifestations (such as uveitis, psoriasis, and inflammatory bowel disease), infections (hospitalized and opportunistic), hematologic malignancy, solid tumors, and non-melanoma skin cancer (i.e. basal and squamous cell cancer).

Sample Size, Data Analysis:

The total number of people included in these three databases is approximately 40 million persons. The AS and the comparator non-AS population will be selected from these databases. We will calculate the age and sex standardized prevalence of AS, the prevalence of comorbidities and the incidence rate (with 95% confidence interval) of outcome of interest by treatment exposures stratified by each data source. Crude and multivariable adjusted hazard ratios of outcomes of interest for each exposure (DMARDs, biologics, and NSAIDs or no AS medication) will be calculated and stratified by each data source using Cox regression. Conditional on the homogeneity of the hazard ratios across data sources, the stratified results may be pooled to provide a narrower confidence interval of hazard ratios. Sandwich estimators will also be used to adjust the variance due to patients contributing multiple episodes. The Lin, Wei, and Ying (1993) method will be used to check the proportional hazard assumption. Hazard ratios for patient characteristics will also be provided.

2 MILESTONES

Milestone	Expected Planned date
Data exaction	October 2016
Start of database analysis for feasibility estimation	January 2017
Protocol Approval	May 2017
End of database analysis	November 2017
Final report of study result	January 2018

3 RATIONALE AND BACKGROUND

Ankylosing Spondylitis (AS), the prototypic form of spondyloarthritis, is an immune-mediated chronic inflammatory disease of the axial skeleton (spine and sacroiliac joints) with variable involvement of peripheral joints and non-articular structures. AS is often accompanied by extra-articular manifestations in the cardiovascular, pulmonary, renal, ophthalmic, gastrointestinal, and neurologic systems. In addition, AS patients can have co-morbidities such as heart disease, serious infections, and malignancies. The extra-articular manifestations and the other co-morbidities may increase morbidity and mortality in AS patients. For this study, we use the term 'comorbidities' to include both the extra-articular manifestations of the disease and the conditions mentioned above (please refer to table 2 for a list of disease manifestations and co-morbidities in AS). In comparison to the general population, an excess mortality has been noted in patients with Ankylosing Spondylitis (AS) (1-3). This increase in mortality observed in AS patients is largely associated with increase in the risk of cardiovascular disease (4-8). Along with cardiovascular disorders such as conduction disturbances and valvular disease (aortic insufficiencies or regurgitation) (9-11), accelerated atherosclerosis also puts these patients at greater risk (12). High mortality in AS patients is also closely related to osteoporotic spinal fracture and related complications (13-15). Cauda equina syndrome and spinal cord compression have been known to occur in late stage AS (16).

Ankylosing spondylitis has also been commonly associated with pulmonary apical fibrocystic disease (unilateral or asymmetrical), but most cases eventually consist of bilateral apical fibrobullous lesions, nodules, fibrosis (17), and bronchiectasis.

Opportunistic superinfections of the upper lobe cavities may also occur in AS patients (18). Renal involvement in AS patients, from most frequent to least frequent include: IgA nephropathy, renal amyloidosis, mesangioproliferative glomerulonephritis, membranous nephropathy, focal segmental glomerulosclerosis, and focal proliferative glomeruleonephritis (19).

Recent research indicates that patients with AS might be at increased risk for developing certain co-morbidities (JA Walsh – personal communication). For example, patients with newly diagnosed AS have a 60% higher rate of developing new cardiovascular disease, when compared to a matched general population (JA Walsh – personal communication).

While the exact pathogenesis of most of these comorbidities and disease manifestations is not known, it is generally believed to be related to the underlying inflammatory process. With the use of the new therapeutic modalities targeting this inflammatory process, such as the use of tumor necrosis factor inhibitors (TNFi) in patients with AS during the past decade, we anticipate a change in the prevalence and incidence of such co-morbidities in AS patients (20). Previous studies investigating comorbidities in AS patients are outdated and were mainly conducted in countries other than the United States. In this study, we aim to assess the current burden of comorbidities in AS patients, and how these new therapeutic modalities have influenced the incidence of specific comorbidities in the AS patients within the United States.

The **first Aim of this study** will be broken into two sections. Aim 1 will analyze the (1) prevalence of AS and (2) the prevalence of key AS-associated co-morbidities and disease manifestations (hereafter referred to as ‘comorbidities’ for sake of convenience), in large cohorts of AS patients identified from three administrative databases. These individuals will be compared to a general population without AS or other inflammatory diseases. The **second Aim of the study** will compare the incidence of AS associated comorbidities as shown in Table 2 in patients starting TNFi and non-biologic disease modifying anti-rheumatic drugs (DMARDs) and no AS-specific therapies except for NSAIDs.

4 RESEARCH QUESTION AND OBJECTIVES

Aim I:

- To assess the prevalence of AS in an analysis of Medicare and the United Healthcare databases.
- To assess the prevalence of key co-morbidities in AS patients and the demographic and socio-economic factors (as available) associated with such comorbidities in all three data sources.
- To compare the prevalence of the AS-related comorbidities between the AS population defined as above, and a random sample of the general population without AS or other inflammatory diseases in all three data sources.

Aim II:

- To compare the prevalence and incidence of AS-related disease manifestations and comorbidities in three groups of AS patients: those on no recorded AS therapy or users of NSAIDs, new users of non-biologic DMARDs, and new users of TNFi.

For this objective, the incidence of 13 AS associated comorbidities and disease manifestations will be evaluated in the AS cohort during the follow-up period.

5 METHODOLOGY

5.1 Study design

This is a retrospective cohort study, which will be conducted in three different data sources including Truven MarketScan, the U.S. Medicare Fee-for-Service Claims data, and the United Health data included in the Multi-Payer Claims Database (MPCD; Optum Insight). Our goals are to identify and characterize patients with AS and to estimate the prevalence of AS associated co-morbidities and the incidence rates of events of interest among AS patients on systemic medications.

5.2 Study population

All AS patients (defined below) from the MPCD (2007-2010, excluding patients person-time of individuals with fee-for-service Medicare to avoid overlap), the Truven MarketScan database during 2010-2014, and the U.S. Medicare claims data during

2006-2014 obtained from the Centers for Medicare & Medicaid Services (CMS). A random sample of the general population in MPCD and Medicare will be used for the non-AS comparator; an age- and gender-matched non-AS comparator cohort will be used for the Truven MarketScan database.

5.2.1 Planned number of patients and treatment groups

Because the study is observational in nature, all eligible patients in the database will be included. The sample size is therefore fixed. Feasibility analyses in each of the data sources identified approximately 80,000, 13,000, and 121,000 patients with AS in the Truven MarketScan database, MPCD, and Medicare, respectively.

5.2.2 Methodological techniques for Patient selection

The AS cohort index date is the first date on which both the AS diagnosis code (from a rheumatologist visit) criterion has been met and the patient has had 6 months of continuous medical and pharmacy coverage. This definition will therefore encompass AS patients with both incident and prevalent disease.

Exposure index date is the first date of a new prescription or administration for an AS treatment in the pharmacy or procedure data after the AS cohort eligibility date. New use of a drug is defined as no past exposure to the specific drug using all available data. Those on no treatment on the AS cohort eligibility date, or only receiving NSAIDs, will contribute to the No AS exposure cohort and their exposure index date will be the same as their AS index date. Patients who are prevalent users of a DMARD or biologic drug on the AS cohort index date will not contribute person time to the specific DMARD or biologic drug in the Aim II analysis but could contribute to other DMARD or biologic drug use if they subsequently initiate a new DMARD or biologic. AS patients may therefore have zero, one or more than one treatment index date. It is possible that the AS cohort eligibility date and the first AS exposure episode date occur on the same calendar day.

5.2.3 Patient selection period

The study period will align with the availability of the data from each source. Those eligible for the AS disease cohort may be eligible to contribute follow-up time to one or more AS exposure episodes.

5.2.4 Baseline (pre-index) period

All available data prior to the index date will be used for selected key covariates (i.e. past biologic or DMARD use, and history of all thirteen outcomes of interest). The 6-month period before the AS cohort eligibility date, and before each treatment episode, will be used to establish the baseline period for other key covariates and be considered the primary baseline period for all analyses.

5.2.5 Follow-up (post-index) period

AS cohort

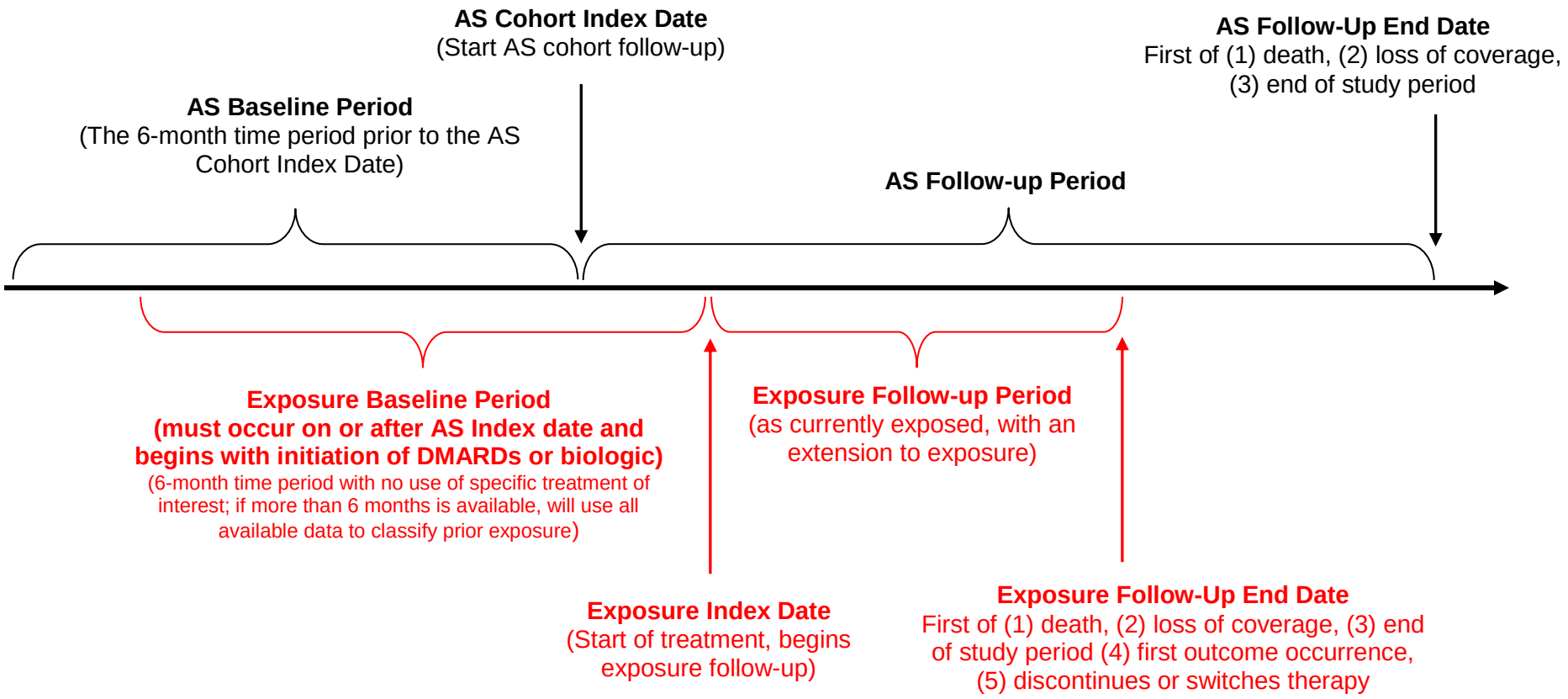
Follow up will start on the first date when both the AS cohort eligibility date and the 6-month coverage requirement have both been met. Follow-up will end at the earliest date of death, lost medical or pharmacy coverage, or end of study period. We will examine various intervals of time (e.g. 2 years, 3 years, see below) over which we will report the point prevalence of the various comorbidities, depending on the length of follow-up in each data source. Patients will contribute only their first enrollment episode to the study; in other words, if they discontinued insurance coverage, and then re-enrolled, only the first would be counted. This generally is relatively uncommon (e.g. <5% of individuals).

AS exposure cohort

Exposure episodes will start at the first prescription for each TNFi and DMARD and end at the earliest date of death, lost medical or pharmacy coverage, end of study period, first outcome occurrence, treatment switch (defined below) or discontinuation (based upon the end of the days' supply plus 90 days).

A sensitivity analysis will be conducted that will consider exposure as first observation carried forward without respect to drug discontinuation. Patients will remain in each exposure group indefinitely until they initiate a new medication higher in the exposure hierarchy (see Section 5.3.1). For example, someone initiating MTX would contribute exposure time to the DMARD category until they initiated a TNFi. They would remain in the TNFi category until the end of the study period, initiating a non-TNFi biologic (e.g. secukinumab) or outcome occurrence. The purpose of this sensitivity analysis is to allow for alternative assumptions regarding the latency between exposure and the outcome. For serious infections, for example, risk is expected to predominantly be associated with only current or recent treatment, whereas for malignancies, the exposure-outcome temporal relationships are less clear.

5.2.6 Schematic diagram – Study period



5.2.7 Inclusion criteria

AS disease cohort

All patients meeting an AS definition in the databases above will be potentially eligible for the AS cohort. The inclusion criteria include:

1) At least one rheumatologist-given diagnosis code for AS. We will descriptively explore the possibility that 2 rheumatologist-given diagnoses may be required, in the absence of any AS-specific treatment other than NSAIDs
2) Patients must have at least 6 months of continuous medical and pharmacy enrollment
3) At least 20 years of age on the AS cohort index date (to avoid overlap of children with JIA/SpA)

Non-AS Disease Comparator cohort

Within MPCD, MarketScan and Medicare, we will construct a comparator population that is free of all AS diagnoses across the entire study period. It would be possible to permit non-AS patients to contribute to the comparator cohort, and then subsequently contribute person-time to the AS cohort, although the extended time frame that it takes most AS patients to be diagnosed makes it probable that patients with this pattern in the data are more likely to have had pre-existing AS and thus should not contribute to the comparator cohort. For Medicare, the comparator cohort is a 5% random sample, and in the MPCD data source, a size-matched random sample, and for MarketScan, an age and gender-matched non-disease cohort. The index date for the comparator cohort is the first date a patient meets the medical and pharmacy coverage requirement. Although follow-up is variable for each person within the datasets and in the two cohorts, we will work to standardize the follow-up time and report in discrete increments of calendar time (described below).

AS Exposure Cohort

In addition to the inclusion criteria in the AS cohort, patients in the AS exposure cohorts are also required to meet the following criteria:

1) At least one prescription for a TNFi biologic or DMARD in these exposure cohorts (**Table 2**). New users of these medications are defined as having no prior use of the index drug using all available data, prior to the drug initiation. By defining new use as no prior use of each specific drug using all available data, patients in both the TNFi and DMARD exposure groups can therefore have prior exposure to other therapies in the same exposure group. For example, a patient initiating certolizumab cannot have ever had past certolizumab exposure, but could have received another TNFi in the past. Note that given the common progression of patients from NSAIDs to DMARDs and/or biologics, DMARD patients will not necessarily be biologic naïve. DMARD patients must not have had biologic prescriptions/infusions in the preceding six months.

2) The drug initiation date (index date) must be on or after the AS index date so as to identify new users.

5.2.8 Exclusion criteria

AS disease cohort

There are no exclusion criteria for the AS cohort.

AS exposure cohort

For all outcomes except serious infection and NMSC, outcome specific analyses of incident conditions will exclude patients who already had the outcome of interest at any time prior to the start of follow-up.

5.3 Variables

5.3.1 Exposure Definition and Measurement

The AS cohort will be the starting cohort for all Aims. As a subgroup for the Aim II analyses, we will identify AS exposure episodes to define cohorts. AS patients may have zero, 1, or more than 1 AS exposure episodes. AS exposures will encompass four

mutually exclusive exposure categories, of which the first three will be explicitly represented in the analysis: 1) no treatment or prescription NSAIDs alone; 2) non-biologic conventional DMARDs; 3) TNFi biologics; and 4) Other biologics and targeted synthetic DMARDs (e.g. apremilast). Group 4 will be examined but will be used only for censoring and will not comprise a separate exposure category. The specific drugs exposures are in Table 1.

The TNFi biologic drugs under study are adalimumab, etanercept, certolizumab, golimumab, and infliximab. The definition of new biologic users is patients with no use of that specific treatment identified using all available data prior to the drug initiation date (the date of first dispensing). Similar definitions will be used for new non-biologic DMARD (nbDMARD) users.

Drug exposures will follow a hierarchy of no treatment/NSAIDs, non-biologic DMARDs, and biologics. For example, if a patient initiates methotrexate, then adds adalimumab, then changes to etanercept, and then secukinumab, they would contribute three treatment episodes, one for each of: MTX, adalimumab, and etanercept, where secukinumab would be a censoring event. The adalimumab and etanercept treatment episodes would contribute two observations to the TNFi biologic exposure group, and the time axis and propensity scores, as well as baseline characteristics, are updated at each of the new drug start times.

Table 1. Examples of Common Drugs used for Ankylosing Spondylitis and their Hierarchical Classification

NSAIDs (prescription only)	**nbDMARDs (non-biologics)**	**Biologics and targeted synthetic DMARDs**
Level I (lowest)	**Level II**	**Level III (highest)**
Celecoxib (Celebrex) Ibuprofen Naproxen (Aleve) Meloxicam Indomethacin Diclofenac (Voltaren) Ketorolac (Toradol) Ketoprofen Etodolac Salsalate Flurbiprofen	Hydroxycholoroquine (plaquenil) Leflunomide Methotrexate Sulfasalazine	Adalimumab(Humira) Certolizumab (Cimzia) Etanercept (Enbrel) Golimumab (Simponi) Infliximab (Remicade) ***The biologics and targeted synthetic DMARDs below are not part of this exposure category, but their dispensing represents a censoring event for all Aim II analyses***. *Abatacept(Orencia)* *Anakinra(Kineret)* *Belimumab(Benlysta)* *Canakinumab(Ilaris)* *Ixekizumab (Taltz)* *Rituximab(Rituxan)* *Secukinumab (Cosentyx)* *Tocilizumab(Actemra)* *Ustekinumab (Stelara)* *Vedolizumab(Entyvio)* *Apremilast (Otezla)* *Tofacitinib (Xeljanz)*

We will use pharmacy data (NDC codes) and drug administration codes (HCPCS codes) to identify exposures to NSAIDs, DMARDs, and biologics among patients with AS. Patients will be categorized as new users of the following therapies in the following hierarchy:

1) Neither DMARD nor biologic exposed. This will include patients on no AS treatment, NSAIDs by prescription (observable in the data) or NSAIDs over the counter (not observable in the data);
2) Systemic non-biologic DMARD therapies;
3) Systemic biologic DMARDs, such as anti-TNFs (etanercept, adalimumab, infliximab, certolizumab, golimumab).

For the current study, a person can only be a new user once for each specific drug. However, a person can be a new user for multiple drugs, even if they are within the same drug class (e.g., TNFi therapy). For example, a patient who previously used etanercept, and now is using adalimumab qualifies as a new user of adalimumab and would contribute two exposure episodes to the TNFi group. The unit of analysis for the AS exposure cohort is at the exposure episode level, not the patient level. Each time that a patient starts a new therapy within the same exposure group, baseline covariates will be updated. Treatment episode censoring will occur if a patient initiates a new therapy at the same level or a higher level in the drug exposure hierarchy. Exposure will also be censored if patients start a non-TNFi biologic (e.g., abatacept, secukinumab). Patients will contribute person-time to the 3 exposure groups only if the 6 months prior to the treatment episode start date is uncontaminated by exposure to any treatment higher in the exposure hierarchy. For example, if a patient is on TNFi and then adds a DMARD within 6 months of the most recent TNFi prescription/infusion, this entire DMARD treatment episode will not be counted in any exposure cohort. The rationale for this rule is to correctly classify DMARD exposure that is free of recent or concomitant biologic use. In addition, if someone is on a biologic and stops taking it and initiates no new treatment, exposure will extend by 90 days after discontinuation, and then 6 months later, the patient will begin to contribute person-time to the lowest level exposure group (no treatment/NSAIDs).

5.3.2 Outcome Definitions and Measurement

The 13 outcomes of interest include disease manifestations and comorbidities (8 categories, see Table 2), hospitalized infection, and malignancy (subdivided as hematologic, solid tumors, and NMSC).

Table 2. AS-specific co-morbidities and disease manifestations of interest

Outcome Categories	Specific Manifestation
Cardiac disease	Conduction block*
	Aortic insufficiency/aortic regurgitation*
	Myocardial infarction
Osteoporotic fracture	Clinical vertebral fracture
	Non-vertebral osteoporotic fracture (hip, pelvis, femur, humerus, distal radius/ulna)
Neurological disease	Cauda equina syndrome*
	Spinal cord compression*
Lung disease	Apical pulmonary fibrosis*
	Interstitial lung disease (sensitive definition)
	Restrictive lung disease*
Kidney disease	IgA nephropathy*
	Amyloidosis*
	Nephrotic syndrome*
PSO/PsA	Psoriasis
	Psoriatic arthritis
Inflammatory bowel disease	Ulcerative colitis
	Crohn's disease
Uveitis	Uveitis*

**No known validated algorithm exists in claims data for this outcome; see below*

The thirteen study outcomes of interest include the 8 related to disease manifestations above (Table 2, left-hand column). For the 8 groupings above, the subtypes of each category (Table 2, right-hand column) will be reported descriptively. In addition, we will study 5 more outcomes including infection (hospitalized, opportunistic), and malignancy (solid tumor, hematologic, and NMSC). To the extent possible, we will rely on published algorithms, but where none exists, for prevalent conditions (i.e., comorbidities, and manifestations of the disease itself); we will require 1 physician diagnosis for these conditions. As part of a sensitivity analysis, we will examine the impact of requiring 2

diagnoses for these conditions, and assign the event date to be the date of the second diagnosis.

5.3.3 Covariates

Variables under consideration to characterize AS cohorts and in the comparator population (for Aim 1) and to adjust for in multivariable analyses (for Aim 2), include but are not limited to the following:

- Age (in 5-year age groups)
- Sex
- Median household income, to the limits of precision available, characterized in quartiles
- Geographic location (e.g., country, state), to the limits of geographic precision
- Year of AS cohort entry
- Concurrent and past use (including duration and dosage) of medications Variables that may be indicators of health status or health seeking behavior, such as health resource utilization (office visits, ER visits, hospitalizations).
- Geographic latitude information that may be relevant for the study of some outcomes (e.g., non-melanoma skin cancer), if available.
- History of TNFi exposure
- History of exposure to other biologics
- History of DMARD use

Further details are available as described in the Supplementary Excel Tables for comorbidities and outcomes of interest.

5.4 Data source and data management

5.4.1 Description of database (s)

Three databases will be used for this project: Marketscan, MPCD, and Medicare FFS.

Truven Health MarketScan® Research Databases (2010-2014): These data are a convenience sample of employer sponsored healthcare plans and include the

Commercial Claims and Encounters (Commercial) Database, Medicare Supplemental and Coordination of Benefits (Medicare Supplemental) Database, and the Lab Database. These databases represent healthcare claims information for individuals enrolled in various employer-sponsored healthcare plans. A subset of these individuals also receives benefits via Medicare supplemental insurance (represented as Coordination of Benefits, or COB, information in the database). The Commercial Database represents individuals covered under various commercial plans such as fee-for-service, capitated payment, preferred provider organizations, point-of-service, indemnity, and health maintenance organizations. In 2011, approximately 34.7 million individuals were included in this database. The Medicare Database also represents individuals with Medicare supplemental insurance that is coordinated within the main employer-sponsored plan. In 2011, approximately 4.1 million individuals were included in this database. Laboratory results for a subset of individuals in the Commercial and Medicare Databases are available in the Lab Database. The Lab Database is linked to information in the Commercial and Medicare Databases via unique enrollee identifiers. In 2011, approximately 2.2 million lab tests for 1.4 million unique individuals were available. The Truven comparator cohort is a 10:1 age and sex- matched sample of non-AS patients.

Multi-Payer Claims Database (2007-2010): Patients included in this study were individuals drawn from a proprietary research database (Optum, derived from United Healthcare data) containing eligibility and pharmacy and medical claims data from a large commercial U.S. health plan. The individuals covered by this health plan are geographically diverse across the United States and cover regions in which United Healthcare draws its membership. The data are de-identified for research purposes. An equally sized comparator population of non-AS patients is available.

Medicare Fee-for-service claims data (2006-2014): Patients will be drawn from Fee-for-Service (FFS) Medicare. The available data include Medicare enrollment information, Medicare parts A & B facility claims, Medicare part B carrier file, and Medicare part D prescription event file. The enrollment file contains data on sex, date of birth, race, survival status, managed care participation, and Part A and B eligibility status for each beneficiary at a person-month level. The record is updated monthly

meaning that any change in managed care participation is identified. The Medicare parts A & B claims files contain information on procedures, outpatient physician services, recorded diagnoses, and cost from inpatient claims (Part A), outpatient professional services (Part B). The Medicare part D prescription event file contains detailed information on all outpatient drug dispensing events including the date of service, quantity dispensed, days supplied, cost information, and the identity of the dispensed medication (e.g. formulation, dosage, etc.) using National Drug Codes (NDC).

5.5 STATISTICS

5.5.1 Definition of analysis set

All patients with AS, or in the non-AS comparator cohort (MPCD and Medicare comparator patients only for the prevalence analysis) will contribute to Aim I analyses. The subgroup of patients with AS who contribute one or more AS exposure episodes will be analyzed in Aim II.

5.5.2 Planned analysis

5.5.2.1 Analysis of the primary outcome measures

Aim 1:

- We will calculate the annual prevalence of AS as a point prevalence. The annual prevalence of AS will be calculated as the number of patients who meet the AS definition in each specific year divided by the number of enrollees observable for the full specific year (e.g., the prevalence of AS in 2014 Medicare equals the number of beneficiaries who meet the AS definition in 2014 divided by the number of beneficiaries in the 5% random sample enrolled in Part A, B and not C for the full year 2014 times 20). The number of individuals who meet this definition will be enumerated and standardized by age and sex using the 2010 U.S. Census (2010 Census Summary File 1, Table P12 (https://factfinder.census.gov/bkmk/table/1.0/en/DEC/10_SF1/P12).

- Calculate the prevalence of each comorbidity over various time intervals, by disease cohort (AS vs. non-AS). We will explore how the prevalence of each condition varies according to the amount of data available. Outcomes will be examined in discrete 12 month increments (i.e. using 12 months of data, 24 months, 36 months, etc.) using the 6 months' baseline plus an additional 6 (1st year), 18 (1^{st} and 2^{nd} year), 30 (1^{st}, 2^{nd}, and 3^{rd} year), etc. months of follow-up.

Aim 2:

- Calculate the crude and adjusted incidence rate (and 95% confidence interval) of events of interest by exposure category, stratified by each data source. The results may be pooled across data sources depending on their homogeneity and based upon consensus from the research team. For example, if incidence rate ratios (IRRs) are consistent within each data source, even if the absolute incidence rates vary across data sources, a pooled effect estimate of the IRRs can be provided, allowing the baseline hazard to vary within each data source.
- Depending on the number of absolute events identified in the data, the 13 outcomes of interest will be examined. Patients with prior evidence (using all available data) of each of the 6 disease manifestations, or prior solid tumor and hematologic malignancy, will be excluded from the analysis of these 8 outcomes. There will be no exclusion for prior infection or NMSC.

  Crude and multivariable adjusted hazard ratios for outcomes of interest for each medication class will be calculated and compared. For the no treatment/NSAID, DMARDs, and biologic medication cohorts, propensity scores will be computed with TNFi as the referent group. Stratification or inverse probability treatment weighting (IPTW) will be used to balance exposure groups for aim 2 outcomes.

5.5.2.2 Statistical Analysis for Aim 1

Using the MPCD and Medicare data, we will calculate annual prevalence of AS as point prevalence. The annual prevalence of AS will be calculated as the number of patients who meet the AS definition in the specific year divided by the number of enrollees

observable for the full specific year (e.g., the prevalence of AS in the 2014 Medicare data equals the number of beneficiary that meet the AS case definition in 2014 divided by the number of beneficiaries in the 5% random sample, enrolled in Part A, B and not C for the full year 2014, times 20) The number of individuals meeting this definition will be enumerated and standardized by age and sex using the 2010 U.S. Census (2010 Census Summary File 1, Table P12 (https://factfinder.census.gov/bkmk/table/1.0/en/DEC/10_SF1/P12). Age categories will conform to the categories reported in the Census table.

For the analysis of the prevalence of various AS-related comorbidities, we will compare beneficiaries' characteristics and health care utilization covariates between different data sources among patients who have AS, and within each data source, compared between AS patients and those who do not have AS.

To compare the prevalence of each comorbidity, prevalence per 100 patients with 95% confidence intervals will be calculated using the 6-month baseline period plus 6 months of follow-up, yielding a prevalence estimate for 12 months. The analysis will be repeated using different lengths of follow-up time in 12-month intervals (e.g., 24 months using the 6-month baseline and 18-month follow-up, 36-months using the 6-month baseline and 30-month follow-up, etc.).

5.5.2.3 Statistical Analysis for Aim 2

Incidence rates (IRs) with 95% confidence intervals for each potential outcome will be calculated. Risk ratios, with 95% confidence intervals, will be estimated as hazard ratios (HRs) using Cox regression models and sandwich estimators will be used to examine the associations between patient characteristics and different outcomes. HRs will be calculated from models that are unadjusted and adjusted for all covariates using PS stratification or weighting using IPTW (see below).

Cox-proportional hazard regression models will be used to calculate hazard ratios between PS-trimmed exposure groups of interest while controlling for potentially confounding covariates. For comparison of events between drug exposure groups, Cox-proportional hazard regression models will be adjusted for baseline prednisone use and

propensity score (probability of initiating a biologic agent in contrast to a comparison regimen) grouped into quintiles. Analyses will be performed using SAS (version 9.3, SAS Institute, Cary, NC, USA).

Although one person can only contribute one observation for each specific drug, our study design allows patients to be in different exposure cohorts; thus, Huber Sandwich Estimators will be used to control for the clustered nature of the data.

5.5.2.4 Propensity Score Methods to Balance AS Exposure Groups

Propensity score analysis will be used to balance comorbidities, demographics and other factors between exposure groups (see Supplementary Tables for covariates). For examination of outcomes in relation to disease-modifying drug exposures, propensity scores (PS) estimating the probability of initiating a treatment in contrast to a comparison regimen will be computed. The scores will be divided into quintiles (i.e., stratification) and graphed for every pairwise contrast. Treatment episodes occurring in the non-overlapping tails of the propensity score distribution will be removed ('trimmed') prior to multivariable modeling given the non-comparability of exposed patients to comparator patients. Depending on the number of outcomes and the within-strata balance, IPTW weights will be used, with weights at the 1% and 99% truncated to avoid observations with undue influence. The rationale for preferring IPTW is that this method is expected to preserve the maximal amount of the data. Moreover, stratification with 3 PS groups is potentially difficult to understand since the strata for the 3 pairwise comparisons will not necessarily be comparable across all 3 exposure groups.

5.5.3 Handling of Missing Data

Since beneficiaries are required to be observable in all three data sets during the study period, we do not have known missing data. We do recognize that claims databases capture only certain types of data and events, and thus, misclassification and residual confounding is possible. Claims data do not capture information on certain disease features (e.g., disease severity, disease duration); there is minimally missing-ness of

the data captured. We do not anticipate that loss due to missing or incomplete data will reduce the size of the patients in this study enough to impact study results.

5.5.4 Sample size

The sample size available in the data is fixed and thus formal power calculations are not included for this analysis. An estimate of the number of patients with AS in the Truven MarketScan, MPSD, and Medicare data sources identified approximately 22,000, 5,000, and 35,000 patients, respectively, available for study. The numbers in the Truven MarketScan and Medicare data are likely large enough to estimate the risk of the outcomes of interest in Aim II.

5.5.5 Strength and Limitations

5.5.5.1 Measurement Error(s)/Misclassification(s)

This study is an analysis of automated medical and prescription claims. While claims data are extremely valuable for the efficient and effective examination of health care outcomes, treatment patterns, health care resource utilization, and costs, all claims databases have certain inherent limitations because the claims collected are for the purpose of payment and not research.

For example, Identification of AS using an ICD-9 diagnosis claim does not ensure the presence of disease. Claims submitted are to obtain reimbursement, not for research purposes. Misclassification of an AS diagnosis using a single claim could allow inclusion of false positives cases and overestimate the prevalence of AS.

In the case of drug administrations identified in the pharmacy records, presence of a claim for a filled prescription does not indicate that the medication was consumed or that it was taken as prescribed. Medications filled over-the-counter (e.g., NSAIDs) or provided as samples by the physician, will not be observed in the claims data. Presence of a diagnosis code on a medical claim may not represent true

presence of a disease, as the diagnosis code may be incorrectly coded or included as a rule-out criteria rather than actual disease.

5.5.5.2 Information Bias

Information bias is a flaw in measuring exposure(s), outcome(s), or covariate(s) that results in differences in the quality or accuracy of information between comparison groups. While there is reason to believe that some outcomes are under reported, there is no evidence that they would be differentially reported in the 3 data sources. Covariates in these analyses are age, sex, and year-of-cohort entry, measured at baseline. There is also no evidence that these simple covariates would be differentially measured in the exposure and comparison cohorts. Likewise, measurement of the diagnosis requires all AS cohorts to be determined with the same algorithm.

5.5.5.3 Selection Bias

In this cohort study, all patients diagnosed with AS are included and the planned analysis allows for censoring. There is the possibility that AS itself could differentially affect disenrollment from the insurer (e.g. loss of employment) for AS patients, patients with more severe AS.

5.5.5.4 Confounding

In non-interventional studies, there is always the possibility of residual confounding. Although many characteristics are likely to differ between the AS exposure cohorts, the only factors that can confound the estimates are those that are associated with both medications to treat AS and the outcomes of interest.

5.5.5.5 External Validity of Study Design

While these data may be generalizable to the commercially insured population and Medicare population, they may not be representative of those whose primary insurance

is through Medicaid. AS patients will be identified based on the presence of claims with diagnosis codes for AS. This may not reflect the actual AS population, especially if a patient's AS is in remission during the study period.

5.5.5.6 Analysis Limitations

Although robust, when not all of the assumptions of a Cox model are met, it is possible that subsequent analyses and risk estimates will be biased. Assumptions regarding the time-independence of the hazard ratio may not be correct, and the proportional hazard assumption will be violated, It is possible that the impact of AS or confounders may vary over time, and the proportional hazard assumption will be tested using the method of Ling, Wei and Ying (1993). If the proportional hazard assumption was violated, a piece-wise hazard ratio will be calculated.

6 PROTECTION OF HUMAN SUBJECTS

6.1 Patient consent for data usage and processing

As all of the data is already collected for the primary purpose of billing and administrative functions, no explicit patient consent is required.

6.1.1 Non-Identifiable patient data

The Marketscan and MPCD databases are Health Insurance Portability and Accountability Act (HIPAA) compliant, and all patient data were de-identified. The Medicare data are considered research-identifiable and are governed by a Data Use Agreement (DUA) by CMS.

6.1.2 Patient identification

Individual patients should not be identifiable in any database used for this project.

7 TERMINATION OF THE STUDY

N/A

8 GOOD PHARMACOEPIDEMIOLOGY PRACTICES

This database study follows the Guidelines for Good Epidemiologic Practice (GEP) practices laid out in 2005 FDA GPP and the 2008 International Society of Pharmacoepidemiology (ISPE) GPP.

9 AUDIT AND INSPECTION

The study owner will permit audits mandated by UCB and by the database owner after reasonable notice.

10 ETHICS AND REGULATORY REQUIREMENTS

10.1 Institutional Review Boards and Independent Ethics Committees

Use of the data and the analysis are governed by IRB approval at each participating university.

11 LISTING OF APPENDIX TABLES PROVIDED AS SEPARATE RESOURCES

Outcome Codebook: ICD9 & HCPCS codes & algorithms for diseases, study outcomes (cross-reference with Table 2), see in Excel file “AS Project Cohort Outcome Codebook-20170409.xlsx”

Covariate Listing: ICD9 and HCPCS codes for covariates of interest (cross reference with Table 2), see in Excel file “AS Project Comorbidity Codebook-20170228.xlsx” & Word file “ALGORITHMS TO ENHANCE SPECIFICITY OF FRACTURE IDENTIFICATION_ 100316.docx”

DMARDS & Biologics, and Oral NSAIDs: NDC codes for medications of interest (cross-reference with Table 1), see in Excel file “AS Project Medicine – DMARDs & BIOs - 20170409.xlsx” & “AS Project Medicine - Oral NSAIDs – 20170406.xlsx”

12 PROTOCOL AMENDMENT

Protocol changes (e.g., changing of definition of AS or clean period for differing new users) may affect the statistical evaluations of sample size and the likelihood of the study fulfilling its primary objective.

Significant changes to the protocol will only be made as an amendment to the protocol and must be approved by UCB prior to being implemented. The format will be as follows:

Rationale for the amendment

Modifications and changes

Major changes to the protocol included the following:

<Complete list as appropriate>

Change #<number>

Heading original text

<Original text >

Has been changed to:

<New text>

13 REFERENCES

1. Bremander A, Petersson IF, Bergman S, Englund M**.** Population-based estimates of common comorbidities and cardiovascular disease in ankylosing spondylitis. Arthritis Care Res (Hoboken). 2011;63(4):550-6.

2. Prati C, Claudepierre P, Pham T, Wendling D**.** Mortality in spondylarthritis. Joint Bone Spine. 2011;78(5):466-70.

3. Gladman DD**.** Mortality in psoriatic arthritis. Clin Exp Rheumatol. 2008;26(5 Suppl 51):S62-5.

4. Peters MJ, van der Horst-Bruinsma IE, Dijkmans BA, Nurmohamed MT**.** Cardiovascular risk profile of patients with spondylarthropathies, particularly ankylosing spondylitis and psoriatic arthritis. Semin Arthritis Rheum. 2004;34(3):585-92.

5. Szabo SM, Levy AR, Rao SR, Kirbach SE, Lacaille D, Cifaldi M, et al. Increased risk of cardiovascular and cerebrovascular diseases in individuals with ankylosing spondylitis: a population-based study. Arthritis Rheum. 2011;63(11):3294-304.

6. Peters MJ, van Eijk IC, Smulders YM, Serne E, Dijkmans BA, van der Horst-Bruinsma IE, et al. Signs of accelerated preclinical atherosclerosis in patients with ankylosing spondylitis. J Rheumatol. 2010;37(1):161-6.

7. Sari I, Okan T, Akar S, Cece H, Altay C, Secil M, et al. Impaired endothelial function in patients with ankylosing spondylitis. Rheumatology (Oxford). 2006;45(3):283-6.

8. Mathieu S, Gossec, Laure, Dougados, Maxime, Soubrier, Martin**.** Myocardial Infarction and Cardiovascular Risk Profile in Ankylosing Spondylitis. A Systematic Review and Meta-Analysis. American College of Rheumatology. Atlanta: Arthritis Rheum; 2010:542.

9. Lehtinen K**.** Mortality and causes of death in 398 patients admitted to hospital with ankylosing spondylitis. Ann Rheum Dis. 1993;52(3):174-6.

10. O'Neill TW, Bresnihan B**.** The heart in ankylosing spondylitis. Ann Rheum Dis. 1992;51(6):705-6.

11. O'Neill TW, King G, Graham IM, Molony J, Bresnihan B**.** Echocardiographic abnormalities in ankylosing spondylitis. Ann Rheum Dis. 1992;51(5):652-4.

12. Mathieu S, Gossec L, Dougados M, Soubrier M**.** Cardiovascular profile in ankylosing spondylitis: a systematic review and meta-analysis. Arthritis Care Res (Hoboken). 2011;63(4):557-63.

13. Briot K, Garnero P, Le Henanff A, Dougados M, Roux C**.** Body weight, body composition, and bone turnover changes in patients with spondyloarthropathy receiving anti-tumour necrosis factor {alpha} treatment. Ann Rheum Dis. 2005;64(8):1137-40.

14. Westerveld LA, Verlaan JJ, Oner FC**.** Spinal fractures in patients with ankylosing spinal disorders: a systematic review of the literature on treatment, neurological status and complications. Eur Spine J. 2009;18(2):145-56.

15. Verlaan JJ, Diekerhof CH, Buskens E, van der Tweel I, Verbout AJ, Dhert WJ, et al. Surgical treatment of traumatic fractures of the thoracic and lumbar spine: a systematic review of the literature on techniques, complications, and outcome. Spine (Phila Pa 1976). 2004;29(7):803-14.

16. Bartleson JD, Cohen MD, Harrington TM, Goldstein NP, Ginsburg WW**.** Cauda equina syndrome secondary to long-standing ankylosing spondylitis. Ann Neurol. 1983;14(6):662-9.

17. Jessamine AG**.** Upper lung lobe fibrosis in ankylosing spondylitis. Can Med Assoc J. 1968;98(1):25-9.

18. Kanathur N, Lee-Chiong T**.** Pulmonary manifestations of ankylosing spondylitis. Clin Chest Med. 2010;31(3):547-54.

19. Strobel ES, Fritschka E**.** Renal diseases in ankylosing spondylitis: review of the literature illustrated by case reports. Clin Rheumatol. 1998;17(6):524-30.

20. van Sijl AM, van Eijk IC, Peters MJ, Serne EH, van der Horst-Bruinsma IE, Smulders YM, et al. Tumour necrosis factor blocking agents and progression of subclinical atherosclerosis in patients with ankylosing spondylitis. Ann Rheum Dis. 2015;74(1):119-23.

21. Lin, D., Wei, L. J., and Ying, Z. (1993), “Checking the Cox Model with Cumulative Sums of Martingale-Based Residuals,” *Biometrika*, 80, 557–572.