TITLE PAGE

PASS information

Title	Hydroxychloroquine safety and potential efficacy as an antiviral prophylaxis in light of potential wide-spread use in COVID-19: a multinational, large-scale network cohort and self-controlled case series study
Protocol version identifier	1.4
Date of last version of protocol	01/04/2020
EU PAS register number	To be completed after protocol finalization
Active substance	Sulfasalazine, Hydroxychloroquine, Azithromycin, Amoxicillin
Medicinal product	Sulfasalazine, Hydroxychloroquine, Azithromycin, Amoxicillin
Research question and objectives	The overarching objective is to evaluate the comparative safety of hydroxychloroquine. Secondly, we will evaluate the association between hydroxychloroquine and the risk of flu and hospitalised pneumonia
Country(-ies) of study	To be confirmed. For now: United Kingdom, Germany, Spain, Netherlands, Japan, and the United States of America
Authors	Daniel Prieto-Alhambra
	Jennifer Lane
	Daniel Morales
	Martijn Schuemie
	James Weaver
	Kristina Fišter
	Marc Suchard
	Sajan Khosla
	Ajit Londhe
	Danielle Newby
	Haini Wen
	Paula Casajust
	Andrea Margulis

1. TABLE OF CONTENTS

The study protocol should include a table of contents. The following table of contents can be used if this guidance serves as a template (select the table of content and press "F9" to update the page numbers).

2. LIST OF ABBREVIATIONS

RA	Rheumatoid Arthritis
DMARD	Disease Modifying Anti-rheumatic drug
csDMARD	Conventional synthetic DMARD
SSZ	Sulfasalazine
HCQ	Hydroxychloroquine
SCCS	Self-controlled case series
SAE	Serious adverse events
AZM	Azithromycin
AMX	Amoxicillin
ARDS	Acute Respiratory Distress Syndrome

3. RESPONSIBLE PARTIES

Responsible party/person	**Institution/Affiliation**
Daniel Prieto-Alhambra*	Pharmaco- and Device Epidemiology, CSM-NDORMS, University of Oxford, Oxford, UK
Jennifer Lane	NDORMS, University of Oxford, Oxford, UK
Martijn Schuemie	Observational Health Data Analytics, Janssen Research and Development, and Department of Biostatistics, UCLA
James Weaver	Observational Health Data Analytics, Janssen Research and Development
Patrick Ryan	Observational Health Data Analytics, Janssen Research and Development, and Department of BioInformatics, Columbia University

* Principal investigator

5. AMENDMENTS AND UPDATES

Number	Date	Section of study protocol	Amendment or update	Reason
None				

7 RATIONALE AND BACKGROUND

Since January 2020 a growing number of infections caused by coronavirus SARS-Cov2 COVID19 has resulted in an unprecedented pressure on healthcare systems worldwide, and a great number of casualties on a global scale. With an approximate 4% mortality based on data from China where the outbreak originated, there is a paucity of data on an international level surrounding the factors associated with disease severity or morbi-mortality.[1] As the number of infected patients continues to increase across the world, useful knowledge can be obtained from previous viral outbreaks including the seasonal flu. Federated access to international data assets mapped to the Observational Medical Outcomes Partnership common data model (OMOP) provides a unique opportunity to make a difference in the current crisis within the Observational Health Data Science and Informatics (OHDSI) community.[2,3]

Several existing medicines have been postulated to be effective against coronavirus, including conventional synthetic disease modifying anti-rheumatic drugs (csDMARDs) usually used as the first line of treatment in autoimmune conditions such as rheumatoid arthritis (RA).[4] Chloroquine and hydroxychloroquine (HCQ) have been proposed as a potential treatment for coronavirus based upon their mechanism of action. Accumulating in the acid vesicles (endosome, golgi vesicles, lysosomes), these agents cause alkalinisation, leading to enzyme dysfunction and preventing endosome mediated viral entry to the cell. [5-8] It is also suggested in vitro that they can prevent glycosylation of virus cell proteins including the ACE2 receptor, inhibiting virus entry and replication.[7,9,10] Chloroquine has been shown in vitro to specifically inhibit SARS-Cov2 COVID-19.[11]

In clinical studies, the addition of HCQ has shown increased early virological response to treatment for chronic hepatitis C, and reduced viral load compared to placebo in patients with HIV infection. [12,13] Treatment with also HCQ saw a lower IL-6 level in HIV patients when compared to Zidovudine treatment.[14] Due to its effects in hepatitis C and HIV, in addition to its known mechanism of action, efforts have focused towards HCQ in the attempt to treat SARS-Cov2 COVID-19. Many trials are currently ongoing using hydroxychloroquine in the treatment of SARS-Cov2 COVID-19.[15-21] Gao *et al*. reported that early results from randomised control trials in China using chloroquine versus placebo have seen reduced disease severity and course without serious adverse effects, although Chen *et al.* have suggested no difference in outcome from conventional treatment.[22,23]

Of those studies that have reported more detailed results, Gautret *et al.* compared HCQ 600mg to those receiving HCQ with azithromycin, suggesting that reduced viral load was seen on day 6 of the study in those undertaking combined treatment.[24] However, a double blind RCT using chloroquine for influenza A and B prevention (including H1N1) in Singapore previously saw no prevention of influenza infection of healthy volunteers when compared to placebo. [25]

All therapies used for RA are associated with both non-serious and serious adverse events (SAEs), although the comparative risk of such events in csDMARDs continues to be conflicting.[26] Patients on methotrexate are frequently counselled regarding an increased risk of infection, however there is little good quality evidence quantifying this risk in the literature with several studies suggesting no increased infection risk with any nonbiologic DMARDs, including HCQ.[27,28] Methotrexate use has also been associated with a type of rare lymphoma, however RA patients with uncontrolled disease are at risk of such cancers regardless of therapy compared to the general population.[29]

In the current climate, a set of studies comparing the relative safety of HCQ is pertinent to address this newfound interest in it as a treatment strategy for SARS-Cov2 COVID-19, especially when combined with other medications. The ability to do this across an international network of real world data also confers increased benefit to the global community through greater potential generalisability of results. In addition, studying the association between HCQ use and the risk of flu and hospitalised pneumonia will inform future trials on the use of HCQ as an anti-COVID19 prophylaxis.

8. RESEARCH QUESTION AND OBJECTIVES

The aim of this study is to firstly assess the safety of hydroxychloroquine (HCQ) used in RA, in addition to the safety of combination treatment of HCQ and azithromycin (AZM). Secondly, to assess the antiviral efficacy of HCQ and preventative potential of HCQ in assessing the incidence of viral respiratory infection and hospitalised pneumonia in patients taking HCQ.

Specifically, the study has the following objectives:

1. To assess the direct and comparative safety of HCQ
2. To assess the comparative safety of HCQ in combination with AZM compared to a combination of HCQ with AMX
3. To assess the comparative risk of the adverse outcome of hospital-treated pneumonia associated with the use of HCQ compared to SSZ, and HCQ with AZM in combination therapy compared to HCQ with AMX combination therapy

9. RESEARCH METHODS

9.1. STUDY DESIGN

Two study designs will be used to undertake the four objectives in a multinational, multi-database network:

1. Self-controlled case series (SCCS) estimating the safety of HCQ (Aim 1)
2. New user cohort study estimating the safety, risk of influenza infections and pneumonia in incident users of HCQ compared to SSZ, and users of the combination of HCQ with AZM versus HCQ with AMX (Aims 1, 2, and 3).

Note that AZM will not be studied in the SCCS analyses because of the expected strong time-varying confounding associated with the infection for which AZM is prescribed.

9.2. SETTING

Participants from at least 4 European countries (United Kingdom, Germany, Spain, Netherlands), the United States of America, and Japan are proposed for inclusion. Additional databases will be analysed using the same analytical packages as they join the distributed data network.

Electronic health records and administrative claims from primary care and secondary care will be utilised.

The study will be conducted using data from a large network of real world data sources previously mapped to the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) in collaboration with the Observational Health Data Sciences and Informatics (OHDSI) and European Health Data and Evidence Network (EHDEN) initiatives.

9.2.1. STUDY PERIOD

The study period, when index events and outcomes of interest can be observed, will start from 01/09/2000 and end at the latest available date for all data sources in 2020.

9.2.2. STUDY POPULATION: INCLUSION/EXCLUSION CRITERIA

Participants will be identified using pre-specified concept sets reviewed by a core team of clinicians, epidemiologists, vocabulary experts, and health data scientists with extensive expertise in the use of the OMOP CDM and the OHDSI tools.

New user exposure cohorts

Exposure cohorts will be defined where treatment initiation is the index event and includes the following criteria:

- History of RA: Have a condition occurrence or observation indicating RA any time before or on the same day as the index event
- Be aged 18 years or over at index event
- Have at least 365 days of continuous observation time prior to index event.

Concept ID	Concept name	Domain
4035611	Seropositive rheumatoid arthritis	Condition
4083556	Seronegative rheumatoid arthritis	Condition
80809	Rheumatoid arthritis	Condition
4102493	Polyneuropathy in rheumatoid arthritis	Condition
4107913	Myopathy due to rheumatoid arthritis	Condition
46273442	Deformity of hand due to rheumatoid arthritis	Condition
4334806	Deformity of foot due to rheumatoid arthritis	Condition
4297650	Cutaneous atrophy due to rheumatoid arthritis	Condition
2107559	Rheumatoid arthritis (RA) disease activity, moderate (RA)	Observation
2107558	Rheumatoid arthritis (RA) disease activity, low (RA)	Observation
2107560	Rheumatoid arthritis (RA) disease activity, high (RA)	Observation
2108721	Patient receiving first-time biologic disease modifying anti-rheumatic drug therapy for rheumatoid arthritis (RA)	Observation
4058299	H/O: rheumatoid arthritis	Observation

2107561	Disease prognosis for rheumatoid arthritis assessed, poor prognosis documented (RA)	Observation
2107572	Disease prognosis for rheumatoid arthritis assessed, good prognosis documented (RA)	Observation

Hydroxychloroquine study

HCQ cohort (Target)

Index event is defined as the first recorded dispensing/prescription of HCQ in a patient's history; inferred persistent exposure by allowing up to 90 day gaps between dispensing/prescription records.

SSZ cohort (comparator)

Index event is defined as the first recorded dispensing/prescription of SSZ in a patient's history; inferred persistent exposure by allowing up to 90 day gaps between dispensing/prescription records.

Hydroxychloroquine with Azithromycin study

HCQ and AZM group (target)

Index event is defined as the first recorded dispensing/prescription of either HCQ or AZM in a patient's history with the requirement that a dispensing/prescription for the non-index exposure is observed during the 30 days before and including the index date. Assuming that HCQ use is chronic in the management of RA, and AZM is an acute prescription for infection treatment, inferred persistent exposure to AZM is assessed by allowing up to 30 days between dispensing/prescription records.

HCQ and Amoxicillin group (comparator)

Index event is defined as the first recorded dispensing/prescription of either HCQ or AMX in a patient's history with the requirement that a dispensing/prescription for the non-index exposure is observed during the 30 days before and including the index date. Assuming that HCQ use is chronic in the management of RA, and AMX is an acute prescription for infection treatment, inferred persistent exposure to AMX is assessed by allowing up to 30 days between dispensing/prescription records.

SCCS exposure cohorts

Additional exposure populations, regardless of indication, will be included for the SCCS. For each exposure population, all prevalent users of HCQ will be included and periods of inferred persistent exposure by allowing up to 90 day gaps between dispensing/prescription records will be constructed. Individual SCCS analyses will therefore be executed separately for each of the proposed study outcomes mentioned below (Section 9.3.2), including both safety events and negative control outcomes.

9.2.3. FOLLOW UP

Cohort studies

The index date is defined by the first dispensing/prescription as described in the cohort definitions above (Section 9.2.2.) Two periods of follow-up will be considered for two types of analyses for the serious adverse effect outcomes:
In an *intention-to-treat analysis*, the analysis follow-up starts 1 day after the index date and continues up until the first of: outcome of interest, loss to follow-up, or 30 days after the index date to resemble the likely duration of COVID19 treatment regimens. Patients are required to have at least 1 day of follow-up.

In an *on-treatment analysis*, the analysis follow-up starts 1 day after the index date and continues until the first of: discontinuation/switching/combined therapy of index therapy plus a lag time of 14 days, outcome of interest, loss to follow-up, or 1826 days (5 years) after the index date. Patients are required to have at least 1 day of follow-up.

Self-controlled case series

Self-controlled case series analyses will be conducted in the general case populations in support of to Aim 1. In these analyses, patient cases are followed for their entire observation time (e.g. from enrolment to disenrollment in an insurance claims database), and incidence rates of each of the study outcomes are calculated and compared in periods of inferred persistent exposure to HCQ and non-exposure periods. We will execute a multivariate SCCS analysis for each exposure population that includes time-varying covariates for other drug exposures. We will fit a multivariate SCCS model with HCQ therapy as the time-varying exposure.

9.3. VARIABLES

9.3.1.- EXPOSURES

Two active comparator analyses will be conducted. First, HCQ (target) will be compared to SSZ (comparator). Second, HCQ used in combination with AZM (target) will be compared to HCQ used in combination with AMX (comparator).

Concept IDs for the four ingredients are below:

Target drug group	Concept ID	Concept name
csDMARD	1777087	Hydroxychloroquine
csDMARD	964339	Sulfasalazine
antibiotic	1734104	Azithromycin
antibiotic	1713332	Amoxicillin

Exposure assessment

As described in the cohort definitions (Section 9.2.2), exposure commences on the first dispensing/prescription record with at least 365 days of prior observation period to increase confidence that the exposure is incident. Exposure interval gaps of ≤90 days (HCQ and SSZ) and of ≤30 days (AZM and AMX) between drug dispensing/prescription records will be allowed and inferred as persistent exposure. In the hydroxychloroquine study, drug discontinuation will also be considered if a patient switches from one study drug to another, or when a concomitant second drug is added, with switching defined as an overlap of 30 days or more between two

different drugs. Patients who switch from target exposure to comparator exposure, or vice versa, will contribute follow-up time to the exposure cohort that they entered first.

9.3.2.- OUTCOMES

Outcome identification and validation

The proposed code lists for the identification of the study population (codes for the identification of RA diagnosis) and for the study outcomes were created by clinicians with experience in the management of RA using ATLAS™, and reviewed by 4 clinicians and 1 epidemiologist. Myocardial infarction and stroke codes were based on a previously published paper.[30]

Face validity for the proposed RA and for each of the outcome cohorts will be reviewed by exploring age- and sex-specific incidence rates compared to previous clinical knowledge and/or existing literature.

Safety outcomes (Aim 1) (Preliminary working list at time of protocol registration)

Concept ID	Name / Description	Concept ID to be Excluded
Episodes of acute respiratory distress		
4191650	Acute respiratory distress	
4195694	Acute respiratory distress syndrome	
318459	Respiratory insufficiency	
Persons with chest pain or angina		
321318	Angina pectoris	
77670	Chest pain	
Venous thromboembolic (pulmonary embolism and deep vein thrombosis) events		
435616	Amniotic fluid embolism	X
435887	Antepartum deep vein thrombosis	X
196715	Budd-Chiari syndrome	X
4062269	Cerebral venous thrombosis in pregnancy	X
442055	Obstetric air pulmonary embolism	X
433832	Obstetric blood-clot pulmonary embolism	X
435026	Obstetric pulmonary embolism	X
440477	Obstetric pyemic and septic pulmonary embolism	X
318137	Phlebitis and thrombophlebitis of intracranial sinuses	X
199837	Portal vein thrombosis	X
438820	Postpartum deep phlebothrombosis	X
440417	Pulmonary embolism	
254662	Pulmonary infarction	
4235812	Septic thrombophlebitis	X
195294	Thrombosed hemorrhoids	X
4187790	Thrombosis of retinal vein	X
444247	Venous thrombosis	
Acute renal failure events		
262	Emergency Room and Inpatient Visit	
9203	Emergency Room Visit	
9201	Inpatient Visit	
197320	Acute renal failure syndrome	

432961	Acute renal papillary necrosis with renal failure	
444044	Acute tubular necrosis	
Persons with end stage renal disease		
443611	Chronic kidney disease stage 5	
193782	End stage renal disease	
4090651	Dialysis finding	
4032243	Dialysis procedure	
45889365	Dialysis Services and Procedures	
Persons with hepatic failure		
377604	Hepatic coma	
4029488	Hepatic encephalopathy	
4245975	Hepatic failure	
4337543	Hepatic necrosis	
Acute pancreatitis events		
262	Emergency Room and Inpatient Visit	
9203	Emergency Room Visit	
9201	Inpatient Visit	
199074	Acute pancreatitis	
Persons with heart failure		
315295	Congestive rheumatic heart failure	X
316139	Heart failure	
Total cardiovascular disease events (ischemic stroke, hemorrhagic stroke, heart failure, acute myocardial infarction or sudden cardiac death)		
262	Emergency Room and Inpatient Visit	
9203	Emergency Room Visit	
9201	Inpatient Visit	
4329847	Myocardial infarction	
314666	Old myocardial infarction	X
4048809	Brainstem death	
321042	Cardiac arrest	
442289	Death in less than 24 hours from onset of symptoms	
4317150	Sudden cardiac death	
4132309	Sudden death	
437894	Ventricular fibrillation	X
372924	Cerebral artery occlusion	
375557	Cerebral embolism	
443454	Cerebral infarction	
441874	Cerebral thrombosis	
376713	Cerebral hemorrhage	
439847	Intracranial hemorrhage	
432923	Subarachnoid hemorrhage	
315295	Congestive rheumatic heart failure	X
316139	Heart failure	
Person with cardiac arrhythmia		
44784217	Cardiac arrhythmia	
38001137	Cardiac arrhythmia & conduction disorders w CC	
38001138	Cardiac arrhythmia & conduction disorders w/o CC/MCC	
315078	Palpitations	
444070	Tachycardia	
21600248	ANTIARRHYTHMICS, CLASS I AND III	
43013024	apixaban	
40228152	dabigatran etexilate	

40241331	rivaroxaban	
1310149	Warfarin	
1309204	Adenosine	
45892847	edoxaban	
45890325	Cardioversion, elective, electrical conversion of arrhythmia	
45890400	Operative tissue ablation and reconstruction of atria, extensive (eg, maze procedure)	
2107068	Operative tissue ablation and reconstruction of atria, performed at the time of other cardiac procedure(s), extensive (eg, maze procedure), with cardiopulmonary bypass (List separately in addition to code for primary procedure)	
4051932	Procedure for arrhythmia	
Persons with bradycardia		
4169095	Bradycardia	
316999	Conduction disorder of the heart	
4171683	Sinus bradycardia	
317302	Sinus node dysfunction	
Gastrointestinal bleeding events		
262	Emergency Room and Inpatient Visit	
9203	Emergency Room Visit	
9201	Inpatient Visit	
4138962	Acute duodenal ulcer without hemorrhage AND without perforation	X
4195231	Acute gastric ulcer without hemorrhage AND without perforation	X
4147683	Acute gastrojejunal ulcer without hemorrhage AND without perforation	
4163865	Acute peptic ulcer without hemorrhage AND without perforation	X
195584	Acute peptic ulcer without hemorrhage AND without perforation but with obstruction	X
40482685	Angiodysplasia of duodenum	
28779	Bleeding esophageal varices	
4222896	Chronic duodenal ulcer without hemorrhage AND without perforation	X
4296611	Chronic gastric ulcer without hemorrhage AND without perforation	X
200769	Chronic gastric ulcer without hemorrhage, without perforation AND without obstruction	X
4177387	Chronic gastrojejunal ulcer without hemorrhage AND without perforation	X
434400	Chronic gastrojejunal ulcer without hemorrhage AND without perforation but with obstruction	X
438795	Chronic gastrojejunal ulcer without hemorrhage, without perforation AND without obstruction	X
4204555	Chronic peptic ulcer without hemorrhage AND without perforation	X
24973	Chronic peptic ulcer without hemorrhage AND without perforation but with obstruction	X
23808	Chronic peptic ulcer without hemorrhage, without perforation AND without obstruction	X
2002608	Control of hemorrhage and suture of ulcer of stomach or duodenum	
198798	Dieulafoy's vascular malformation	
4198381	Duodenal ulcer disease	
4209746	Duodenal ulcer without hemorrhage AND without perforation	X
4112183	Esophageal varices with bleeding, associated with another disorder	
2108900	Esophagogastroduodenoscopy, flexible, transoral; with control of bleeding, any method	
2108878	Esophagoscopy, flexible, transoral; with control of bleeding, any method	
4265600	Gastric ulcer	

4248429	Gastric ulcer without hemorrhage AND without perforation	X
192671	Gastrointestinal hemorrhage	
4101104	Gastrojejunal ulcer without hemorrhage AND without perforation	X
443530	Hematochezia	X
197925	Hemorrhage of rectum and anus	X
4027663	Peptic ulcer	
4291028	Peptic ulcer without hemorrhage AND without perforation	X
All-cause mortality		
Cardiovascular-related mortality		
4329847	Myocardial infarction	
314666	Old myocardial infarction	X
376713	Cerebral hemorrhage	
439847	Intracranial hemorrhage	
432923	Subarachnoid hemorrhage	
4048809	Brainstem death	
321042	Cardiac arrest	
442289	Death in less than 24 hours from onset of symptoms	
4317150	Sudden cardiac death	
4132309	Sudden death	
437894	Ventricular fibrillation	X
315295	Congestive rheumatic heart failure	X
316139	Heart failure	
262	Emergency Room and Inpatient Visit	
9203	Emergency Room Visit	
9201	Inpatient Visit	
372924	Cerebral artery occlusion	
375557	Cerebral embolism	
443454	Cerebral infarction	
441874	Cerebral thrombosis	
Transient ischemic attack events		
262	Emergency Room and Inpatient Visit	
9203	Emergency Room Visit	
9201	Inpatient Visit	
373503	Transient cerebral ischemia	
Stroke (ischemic or hemorrhagic) events		
262	Emergency Room and Inpatient Visit	
9203	Emergency Room Visit	
9201	Inpatient Visit	
372924	Cerebral artery occlusion	
375557	Cerebral embolism	
376713	Cerebral hemorrhage	
443454	Cerebral infarction	
441874	Cerebral thrombosis	
439847	Intracranial hemorrhage	
432923	Subarachnoid hemorrhage	
Acute myocardial infarction events		
262	Emergency Room and Inpatient Visit	
9203	Emergency Room Visit	
9201	Inpatient Visit	
4329847	Myocardial infarction	
314666	Old myocardial infarction	X

Influenza infection (aim 2)

Episodes of influenza		
4266367	Influenza	
4319159	Influenza-like illness	
4153160	Influenza-like symptoms	
4099911	Parainfluenza virus bronchitis	
439857	Parainfluenza virus pneumonia	
Episodes of influenza symptoms or complications		
4041664	Difficulty breathing	
4266367	Influenza	
4319159	Influenza-like illness	
4153160	Influenza-like symptoms	
4099911	Parainfluenza virus bronchitis	
439857	Parainfluenza virus pneumonia	
444413	Febrile convulsion	
437663	Fever	
254761	Cough	
4089228	Sputum finding	
255848	Pneumonia	
442774	Intermittent claudication	X
442752	Muscle pain	

Hospitalisation for pneumonia (Aim 3)

Concept ID	Name / Description	Concept id to be Excluded
Hospitalizations with pneumonia		
255848	Pneumonia	
262	Emergency Room and Inpatient Visit	
9201	Inpatient Visit	

Negative control outcomes

A list of negative control outcomes will also be assessed for which there is no causal relationship with choice of medication after a diagnosis of rheumatoid arthritis. These outcomes were identified using a semi-automatic process based on data extracted from literature, product labels and spontaneous reports followed by manual review by 2 clinicians.[31] The list is available in Annex 2.

9.3.3.- Covariates

Cohort studies

The following consistently extracted set of baseline patient characteristics will be constructed for inclusion as potentially confounding covariates in the regularized, logistic regression PS model.[32] From this large set of typically tens of thousands of covariates, key predictors of exposure classification will be selected for the propensity score (See Section 9.7.). Note that not all data sources necessarily include data for all covariates. Covariates to be included:

- Demographics (age in 5-year bands, sex, race, ethnicity, index year, index month)

- All conditions occurrence records aggregated to SNOMED clinical finding level during the following lookback windows:
    - in 365 days prior to and including index date
    - in 30 days prior to and including index date
- All drug exposure records aggregated to RxNorm ingredient level and ATC classes during the following lookback windows:
    - in 365 days prior to and including index date
    - in 30 days prior to and including index date
    - persistent exposure that overlaps index date
- All procedure occurrence records during the following lookback windows:
    - in 365 days prior to and including index date
    - in 30 days prior to and including index date
- Measurements (including laboratories) within, above, and below normal range during the following lookback window:
    - in 365 days prior to and including index date
- Device exposure records during the following lookback windows:
    - in 365 days prior to and including index date
    - in 30 days prior to and including index date
- Comorbidity or risk scores including:
    - Charlson
    - DCSI
    - CHADS2
    - CHADS2VASc

SCCS

The effects of age and season will be assumed constant within each calendar month, and will each be modelled using bicubic splines with 5 knots. For more detail, see the SelfControlledCaseSeries package. (http://ohdsi.github.io/SelfControlledCaseSeries)

9.4. DATA SOURCES

This study will aim to be conducted using routinely collected data from different data sources that participate in the OHDSI and/or EHDEN initiatives.

These databases will provide representative clinical information as collected in actual practice conditions in different European healthcare settings, US, Australia and Japan routine practice. Further databases will be added as they are made available to this initiative.

The proposed databases have been proposed based upon their participation in the OHDSI and EHDEN initiatives after mapping to the OMOP common data model. Where possible, data will be accessed remotely by participants from data partner institutions in EHDEN (SIDIAP, IPCI, CPRD), from study investigators from Janssen Pharmaceuticals (CCAE, Optum, MDCR, MDCD, JMDC, PanTher) and from IQVIA (THIN UK, IQVIA US Ambulatory EMR, IQVIA Australia EMR, LPD Belgium EMR, IQVIA Disease Analyser France EMR, IQVIA Disease Analyser Germany EMR, and IQVIA Hospital US Charge Master).

Data available to Janssen have been described elsewhere and include US claims and EMR, and Japanese claims.[30] Other participating databases are detailed in the table below, and include electronic medical records and claims from Europe and the US.

All analyses will be conducted in a federated manner using tools previously validated and tested in a number of studies conducted by the OHDSI community.

Table 9.4: Overview of the considered databases

Source Name	Patient Count	History	Patient Type	Data collection
IQVIA US Ambulatory EMR	49m	2006 –	Outpatient / General population	Dataset consists of longitudinal, de-identified ambulatory EHR data
IQVIA Australia EMR	6m	2006 –	Outpatient / General population	Anonymized patient records collected from Patient Management software used by GPs during an office visit to document patients' clinical records
LPD Belgium EMR	2m	2005 –	Outpatient / General population	Medical contacts and diagnoses, test results and drugs associated with them. Only outpatient
IQVIA Disease Analyser France EMR	10m	1997 –	Outpatient / General population Patients seen in the primary care setting	Anonymized patient records collected from Patient Management software used by GPs during an office visit to document patients' clinical records
IQVIA Disease Analyser Germany EMR	37m	1992 –	Outpatient / General population Public and private insurance	Anonymized patient records collected from Patient Management software used by GPs and selected specialists to document patients' medical records within their office-based practice during a visit
IQVIA Hospital US Charge Master	86m	2007 –	Inpatient & outpatient hospital encounters, including Emergency Room visits / General population	Anonymized patient level data are sourced from hospital charge detail masters (CDM) and collected from resource management software within short-term, acute-care and non-federal hospitals
IQVIA UK THIN IMRD EMR	15m	1989 –	General population / Primary care records with hospitalization / referral information	Pseudonymized Electronic Medical Records collected from Patient Management software used within UK Primary Care
Estonian Health Information System	1.4m	2012-2016	All inpatient and outpatient discharge summaries, general population	Pseudonymized patient level health records from central e-health database where submitting the records is mandatory for all healthcare service providers in Estonia
Integrated Primary Care Information	2.5m	1996	Patients seen in Primary Care setting	The Integrated Primary Care Information (IPCI) database is a Dutch database containing the complete medical record of more than 2.5 million patients provided by more than 450 GPs geographically spread over the Netherlands. (12) In the Netherlands, all citizens are registered with a GP practice which acts as a gatekeeper in a two-way exchange of information with secondary care.

SIDIAP	6m	2006	Primary care linked (partially) to inpatient data	The Sistema d'Informacio per al Desenvolupament de l'Investigacio en Atencio Primaria (SIDIAP) is a primary care records database that covers >80% of the population of Catalonia, North-East Spain. Healthcare is universal and tax-payer funded in the region, and primary care physicians are gatekeepers for all care and responsible for repeat prescriptions.

9.5. STUDY SIZE

Since this study will be undertaken using routinely collected data, all patients meeting the eligibility criteria above will be included. No *a priori* sample calculation was performed; instead, a minimum detectable relative risk (MDRR) will be calculated for each target-comparator-outcome analysis in each of the available databases. Analyses are required to have >0 events observed during follow-up in both target or comparator cohorts in order to produce an estimate and standard error. Given at least 1 event is observed, a large MDRR in a single data source could contribute an underpowered estimate to a meta-analytic estimate provided adequate study diagnostics criteria are met (See Section 9.7.)

9.6. DATA MANAGEMENT

All data extraction and curation will be conducted using the ATLAS tool, an open source software platform by the OHDSI community, as well as the OHDSI Methods Library, a set of R packages developed and maintained by the OHDSI community.
(https://ohdsi.github.io/MethodsLibrary)

The process will follow the steps described here:

1. Define concept set expressions that consist of the source codes used to record clinical observations in disparate data sources
2. Define the target and comparator exposure cohorts used as input to subsequent analytic routines
3. Ascertain outcome populations
4. Review of cohort diagnostics for study feasibility and clinical face validity (e.g. cohort sizes, age and sex-specific incidence rates, index event source code prevalence, clinical characteristics)

9.7. DATA ANALYSIS

SAFETY ANALYSES (AIMS 1 AND 2)

Comparative Cohort analysis in RA

The comparative safety of HCQ study in subjects with RA will be assessed through a comparative cohort analysis, compared against SSZ as an active comparator. The comparative safety of HCQ + AZM combination therapy will be compared against HCQ + AMX therapy in subjects with RA. Individuals with a history of the outcome occurring within the 30 days before index will be excluded from the analyses.

Analyses will use the CohortMethod package (https://ohdsi.github.io/CohortMethod/). This analytic package uses a large-scale propensity score constructed through the Cyclops package (https://ohdsi.github.io/Cyclops), based on many baseline covariates derived from the data, including all drugs, condition, and procedures observed prior to the treatment initiation, as well as summary scores such as the Charlson Comorbidity Index.[33]

Propensity score (PS) stratification will be used as the analytic strategy to adjust for imbalance between exposure cohorts in a comparison. The PS will be estimated using a large-scale regularized logistic regression fitted with a Laplace prior (LASSO) and with the optimal hyperparameter determined through 10-fold cross validation. The predictor variables included will be based on all observed patient characteristics and covariates available at each data source and extracted as described above (See Section 9.3.3.). We will exclude all covariates that occur in fewer than 0.1% of patients within the target and comparator cohorts prior to PS model fitting for computational efficiency. Patients in the target and comparator cohorts will be stratified into 5 PS quintiles. We will compute and plot the propensity score distribution and assess covariate balance expressed as the standardized difference of the mean for every covariate before and after propensity score adjustment. We will consider any standardized difference > 0.1 to indicate non-negligible imbalance between exposure cohorts.[34]

We compare the target cohort with the comparator cohort for the hazards of outcome during the follow-up periods by applying a univariate Cox proportional hazards model conditioned on the PS strata with treatment allocation as the sole explanatory variable.

SCCS

Safety of HCQ therapy will also be assessed separately, regardless of indication, using SCCS analysis, comparing exposed and unexposed time periods within the same individuals. The method is self-controlled in that it makes within-person comparisons of event rates during periods of hypothesized increase risk with other periods of baseline risk, with eliminates all time-invariant confounding. Because we do not compare between persons, the SCCS is robust to between-person differences, even including unmeasured differences (like genetics). However, the method is vulnerable to time-varying confounders: the time of exposure may be incomparable to the time when not exposed. To adjust for this, we will include many time-varying co-variates in the models, including age, season, and other drug exposures (and by proxy underlying conditions treated by drug exposures).

The effect estimate is the IRR, which is derived from a Poisson regression that is conditioned on an outcome event having been observed. Hence, it is a case-only design where the population for which the inference is made is the intersection of the exposed population and case population. The method is considered a series because it uses longitudinal time-stamped exposure and event information between fixed study start and end points. The validity of the SCCS effect measure relies on the main assumptions that events do not influence subsequent exposures, events do not influence the length of a patient's observation period with no time trend in outcome, and events arise in a non-homogenous Poisson process.[35,36] The event rate is modeled as

$$\lambda_{ij} = \exp(\phi_i + \beta_j)$$

where λ_{ij} is the event rate for the ith person with the jth level exposure, ϕ_i is the individual effect for the ith person, and β_j is the jth level exposure effect that acts multiplicatively relative

to β_0=0. The case series likelihood function used to estimate β_j takes a ratio form in which the individual effect ϕ_i cancels out as a nuisance parameter, rendering the effect β_j implicitly controlled for individual effects. The estimate log(β_j) is the IRR and is interpreted as the effect of the jth level of the exposure on the outcome relative to the baseline risk.[37] We will adjust for age and season to reduce the risk of underlying confounding. In a separate sensitivity analysis, we will additionally adjust for time-dependent observation.[38]

The SCCS is run using the freely available package (https://ohdsi.github.io/SelfControlledCaseSeries/). This package allows the modelling of age and season as splines, and supports the multiple SCCS model (MSCCS), where all other drug exposures are included as time-varying covariates.[39] A large-scale regularized conditional Poisson regression fits the outcome model using the Cyclops package, with a hyperparameter selected through 10-fold cross-validation.

RISK OF SYSTEMIC VIRAL INFECTIONS AND OF HOSPITAL-TREATED PNEUMONIA (AIM 3)

The comparative and direct risk of systemic viral infections (flu) associated with the use of HCQ and HCQ + AZM will be assessed using both a comparative cohort study (comparison versus SSZ for HCQ, and HCQ + AMX for HCQ + AZM) and a SCCS (within-subject comparisons for each target exposure population separately), with the same methodology described for Aim 1.

The comparative and direct risk of hospital-treated pneumonia associated with the use of HCQ and HCQ + AZM will be assessed using both a comparative cohort study (comparison versus SSZ for HCQ, and HCQ + AMX for HCQ + AZM) and a SCCS (within-subject comparisons for each target exposure separately) with the same methodology described for Aim 1.

EVIDENCE EVALUATION

In addition to the design-specific diagnostics, such as the covariate balance computed for the comparator cohort design, we will estimate overall residual bias in all designs using negative controls. An assessment of negative control outcomes (Annex 3) will be used to assess whether there is residual confounding after propensity score adjustment. An empirical null distribution will be fitted to the effect size estimates of the negative controls, allowing for quantification of residual bias and calibration of hazard ratios, confidence intervals, and p-values. If there is evidence of residual confounding and there is a sufficient number of control events, estimates will be calibrated.

Study diagnostics (power, propensity score distribution, covariate balance, empirical null distribution) will be evaluated by clinicians and epidemiologists to determine which database-target-comparator-outcome-analysis variants will produce unbiased estimates. Database-target-comparator-analysis variants with 0 outcomes in the time-at-risk window or contained analyses with baseline covariate with standardized mean difference>0.1 after stratification will be excluded from analysis. Study diagnostics for all database-target-comparator-outcome-analysis will be provided as part of study, regardless of which effect estimation results are unblinded. The main models will be adjusted for unbalanced PS-variables at baseline.

All analysis code will be completed and version controlled at https://github.com/ohdsi-studies/Covid19DrugRepurposing prior to unblinding estimation results. All study diagnostics are available for exploration at https://data.ohdsi.org/Covid19EstimationHydroxychloroquine/.

All the proposed analyses will be conducted for each database separately, with estimates combined in fixed effects meta-analysis methods where I2 is <=40%. No meta-analysis will be conducted where I2 for a given drug-outcome pair is >40%.

9.8. LIMITATIONS OF THE RESEARCH METHODS

Selection bias

Selection bias might arise as the consequence of including subjects with a specific period of time available in the data. Attrition tables will be provided to report on the impact of such exclusion criteria.

Information bias

Information bias may occur due to the incorrect identification of exposure, outcomes or co-variates. With regards to exposure, misclassification may occur due to the patient not fulfilling the prescription (primary non-adherence) or in relation with non-compliance. Hence an overestimate of utilization of the study drugs can happen, expectedly leading to non-differential misclassification.

In addition, lack or incomplete recording of safety events may lead to misclassification of the proposed safety endpoints.

Confounding

As confounding by indication may produce differences in baseline characteristics between the comparator and target cohorts, we will use several methods to deal with confounding:

1. Restriction: comparative studies will be conducted only in subjects previously diagnosed with RA and using any of the drugs of interest as a first line treatment. In addition, we will trim the <5% and >95% percentiles of the preference score to maximise equipoise in the study population.
2. Propensity score stratification: we will stratify by PS quintiles to reduce confounding by indication.
3. Matching: for the comparative studies we will use propensity score matching (1:1) to minimise confounding related to all observed confounders.
4. Negative control outcome analyses will be used to identify any residual unobserved confounding in the propensity score analyses. If this analysis suggests the presence of relevant unresolved confounding then further analyses will not be completed.

10. PROTECTION OF HUMAN SUBJECTS

For this study, participants from numerous healthcare databases will be studied. The use of the OMOP common data model and OHDSI tools will enable the federated analysis of these different databases without changing access rights to patient-level data.

11. MANAGEMENT AND REPORTING OF ADVERSE EVENTS/ADVERSE REACTIONS

According to the new guidelines for good pharmacovigilance practice (EMA/873138/2011) there is no requirement for expedited reporting of adverse drug reactions from studies with secondary use of data (such as electronic health care databases). All the identified adverse events/reactions will be summarized in the resulting manuscript/s and/or interactive web-based report of all conducted analyses.

12. Plans for disseminating and communicating study results

Dissemination activities will be of a scientific nature (articles in scientific journals, presentations at conferences, etc.). Our aim is for these studies to be made available as soon as possible in order to support treatment decisions in the global Covid-19 pandemic.

13. References

1. WHO. Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19). Geneva: WHO, 2020.
2. OHDSI. Observational Health Data Sciences and Informatics. 2020. https://www.ohdsi.org [last accessed 24.03.2020].
3. OHDSI. Data standardisation. https://www.ohdsi.org/data-standardization/ [Last accessed 24.03.2020].
4. Smolen JS, Landewe R, Breedveld FC, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2013 update. *Ann Rheum Dis* 2014; **73**(3): 492-509.
5. Colson P, Rolain J-M, Lagier J-C, Brouqui P, Raoult D. Chloroquine and hydroxychloroquine as available weapons to fight COVID-19. *International journal of antimicrobial agents* 2020: 105932-.
6. Savarino A, Boelaert JR, Cassone A, Majori G, Cauda R. Effects of chloroquine on viral infections: An old drug against today's diseases? *Lancet Infectious Diseases* 2003; **3**(11): 722-7.
7. Fedson DS. Confronting an influenza pandemic with inexpensive generic agents: can it be done? *The Lancet Infectious Diseases* 2008; **8**(9): 571-6.
8. Vigerust DJ, Shepherd VL. Virus glycosylation: role in virulence and immune interactions. *Trends in Microbiology* 2007; **15**(5): 211-8.
9. Devaux CA, Rolain J-M, Colson P, Raoult D. New insights on the antiviral effects of chloroquine against coronavirus: what to expect for COVID-19? *International journal of antimicrobial agents* 2020: 105938-.
10. Savarino A, Di Trani L, Donatelli I, Cauda R, Cassone A. New insights into the antiviral effects of chloroquine. *The Lancet Infectious diseases* 2006; **6**(2): 67-9.
11. Wang M, Cao R, Zhang L, et al. Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro. *Cell Res* 2020; **30**(3): 269-71.
12. Helal GK, Gad MA, Abd-Ellah MF, Eid MS. Hydroxychloroquine augments early virological response to pegylated interferon plus ribavirin in genotype-4 chronic hepatitis C patients. *Journal of medical virology* 2016; **88**(12): 2170-8.
13. Sperber K, Louie M, Kraus T, et al. Hydroxychloroquine treatment of patients with human immunodeficiency virus type 1. *Clin Ther* 1995; **17**(4): 622-36.
14. Sperber K, Chiang G, Chen H, et al. Comparison of hydroxychloroquine with zidovudine in asymptomatic patients infected with human immunodeficiency virus type 1. *Clin Ther* 1997; **19**(5): 913-23.
15. ChiCTR2000029542. Study for the efficacy of chloroquine in patients with novel coronavirus pneumonia (COVID-19). 2020.
16. ChiCTR2000029740. Efficacy of therapeutic effects of hydroxycholoroquine in novel coronavirus pneumonia (COVID-19) patients(randomized open-label control clinical trial). 2020.
17. ChiCTR2000029740, The First Hospital of Peking University Y. Efficacy of therapeutic effects of hydroxycholoroquine in novel coronavirus pneumonia (COVID-19) patients(randomized open-label control clinical trial). 2020.
18. ChiCTR2000029559, Renmin Hospital of Wuhan University N. Therapeutic effect of hydroxychloroquine on novel coronavirus pneumonia (COVID-19). 2020.
19. ChiCTR2000029559. Therapeutic effect of hydroxychloroquine on novel coronavirus pneumonia (COVID-19). 2020.

20. ChiCTR2000029741, The Fifth Affiliated Hospital Sun Yat-Sen University Y. Efficacy of Chloroquine and Lopinavir/ Ritonavir in mild/general novel coronavirus (CoVID-19) infections: a prospective, open-label, multicenter randomized controlled clinical study. 2020.
21. ChiCTR2000029609, The Fifth Affiliated Hospital of Sun Yat-Sen University Y. A prospective, open-label, multiple-center study for the efficacy of chloroquine phosphate in patients with novel coronavirus pneumonia (COVID-19). 2020.
22. Gao J, Tian Z, Yang X. Breakthrough: Chloroquine phosphate has shown apparent efficacy in treatment of COVID-19 associated pneumonia in clinical studies. *Biosci Trends* 2020; **14**(1): 72-3.
23. NCT04261517. Efficacy and Safety of Hydroxychloroquine for Treatment of Pneumonia Caused by 2019-nCoV (HC-nCoV). 2020. https://clinicaltrials.gov/show/NCT04261517; http://subject.med.wanfangdata.com.cn/UpLoad/Files/202003/43f8625d4dc74e42bbcf24795de1c77c.pdf.
24. GAUTRET P, LAGIER JC, PAROLA P, et al. Hydroxychloroquine and Azithromycin as a treatment of COVID-19: preliminary results of an open-label non-randomized clinical trial. *medRxiv* 2020: 2020.03.16.20037135.
25. Paton NI, Lee L, Xu Y, et al. Chloroquine for influenza prevention: a randomised, double-blind, placebo controlled trial. *Lancet Infect Dis* 2011; **11**(9): 677-83.
26. Costello R, David T, Jani M. Impact of Adverse Events Associated With Medications in the Treatment and Prevention of Rheumatoid Arthritis. *Clin Ther* 2019; **41**(7): 1376-96.
27. Lacaille D, Guh DP, Abrahamowicz M, Anis AH, Esdaile JM. Use of nonbiologic disease-modifying antirheumatic drugs and risk of infection in patients with rheumatoid arthritis. *Arthritis Rheum* 2008; **59**(8): 1074-81.
28. Salliot C, van der Heijde D. Long-term safety of methotrexate monotherapy in patients with rheumatoid arthritis: a systematic literature research. *Ann Rheum Dis* 2009; **68**(7): 1100-4.
29. Smitten AL, Simon TA, Hochberg MC, Suissa S. A meta-analysis of the incidence of malignancy in adult patients with rheumatoid arthritis. *Arthritis Res Ther* 2008; **10**(2): R45.
30. Suchard MA, Schuemie MJ, Krumholz HM, et al. Comprehensive comparative effectiveness and safety of first-line antihypertensive drug classes: a systematic, multinational, large-scale analysis. *Lancet* 2019; **394**(10211): 1816-26.
31. Voss EA, Boyce RD, Ryan PB, van der Lei J, Rijnbeek PR, Schuemie MJ. Accuracy of an automated knowledge base for identifying drug adverse reactions. *J Biomed Inform* 2017; **66**: 72-81.
32. Tian Y, Schuemie MJ, Suchard MA. Evaluating large-scale propensity score performance through real-world and synthetic data experiments. *Int J Epidemiol* 2018; **47**(6): 2005-14.
33. Suchard MA, Simpson SE, Zorych I, Ryan P, Madigan D. Massive parallelization of serial inference algorithms for a complex generalized linear model. *ACM Trans Model Comput Simul* 2013; **23**(1).
34. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat Med* 2009; **28**(25): 3083-107.
35. Whitaker HJ, Ghebremichael-Weldeselassie Y, Douglas IJ, Smeeth L, Farrington CP. Investigating the assumptions of the self-controlled case series method. *Stat Med* 2018; **37**(4): 643-58.
36. Petersen I, Douglas I, Whitaker H. Self controlled case series methods: an alternative to standard epidemiological study designs. *BMJ* 2016; **354**: i4515.
37. Whitaker HJ, Farrington CP, Spiessens B, Musonda P. Tutorial in biostatistics: the self-controlled case series method. *Stat Med* 2006; **25**(10): 1768-97.
38. Farrington CP, Anaya-Izquierdo K, Whitaker HJ, Hocine MN, Douglas I, Smeeth L. Self-Controlled Case Series Analysis with Event-Dependent Observation Periods. *Journal of the American Statistical Association* 2011; **106**(494): 417-26.
39. Simpson SE, Madigan D, Zorych I, Schuemie MJ, Ryan PB, Suchard MA. Multiple self-controlled case series for large-scale longitudinal observational databases. *Biometrics* 2013; **69**(4): 893-902.

ANNEX 1. ENCEPP CHECKLIST FOR STUDY PROTOCOLS

Study title: Hydroxychloroquine safety and potential efficacy as an antiviral prophylaxis in light of potential wide-spread use in COVID-19: a multinational, large-scale network cohort and self-controlled case series study

EU PAS Register® number:
Study reference number (if applicable):

Section 1: Milestones	**Yes**	**No**	**N/A**	**Section Number**
1.1 Does the protocol specify timelines for				
1.1.1 Start of data collection[1]	☒	☐	☐	9.2.1
1.1.2 End of data collection[2]	☒	☐	☐	
1.1.3 Progress report(s)	☐	☐	☒	
1.1.4 Interim report(s)	☐	☐	☒	
1.1.5 Registration in the EU PAS Register®	☐	☐	☒	
1.1.6 Final report of study results.	☐	☐	☒	

Comments:

Section 2: Research question	**Yes**	**No**	**N/A**	**Section Number**
2.1 Does the formulation of the research question and objectives clearly explain:	☒	☐	☐	7
2.1.1 Why the study is conducted? (e.g. to address an important public health concern, a risk identified in the risk management plan, an emerging safety issue)	☒	☐	☐	
2.1.2 The objective(s) of the study?	☒	☐	☐	8
2.1.3 The target population? (i.e. population or subgroup to whom the study results are intended to be generalised)	☒	☐	☐	9.2.2
2.1.4 Which hypothesis(-es) is (are) to be tested?	☒	☐	☐	8
2.1.5 If applicable, that there is no *a priori* hypothesis?	☐	☐	☒	

Comments:

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

Section 3: Study design		**Yes**	**No**	**N/A**	**Section Number**
3.1	Is the study design described? (e.g. cohort, case-control, cross-sectional, other design)	☒	☐	☐	9.1
3.2	Does the protocol specify whether the study is based on primary, secondary or combined data collection?	☒	☐	☐	9.1
3.3	Does the protocol specify measures of occurrence? (e.g., rate, risk, prevalence)	☒	☐	☐	9.7
3.4	Does the protocol specify measure(s) of association? (e.g. risk, odds ratio, excess risk, rate ratio, hazard ratio, risk/rate difference, number needed to harm (NNH))	☒	☐	☐	9.7
3.5	Does the protocol describe the approach for the collection and reporting of adverse events/adverse reactions? (e.g. adverse events that will not be collected in case of primary data collection)	☐	☐	☒	

Comments:

Section 4: Source and study populations		**Yes**	**No**	**N/A**	**Section Number**
4.1	Is the source population described?	☒	☐	☐	9.2.2
4.2	Is the planned study population defined in terms of:				
	4.2.1 Study time period	☒	☐	☐	9.2
	4.2.2 Age and sex	☒	☐	☐	
	4.2.3 Country of origin	☒	☐	☐	
	4.2.4 Disease/indication	☒	☐	☐	
	4.2.5 Duration of follow-up	☒	☐	☐	
4.3	Does the protocol define how the study population will be sampled from the source population? (e.g. event or inclusion/exclusion criteria)	☒	☐	☐	

Comments:

Section 5: Exposure definition and measurement		**Yes**	**No**	**N/A**	**Section Number**
5.1	Does the protocol describe how the study exposure is defined and measured? (e.g. operational details for defining and categorising exposure, measurement of dose and duration of drug exposure)	☒	☐	☐	9.2.2 & 9.3.1

Section 5: Exposure definition and measurement		**Yes**	**No**	**N/A**	**Section Number**
5.2	Does the protocol address the validity of the exposure measurement? (e.g. precision, accuracy, use of validation sub-study)	☒	☐	☐	9.2.2
5.3	Is exposure categorised according to time windows?	☒	☐	☐	9.2.2
5.4	Is intensity of exposure addressed? (e.g. dose, duration)	☒	☐	☐	9.2.2
5.5	Is exposure categorised based on biological mechanism of action and taking into account the pharmacokinetics and pharmacodynamics of the drug?	☐	☐	☒	
5.6	Is (are) (an) appropriate comparator(s) identified?	☒	☐	☐	9.2.2 & 9.3.1

Comments:

Section 6: Outcome definition and measurement		**Yes**	**No**	**N/A**	**Section Number**
6.1	Does the protocol specify the primary and secondary (if applicable) outcome(s) to be investigated?	☒	☐	☐	9.3.2
6.2	Does the protocol describe how the outcomes are defined and measured?	☒	☐	☐	9.3.2
6.3	Does the protocol address the validity of outcome measurement? (e.g. precision, accuracy, sensitivity, specificity, positive predictive value, use of validation sub-study)	☒	☐	☐	9.3.2
6.4	Does the protocol describe specific outcomes relevant for Health Technology Assessment? (e.g. HRQoL, QALYs, DALYS, health care services utilisation, burden of disease or treatment, compliance, disease management)	☐	☐	☒	

Comments:

Section 7: Bias		**Yes**	**No**	**N/A**	**Section Number**
7.1	Does the protocol address ways to measure confounding? (e.g. confounding by indication)	☒	☐	☐	9.7
7.2	Does the protocol address selection bias? (e.g. healthy user/adherer bias)	☒	☐	☐	9.8
7.3	Does the protocol address information bias? (e.g. misclassification of exposure and outcomes, time-related bias)	☒	☐	☐	9.8

Comments:

Section 8: Effect measure modification	**Yes**	**No**	**N/A**	**Section Number**
8.1 Does the protocol address effect modifiers? (e.g. collection of data on known effect modifiers, sub-group analyses, anticipated direction of effect)	☐	☐	☒	

Comments:

Section 9: Data sources	**Yes**	**No**	**N/A**	**Section Number**
9.1 Does the protocol describe the data source(s) used in the study for the ascertainment of:				
9.1.1 Exposure? (e.g. pharmacy dispensing, general practice prescribing, claims data, self-report, face-to-face interview)	☒	☐	☐	9.4
9.1.2 Outcomes? (e.g. clinical records, laboratory markers or values, claims data, self-report, patient interview including scales and questionnaires, vital statistics)	☒	☐	☐	9.4
9.1.3 Covariates and other characteristics?	☒	☐	☐	9.4
9.2 Does the protocol describe the information available from the data source(s) on:				
9.2.1 Exposure? (e.g. date of dispensing, drug quantity, dose, number of days of supply prescription, daily dosage, prescriber)	☒	☐	☐	9.3
9.2.2 Outcomes? (e.g. date of occurrence, multiple event, severity measures related to event)	☒	☐	☐	9.3
9.2.3 Covariates and other characteristics? (e.g. age, sex, clinical and drug use history, co-morbidity, co-medications, lifestyle)	☒	☐	☐	9.3
9.3 Is a coding system described for:				
9.3.1 Exposure? (e.g. WHO Drug Dictionary, Anatomical Therapeutic Chemical (ATC) Classification System)	☒	☐	☐	9.3
9.3.2 Outcomes? (e.g. International Classification of Diseases (ICD), Medical Dictionary for Regulatory Activities (MedDRA))	☒	☐	☐	9.3
9.3.3 Covariates and other characteristics?	☒	☐	☐	9.3
9.4 Is a linkage method between data sources described? (e.g. based on a unique identifier or other)	☒	☐	☐	9.3

Comments:

Section 10: Analysis plan	**Yes**	**No**	**N/A**	**Section Number**
10.1 Are the statistical methods and the reason for their choice described?	☒	☐	☐	9.3 & 9.7
10.2 Is study size and/or statistical precision estimated?	☐	☐	☒	9.3
10.3 Are descriptive analyses included?	☐	☐	☒	
10.4 Are stratified analyses included?	☐	☐	☒	
10.5 Does the plan describe methods for analytic control of confounding?	☒	☐	☐	9.7
10.6 Does the plan describe methods for analytic control of outcome misclassification?	☒	☐	☐	9.8
10.7 Does the plan describe methods for handling missing data?	☐	☐	☒	
10.8 Are relevant sensitivity analyses described?	☒	☐	☐	9.7

Comments:

Section 11: Data management and quality control	**Yes**	**No**	**N/A**	**Section Number**
11.1 Does the protocol provide information on data storage? (e.g. software and IT environment, database maintenance and anti-fraud protection, archiving)	☐	☐	☒	
11.2 Are methods of quality assurance described?	☒	☐	☐	9.7
11.3 Is there a system in place for independent review of study results?	☒	☐	☐	9.7

Comments:

Section 12: Limitations	**Yes**	**No**	**N/A**	**Section Number**
12.1 Does the protocol discuss the impact on the study results of:				
12.1.1 Selection bias?	☒	☐	☐	
12.1.2 Information bias?	☒	☐	☐	9.8
12.1.3 Residual/unmeasured confounding? (e.g. anticipated direction and magnitude of such biases, validation sub-study, use of validation and external data, analytical methods).	☒	☐	☐	
12.2 Does the protocol discuss study feasibility? (e.g. study size, anticipated exposure uptake, duration of follow-up in a cohort study, patient recruitment, precision of the estimates)	☒	☐	☐	9.6

Comments:

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

Comments:

Section 14: Amendments and deviations	**Yes**	**No**	**N/A**	**Section Number**
14.1 Does the protocol include a section to document amendments and deviations?	☒	☐	☐	5

Comments:

Section 15: Plans for communication of study results	**Yes**	**No**	**N/A**	**Section Number**
15.1 Are plans described for communicating study results (e.g. to regulatory authorities)?	☒	☐	☐	12
15.2 Are plans described for disseminating study results externally, including publication?	☒	☐	☐	12

Comments:

Name of the main author of the protocol: Daniel Prieto-Alhambra

Date: 02/04/2020

Signature: _____

ANNEX 2: NEGATIVE CONTROL OUTCOME LIST

Concept ID	Concept Name
378256	Abnormal reflex

4092879	Absent kidney
433753	Alcohol abuse
321689	Apnea
78200	Benign mammary dysplasia
4195873	Breath smells unpleasant
443792	Calculus of bile duct
434327	Cannabis abuse
197318	Cholesterolosis of gallbladder
432303	Cocaine abuse
439125	Complete trisomy 21 syndrome
433270	Cord entanglement without compression
4311591	Cramp in limb
441267	Cystic fibrosis
436233	Delayed milestone
40486120	Delay in physiological development
439791	Emotional upset
374801	Foreign body in ear
259995	Foreign body in orifice
196456	Gallstone
4166231	Genetic predisposition
434164	Glycosuria
4163735	Hemochromatosis
439871	Hemospermia
4058388	Hypertrophic scar
435522	Hypervitaminosis D
443236	Hypnotic or anxiolytic dependence
4098604	Hypomagnesemia
435371	Hypothermia
443447	Iatrogenic hypotension
374375	Impacted cerumen
4344500	Impingement syndrome of shoulder region
440382	Learning difficulties
435516	Lipoprotein deficiency disorder
438808	Mammary duct ectasia
441553	Myoclonus
4119307	Neurogenic claudication
4209423	Nicotine dependence
438130	Opioid abuse
313601	Oxygen supply absent
44782778	Pain disorder with psychological factor
4091513	Passing flatus
4022076	Patient dependence on care provider
439971	Poisoning by anticoagulant
46286594	Problem related to lifestyle
199876	Prolapse of female genital organs
4049367	Psychologic conversion disorder
440068	Psychosexual dysfunction
436246	Reduced libido
73754	Restless legs
4168212	Restlessness and agitation
138821	Seborrhea
4198492	Shoulder joint unstable
25518	Sickle cell trait

4176908	Snapping thumb syndrome
4248728	Snoring
138278	Sprains and strains of joints and adjacent muscles
4008710	Stenosis due to any device, implant AND/OR graft
40479573	Stimulant abuse
40483172	Stimulant dependence
440233	Strain of supraspinatus muscle AND/OR tendon
4194160	Thyroid function tests abnormal
4216708	Urgent desire for stool
4275889	Visual hallucinations
4193634	Worried