TITLE PAGE

PASS information

Title	Project **Sc(y)lla**: SARS-Cov-2 Large-scale Longitudinal Analyses on the comparative safety and effectiveness of treatments under evaluation for COVID-19 across an international observational data network
Protocol version identifier	1.1
Date of last version of protocol	28August2020
EU PAS register number	To be completed after protocol finalization
Active substance	
Medicinal product	
Research question and objectives	The overarching objective is to evaluate the comparative effects of COVID-19 treatments
Country(-ies) of study	France, Germany, Netherlands, South Korea, Spain, UK, and the USA. Others might join in the future.
Authors	Patrick Ryan
	Daniel Prieto-Alhambra
	on behalf of the OHDSI COVID consortium

1. TABLE OF CONTENTS

2. List of abbreviations

SARS-COV-2	Severe acute respiratory syndrome coronavirus 2
COVID-19	Coronavirus disease 2019
OMOP	Observational Medical Outcomes Partnership
OHDSI	Observational Health Data Sciences and Informatics
EHDEN	European Health Data and Evidence Network
CDM	Common Data Model
IL	Interleukin
TNF	Tumor necrosis factor
ACE	Angiotensin Converting Enzyme
RCT	Randomized clinical trial
EHR	Electronic health record
UK	United Kingdom
RWD	Real world data
RWE	Real world evidence

3. Responsible parties

Responsible party/person	**Institution/Affiliation**
Daniel Prieto-Alhambra*	Pharmaco- and Device Epidemiology, CSM-NDORMS, University of Oxford, Oxford, UK
Patrick Ryan*	Observational Health Data Analytics, Janssen Research and Development, and Department of Biomedical Informatics, Columbia University, New York, NYC, USA
Peter Rijnbeek	Medical Informatics, Erasmus University Medical Centere , Rotterdam, The Netherlands
Marc Suchard	Fielding School of Public Health, University of California at Los Angeles, Los Angeles, CA, USA

George Hripcsak	Department of Biomedical Informatics, Columbia University, New York, NYC, USA

* 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 by severe acute respiratory syndrome coronavirus 2 (SARS-Cov-2) has resulted in an unprecedented pressure on healthcare systems worldwide(1-3). The resulting disease, coronavirus disease 2019 (COVID-19), has resulted in a great number of casualties worldwide(4). While the number of infected patients continues to increase across the world, routinely collected health data will accumulate, therefore becoming a source of information for the generation of reliable evidence. Federated access to international data assets mapped to the Observational Medical Outcomes Partnership (OMOP) common data model (CDM) provides a unique opportunity to make a difference in the current crisis.

Numerous and valuable pieces of evidence have recently been generated based on historical data during an intensive international virtual Study-A-Thon hosted by the Observational Health Data Science and Informatics (OHDSI) community(5, 6). More importantly, this initiative fostered and further stimulated the setting up of a growing international network of real-world data partners, that will be updated regularly in the weeks to come. It is through this continuous international

effort that we will aim to study the safety and potential effectiveness of existing medicines are used and will be used for the treatment of COVID-19 at a global scale.

A recent review paper has summarised key emerging therapies being repurposed as treatments for the management of COVID-19, including disease-modifying anti-rheumatic drugs (hydroxychloroquine, IL6 inhibitors, JAK-1 inhibitors), systemic steroids, antibiotics (e.g. azithromycin), and anti-retrovirals, amongst others (7). Many additional medicines are under discussion or being trialled in new studies almost daily in different parts of the world (8), therefore calling for a continuous thorough but rapid evaluation of their safety and anti-viral effectiveness. More recent guidance from NIH (link) and others have resulted in the further division of therapies into the following classes: a) antiviral therapy; b) Immune-based therapy; c) Antithrombotic therapy; and d) Concomitant medications (Angiotensin-Converting Enzyme (ACE) Inhibitors, Angiotensin Receptor Blockers (ARBs); Statins; others). Many additional therapies are currently under investigation in clinical trials that do not fall under any of the categories above, and preliminary findings show common use of various antibiotics in routine practice settings.

Until there is sufficient evidence to inform more definitive treatment guidelines (https://covid19treatmentguidelines.nih.gov/), all treatments under investigation or in compassionate use for this new disease require ongoing monitoring for safety and effectiveness across various healthcare contexts.

The OHDSI community offers the potential to conduct this ongoing monitoring in observational data in the framework of scientifically stringent observational investigations. This protocol outlines the analytic plan to enable this monitoring.

8. RESEARCH QUESTION AND OBJECTIVES

The aim of this study is to assess the comparative safety and effectiveness of all emerging drug therapies used in COVID-19 treatment.

Specifically, the study has the following objectives:

1. To assess comparative effectiveness and safety among treatments administered during hospitalization and prior to intensive services
2. To assess comparative effectiveness and safety among treatments administered after COVID-19 positive testing or diagnosis in outpatient setting without prior hospitalization

Out of scope:

3. To assess comparative effectiveness and safety among treatments administered prior to COVID-19 positive testing and hospitalization (e.g. prophylaxis treatment)
4. To assess comparative effects of COVID-19 treatments compared to no treatment
5. To assess comparative effectiveness and safety among treatments administered during hospitalization after initiating intensive services

9. RESEARCH METHODS

9.1. STUDY DESIGN

A series of multinational, multi-database network comparative cohort studies will be conducted, to include:

1. New user comparative study cohorts to compare outcomes between different treatments hypothesized to have anti-viral activity for SARS-COV-2 and being currently investigated and/or used as primary COVID-19 treatment
2. New user comparative cohort study estimating the incidence of outcomes between treatments currently investigated and/or used in routine practice as concomitant therapy for COVID-19 complications

3. New user comparative cohort study estimating the incidence of outcomes between treatments currently investigated and/or used in routine practice as antibiotic therapy for COVID-19
4. New user comparative cohort study estimating the incidence of outcomes between treatments currently investigated and/or used in routine practice as immune-based therapy for COVID-19
5. New user comparative cohort study estimating the incidence of outcomes between treatments currently investigated and/or used in routine practice as antithrombotic therapy for COVID-19
6. New user comparative cohort study estimating the incidence of outcomes between treatments currently investigated and/or used in routine practice as concomitant cardiovascular prevention (anti-hypertensive or statin) therapy for COVID-19
7. New user comparative cohort study estimating the incidence of outcomes between treatments currently investigated and/or used in routine practice as other concomitant therapy/ies for COVID-19

9.2. SETTING

Participants from at least 7 European countries (Belgium, Netherlands, Germany, France, Italy, Spain, and the UK), the United States of America, and South Korea will be included. Additional data sources might be added on a voluntary collaboration basis by invitation through the OHDSI and European Health Data and Evidence Network (EHDEN) research networks. Electronic health records (EHR) and administrative claims from primary care and secondary care will be utilised. Data will be linked to additional data sources including laboratory test data (e.g. Catalan central registry of COVID-19 PCR tests), hospital administrative records (eg English Hospital Episode Statistics Admitted Patient Care database), and other relevant data sources where relevant and possible.

The study will be conducted using data from multiple real world data (RWD) sources previously mapped to the OMOP CDM in collaboration with the OHDSI and EHDEN initiatives.

9.2.1. STUDY PERIOD

The study period, when index events and outcomes of interest can be observed, will start from 01/01/2020 and end at the latest available date for all data sources in 2020 or beyond. Data will be updated regularly where possible, and as frequently as possible, in collaboration with local data partners.

9.2.2. STUDY POPULATION: INCLUSION/EXCLUSION CRITERIA

Participants will be identified using pre-specified cohort definitions 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 drug user study cohorts

1. To assess comparative effectiveness and safety among treatments administered during hospitalization and prior to intensive services

Study cohorts will be defined, to include subjects for all cohorts that, at time of initiation of an index treatment:

- Are aged 18 or over at cohort entry
- Have at least 365 days of continuous observation time prior to cohort entry
- Have 0 prior exposures to index treatment in the 365 days prior to index
- Have at least 1 COVID-19 diagnosis or positive test results in the 30 days prior to or on index
- Are hospitalized on index, defined by an inpatient visit with an admission date in the 30 days prior to or on index and no corresponding discharge date prior to or on index
- Have 0 intensive services in the 30 days prior to or on index

2. To assess comparative effectiveness and safety among treatments administered after COVID-19 positive testing or diagnosis in outpatient setting without prior hospitalization

Study cohorts will be defined, to include subjects for all cohorts that, at time of initiation of an index treatment:

- Be aged at or over 18 at cohort entry
- Have at least 365 days of continuous observation time prior to cohort entry
- Have 0 prior exposures to index treatment in the 365 days prior to index
- Have at least 1 COVID-19 diagnosis or positive test results in the 30 days prior to or on index
- Have 0 inpatient visits in the 30 days prior to or on index

For each research question above, two sets of treatments will be evaluated: drugs repurposed with in vitro activity against COVID-19 and concomitant therapies(7). The latter will be further subdivided for within group comparison into the following as per more recent guidelines: a) antibiotics; b) immune-based; c) antithrombotic; d) concomitant cardiovascular prevention; e) other concomitant therapies.

These treatments were initially identified through systematic review of CDC treatment guidelines, clinicaltrials.gov, or other authoritative sources. We augmented with all drugs on the COVID19 RCT

tracker (https://www.covid-trials.org/)(8) with >=2 trials that are marketed products with RxNorm concept are being included, as of 4May2020. We solicited feedback from industry, academia, and government agencies to be as inclusive as possible. The full list of products is available in Appendix 1.

The list may be updated based on new information, including if additional products are identified as initiated by ≥5% of patients in the study population of patients with COVID-19.

We will not attempt to construct an 'untreated' cohort, because of concerns of intractable confounding by indication and risk of immortal time bias that can be induced by looking across a hospital admission to classify exposure.

9.2.3. Follow up

Index date is defined by the first prescription/dispensation of a treatment.

Three periods of follow-up will be considered for all outcomes:

In a *fixed 7-day time-at-risk* analysis, the analysis follow-up starts 1 day after therapy initiation and continues up until the first of: 7 days after therapy initiation, death, or end of observation period.

In a *fixed 30-day time-at-risk* analysis, the analysis follow-up starts 1 day after therapy initiation and continues up until the first of: 30 days after therapy initiation, death, or end of observation period.

In an *on-treatment* analysis, the analysis follow-up starts 1 day after therapy initiation and continues until the first of: discontinuation of treatment, death, or end of observation period.

9.3. Variables

9.3.1. Exposures

Each treatment listed in section 9.2.2 will be defined based on standard concepts in the OMOP Standardized Vocabularies. The analysis package will provide the final cohort definitions, concept sets and associated source codes as provided by the standardized vocabularies.

Exposure assessment

Exposure to a treatment will commence on the date of the first qualifying record, subject to satisfying all inclusion criteria listed in section 9.2.2. Each drug exposure record has a start date and inferred end date, which is either explicitly entered or derived from other available information, such as days supply or refills. Exposure will be inferred to continue until treatment discontinuation, defined as the end date following

the last of one or more records for the treatment that occur within a persistent window. A persistence window of 7 days between drug utilization records for each study drug will be allowed considered as continuous exposure. Stockpiling will be ignored. We will not perform analyses of dose.

9.3.2. OUTCOMES

Each research question has a different set of effectiveness outcomes, while all questions share a common set of safety outcomes. All outcomes are listed in Appendix 2. These cohorts have been defined using ATLAS, an open-source analysis application developed by the OHDSI community (atlas.ohdsi.org). The analysis package will provide the final cohort definitions, concept sets and associated source codes. Many safety outcomes were defined and applied in a previously published paper by the OHDSI community [https://www.ncbi.nlm.nih.gov/pubmed/31668726].

9.3.3. NEGATIVE CONTROLS

We will use a sample of negative controls as a study diagnostic to evaluate and calibrate for residual systematic error in each analysis strategy applied to each database. Negative controls are exposure-outcome pairs for which there is no expected causal relationship, such that unbiased analyses can be expected to generate effect estimates consistent with relative risk = 1. Candidate negative controls will be identified a priori based on existing literature , and confirmed through clinical review.

9.3.3. COVARIATES

Baseline covariates will be defined by observations prior to the index date. Specific pre-index characteristics to use for confounding adjustment include:

- Age = year (cohort start date) – year of birth, categorized into 5-year intervals
- Biologic sex
- Index month
- Condition groups (SNOMED + descendants), >=1 occurrence during the 30 days prior to 1 day prior to index
- Condition groups (SNOMED + descendants), >=1 occurrence during the 365 days prior to 1 day prior to index
- Drug era groups (ATC/RxNorm + descendants), >=1 day during the interval of 30 day prior to 1 day prior to index which overlaps with at least 1 drug era
- Drug era groups (ATC/RxNorm + descendants), >=1 day during the interval of 365 day prior to 1 day prior to index which overlaps with at least 1 drug era

Cohort-based features to add as custom covariates:

- Hypertension
- Type 2 diabetes mellitus
- Chronic obstructive pulmonary disease (COPD) without asthma
- Asthma without COPD
- Chronic kidney disease
- Heart disease
- Cancer
- Obese
- Smoker
- Autoimmune condition (T1DM, RA, PsO, PsA, MS, SLE, Addisons, Graves, Sjogrens, Hashimoto, Myasthenia gravis, Autoimmune vasculitis, Pernicious anemia, Celiac disease, Scleroderma, Sarcoid, Ulcerative colitis, Crohns)

For research question #2: "To assess comparative effectiveness and safety among treatments administered after COVID-19 positive testing and prior to hospitalization", pre-index covariates about condition and drugs will include observations on the index date.

9.4. DATA SOURCES

This study will be conducted using routinely collected data from different data sources that participate in the OHDSI or EHDEN initiatives. These databases will provide representative clinical information as collected in actual routine practice conditions in different European , North America, and Asia-Pacific healthcare settings, including electronic health record and administrative claims data from primary and secondary care. All interested researchers with access to data sources with qualifying patient-level data converted to OMOP CDM will be encouraged to participate. Data partners will run cohort diagnostics for all the Scylla exposure and outcome cohorts before conducting the analyses. This will be implemented as part of the analytical package to be distributed to partners in the network.

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 size calculation was performed; instead, a minimum detectable rate ratio (MDRR) will be estimated for each drug pair-outcome

analysis in each of the available databases. All exposure cohorts must have greater than 5 patients for outcome analyses to be performed. Analyses will not produce effect estimates if there are no outcomes in either arm, otherwise will produce estimates even if unpowered. Analyses will be run with accumulating data over time to improve precision of estimate. However, with accumulating data some analyses may later become possible to conduct.

9.6. DATA MANAGEMENT

All data extraction and curation will be conducted using the ATLAS tool, an open access software generated by the OHDSI community, as well as the OHDSI Health Analytics Data-to-Evidence Suite (HADES), a set of R packages developed and maintained by the OHDSI community. (https://ohdsi.github.io/Hades/)

The process will follow the steps described here:

1. Identification of the study populations
2. Identification of the comparator and treatment cohorts
3. Identification of the different outcome cohorts
4. Review of cohort diagnostics including age and sex-specific incidence rates for face validity

The different study cohorts will be identified after searching the OMOP vocabulary by data scientists with experience with the use of OMOP and ATLAS, in collaboration with 2 clinicians and clinical epidemiologists.

Cohort definitions will be exported from ATLAS, packaged in study R packages relying on the OHDSI Methods Library, and shared with each of the data partners for a consistent extraction and curation of the population, exposures and outcomes of interest.

9.7. DATA ANALYSIS

Comparative Cohort Analysis

One study design, with multiple different analysis variants, will be conducted after appropriate diagnostics to rule out power and/or confounding issues.

For each research question in section 8, we have identified treatments that 1) have in vitro antiviral activity for SARS-COV-2 virus, or 2) are considered concomitant therapies for the COVID-19 disease. The latter are subdivided into antithrombotics, antibiotics, immune-based therapies, concomitant anti-hypertensive, concomitant anti-diabetics, concomitant statin, and other concomitant treatments (see Appendix 1). For each question, we will make pairwise comparisons of all treatments within these eight categories of therapies (e.g. compare all drugs with antiviral activity against each other, and separately compare all antithrombotic therapies with each other). For each comparison, we will estimate and

compare the incidences of each outcome in Appendix 2 during the time-at-risk windows defined in 9.2.3.

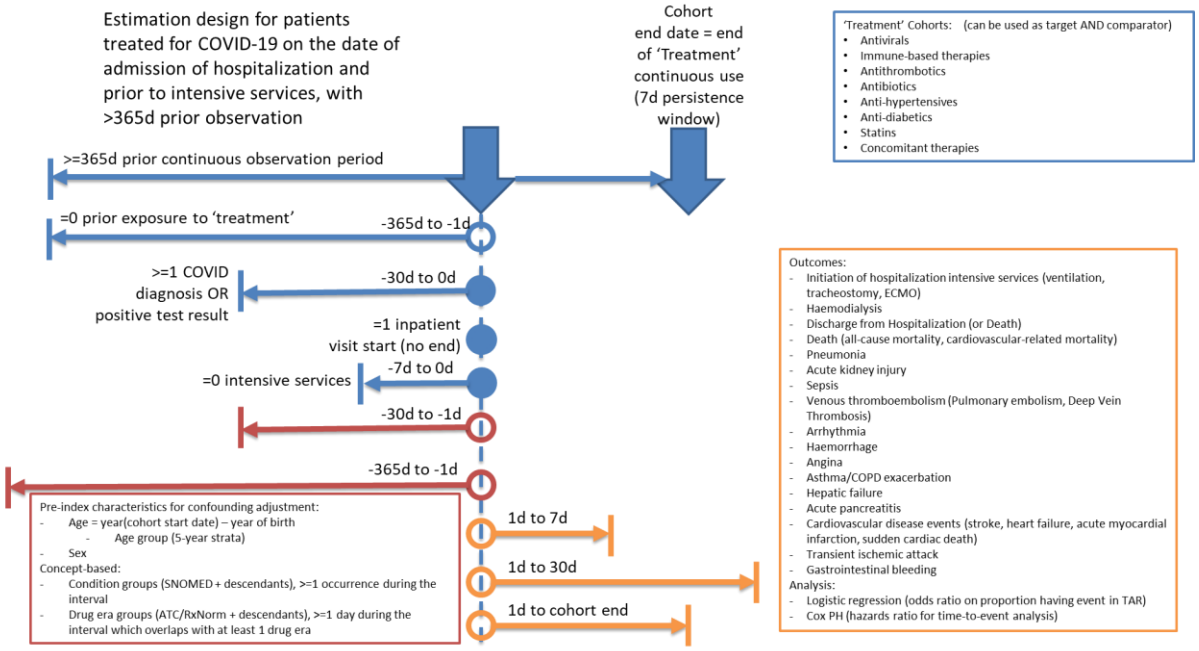
Only the first exposure in the proposed comparisons/groups per person will be included. We will require 365 days as the minimum continuous observation time prior to index date for a person to be included in the cohort. In addition, an analysis will be conducted looking at therapies started on hospital admission date where no previous lookback will be required, with all covariates (potential confounders) assessed on admission date.

If persons qualify for both the target and comparator cohorts, they will be only included in the first cohort they qualify for, and their time-at-risk be censored when the new time-at-risk start to prevent overlap. We will restrict the analysis to the period when both exposures are observed. We remove persons that have the outcome prior to the risk window start.

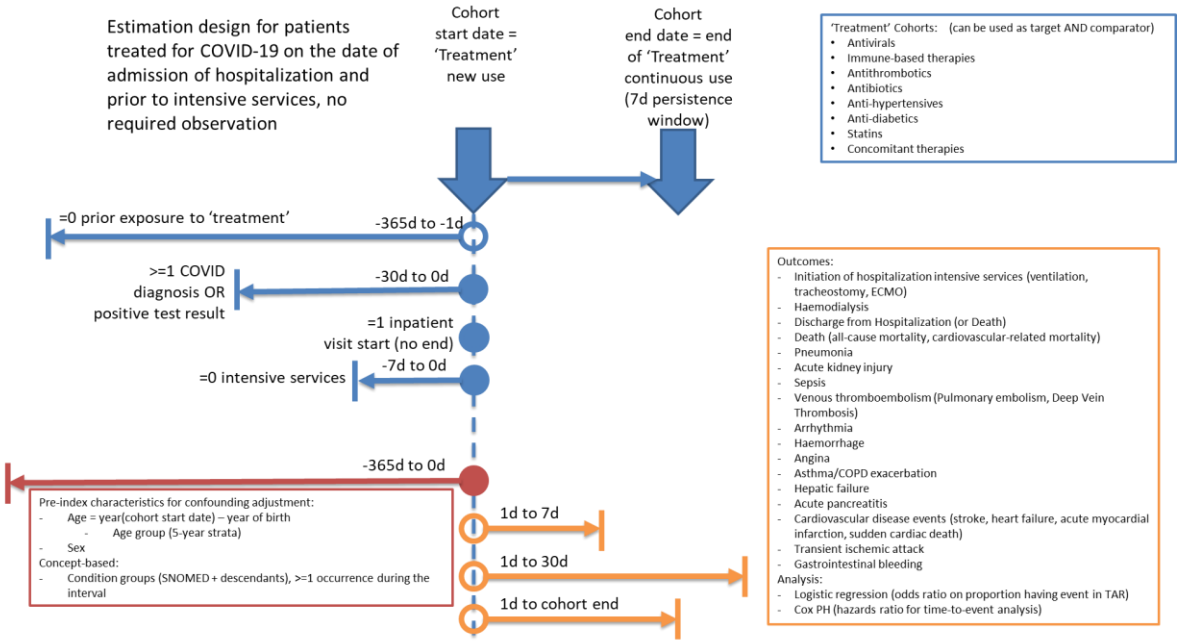
We have completed a characterization of the proposed drug user cohorts, and full results are presented in an interactive web app, and used to inform the current version of study protocol (Ver 1.1).

The figures below summarize the exposures, outcomes, time-at-risk, and pre-index covariates for each question:

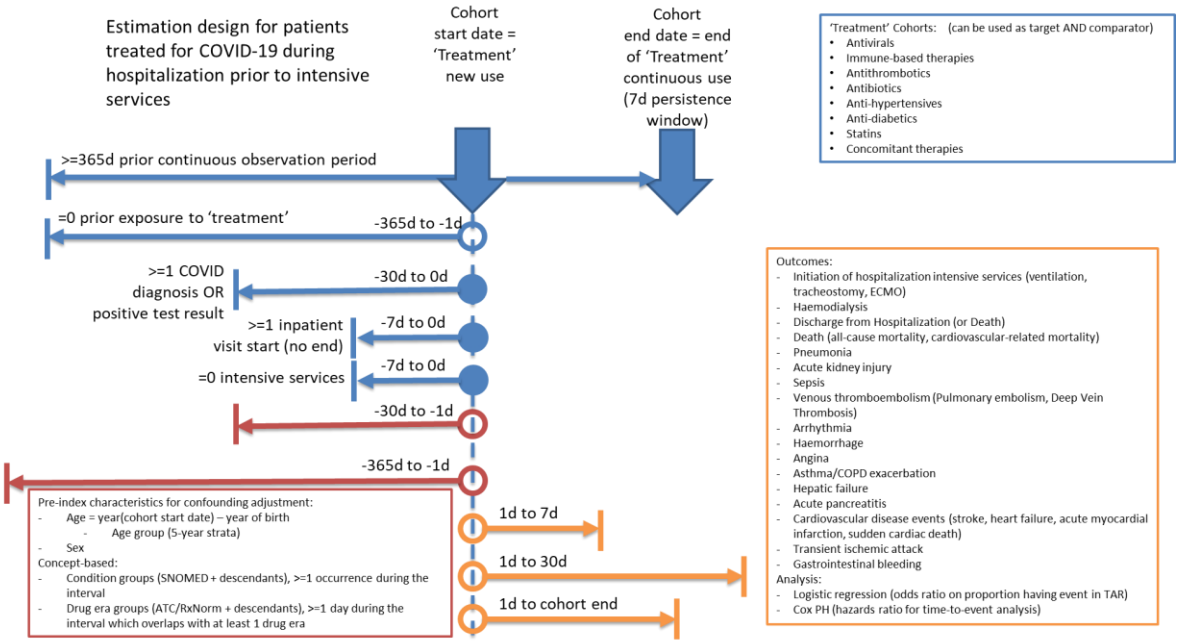
1. To assess comparative effectiveness and safety among treatments administered on the date of admission of hospitalization and prior to intensive services



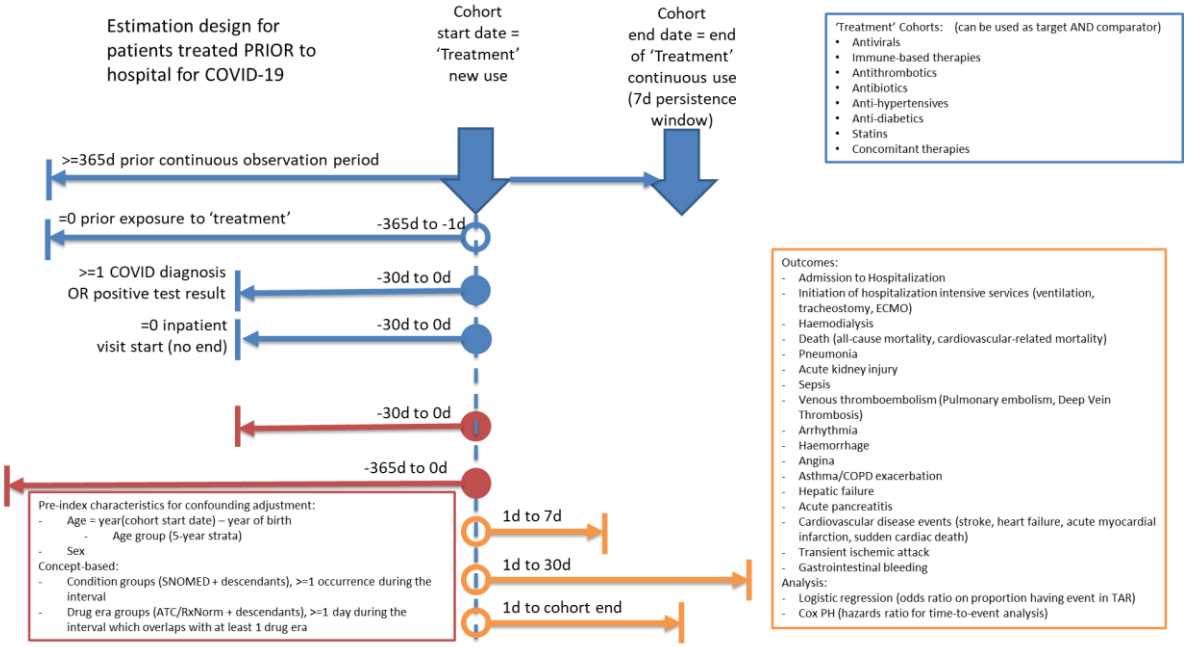
2. To assess comparative effectiveness and safety among treatments administered on the date of admission of hospitalization and prior to intensive services, with no prior observation



3. To assess comparative effectiveness and safety among treatments administered during hospitalization and prior to intensive services



4. To assess comparative effectiveness and safety among treatments administered after COVID positive testing and prior to hospitalization



Two alternative outcome models will be used for each target-comparator-outcome-'time-at-risk' combination:

1. logistic regression to estimate the incidence odds ratio based on the proportion of persons observed to experience the outcome during the time-at-risk.
2. Cox proportional hazards model to estimate the hazard ratio based on the time-to-event, in those databases where outcome dates are accurately captured.

For rare outcomes, the incidence rate ratio and the incidence odds ratio will be quite similar.

For each outcome model, patients will be excluded from the cohorts if they have previously experienced the outcome any time prior to index.

The primary confounding adjustment strategy that will be employed is propensity score adjustment through large-scale modelling, using demographics all condition and drug groupings in the 30 days and 365 days pre-index intervals as baseline covariates(9, 10). A similar approach has been recently used to test the comparative effectiveness and safety of anti-hypertensives(11). In the analysis of inpatient (pre-intensive care) treatments started on the

date of admission amongst data with no prior observation time, propensity scores will be estimated based on condition covariates recorded on the day of hospitalization. Propensity scores will not be trimmed in any of the analyses.

Propensity scores will be used for exposure 1:1 matching (using a caliper = 0.2 * standardized logit) with unconditional outcome modelling, 1:100 matching with conditional outcome modelling, and stratification (5 strata based on all persons in both target and comparator cohorts) with conditional outcome modelling, Crude estimates will be produced to evaluate the extent of estimate stability following adjustment.

Study diagnostics will be applied blinded to outcome results to all analysis variants to evaluate confidence in the results. These diagnostics will include:

- Empirical equipoise: % of persons in target and comparator with preference score (scaled propensity score) between 0.25 and 0.75: passed diagnostic if equipoise >= 50%
- Covariate balance: Standardized mean difference (SMD) for all pre-index covariates after adjustment; passed diagnostic if SMD for all pre-index covariates <= 0.1
- Negative control calibration: % of negative controls with statistically significant estimates ($p<0.05$); passed diagnostics if <=5%
- Empirical null distribution: the mean and standard deviation for the error distribution, as estimated from the negative control sample; pass diagnostic if abs(mean) < 0.1

Additional diagnostics may be considered.

If diagnostics do not pass or there is otherwise concern about residual systematic error, then post-hoc sensitivity analyses that employ alternative confounding adjustment strategies may be considered. All results will be publicly disclosed, including estimates where study diagnostics failed, but analyses with failed diagnostics will be noted as such.

All comparative cohort analyses we will rely on the CohortMethod package (https://ohdsi.github.io/CohortMethod/).

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

All the proposed analyses will be conducted for each database separately, with estimates combined in DerSimonian-Lard random effects meta-analysis methods when I2 is <=40%. No meta-analysis will be conducted when I2 for a given drug-outcome pair is >40%. Network meta-analysis will additionally be performed.

Analyses will be performed by participating data partners on a regular interval, and results will be compiled. At each iteration, source-specific study diagnostics will be reviewed prior to unblinding effect estimates, and combining across data sources. Full analysis will be performed at each time interval on cumulative data, with new results replacing old results for a given source. The number of statistical tests performed across data looks will be summarized, but no adjustment for multiplicity will be applied.

9.8. LIMITATIONS OF THE RESEARCH METHODS

Selection bias

Selection bias might arise as the consequence of including subjects with a specific period of prior observation 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 covariates. 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.

Experimental treatments used during randomized trials may not be captured in source data, which would result in exposure misclassification.

In addition, lack or incomplete recording of safety events may lead to misclassification of the proposed safety endpoints. All contributing data sources will run cohort diagnostics to explore the feasibility of identifying all the proposed study outcomes. This has been completed for some of the contributing databases, and is reported for inspection here: https://data.ohdsi.org/ScyllaCharacterizationDiagFeature/

Confounding

Confounding may occur if there are differences in (observed or unmeasured) baseline characteristics between the comparator and target cohorts which are also associated with outcome. Analysis strategies outlined in 9.7 offer approaches to identify and adjust for confounding, but may be insufficient if there is model misspecification or insufficient sample and/or data to precisely ascertain baseline covariates or to ascertain some factors at all, which can result in residual confounding and some remaining confounding bias. There may be unmeasured confounding associated with COVID severity.

9.9 PROTOCOL FOR ADDRESSING STUDY FAILURES

Because of the nature of this study, with new OMOP data sites, newly coded COVID-19 data, new analytic methods, and vastly shortened timelines, there is likely to be a large proportion of sites with zero cases or failed analyses. We therefore here pre-specify our approach to address study failures, while still conforming to the LEGEND principles.

1. On discovery that one or more sites fail to recover cases despite knowledge that such cases exist in the databases, the sites will investigate the source of the case loss and report back centrally for the phenotyping group to reassess the phenotype definition and adjust the common central

definition, asking all sites to rerun the study on the new phenotypes. Site-specific phenotype definitions will not be permitted for this estimation study.

2. If a site fails to produce estimates due to low sample size and cases are accumulating at the site, then a reanalysis will be permitted with a 50% increase in the number of cases.
3. If a particular analytic method fails to produce estimates for the majority of sites, that method will be pulled from the analysis.
4. If almost all analytic methods fail to produce estimates for almost all hypotheses at almost all sites, then the analytic methods will be redesigned centrally for adjustment to the study protocol and a restart of the study.

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. Each data partner is required to provide statement about IRB approval or exemption to participate.

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. Docherty AB, Harrison EM, Green CA, Hardwick HE, Pius R, Norman L, et al. Features of 16,749 hospitalised UK patients with COVID-19 using the ISARIC WHO Clinical Characterisation Protocol. medRxiv. 2020.
2. Grasselli G, Zangrillo A, Zanella A, Antonelli M, Cabrini L, Castelli A, et al. Baseline Characteristics and Outcomes of 1591 Patients Infected With SARS-CoV-2 Admitted to ICUs of the Lombardy Region, Italy. JAMA. 2020.
3. Guan W-j, Ni Z-y, Hu Y, Liang W-h, Ou C-q, He J-x, et al. Clinical Characteristics of Coronavirus Disease 2019 in China. New England Journal of Medicine. 2020;382(18):1708-20.
4. COVID-19 Map - Johns Hopkins Coronavirus Resource Center.
5. Burn E, You SC, Sena A, Kostka K, Abedtash H, Abrahao MTF, et al. An international characterisation of patients hospitalised with COVID-19 and a comparison with those previously hospitalised with influenza. medRxiv. 2020.
6. Lane JCE, Weaver J, Kostka K, Duarte-Salles T, Abrahao MTF, Alghoul H, et al. Safety of hydroxychloroquine, alone and in combination with azithromycin, in light of rapid wide-spread use for COVID-19: a multinational, network cohort and self-controlled case series study. medRxiv. 2020.
7. Sanders JM, Monogue ML, Jodlowski TZ, Cutrell JB. Pharmacologic Treatments for Coronavirus Disease 2019 (COVID-19): A Review. JAMA. 2020.
8. Thorlund K, Dron L, Park J, Hsu G, Forrest JI, Mills EJ. A real-time dashboard of clinical trials for COVID-19. Lancet Digital Health. 2020.
9. 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).
10. 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.
11. Suchard MA, Schuemie MJ, Krumholz HM, You SC, Chen R, Pratt N, 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.

APPENDIX 1: TREATMENTS TO EVALUATE

Each row represents a distinct treatment cohort. Each column represents the treatment decision context for which a therapy will be under investigation. Within each decision context, all drugs with in vitro antiviral activity against SARS-COV-2 virus (AV) will be compared against each other; all immune-based (IB) therapies against each other; anti-thrombotics (AT) against each other; antibiotics (AB) against each other; cardiovascular prevention therapies (CV) against each other; statins (S) against each other; anti-diabetic (AD) against each other; and all other concomitant (C) against each other.

Category	Exposure
AB	Amoxicillin
AB	Ceftriaxone
AB	Fluoroquinolones
AB, AV	Azithromycin
AD	DPP-4 inhibitors
AD	GLP1 inhibitors
AD	metformin
AD	SGLT2 inhibitors
CV	ACE inhibitors
CV	Angiotensin receptor blockers (ARBs)
CV	Losartan
CV	Non-Dihydropyridine Calcium Channel Blocker (ndCCB)
CV	Thiazide or thiazide-like diuretics
CV	Statins
AT	acenocoumarol
AT	alteplase
AT	apixaban
AT	aspirin
AT	bemiparin
AT	clopidogrel
AT	dabigatran
AT	edoxaban
AT	enoxaparin

AT	heparin
AT	prasugrel
AT	rivaroxaban
AT	warfarin
AV	Chloroquine
AV	emtricitabine
AV	favipiravir
AV	HIV Protease inhibitors
AV	Hydroxychloroquine
AV	Hydroxychloroquine + Amoxicillin
AV	Hydroxychloroquine + Azithromycin
AV	Hydroxychloroquine + ceftriaxone
AV	Hydroxychloroquine + Fluoroquinolones
AV	Interferon alfa 2a
AV	Interferon beta 1a
AV	Interferon beta 1b
AV	itraconazole
AV	Ivermectin
AV	oseltamivir
AV	raloxifene
AV	Remdesivir
AV	ribavirin
AV	ritonavir/lopinavir
C	abiraterone
C	Alpha-1 blockers (Doxazosin, prazosin, terazosin, tasulosin, silodosin, alfuzosin)
C	Antifibrinolytics (tranexamic acid, aminocaproic acid)
C	BCG vaccine
C	bicalutamide
C	Colchicine
C	doxazosin
C	prazosin
C	tamsulosin
C	terazosin
C	tranexamic acid
IB	adalimumab
IB	Anakinra

IB	baricitinib
IB	Bevacizumab
IB	cholecalciferol
IB	dexamethasone
IB	eculizumab
IB	etanercept
IB	famotidine
IB	Fingolimod
IB	H2 receptor antagonist
IB	Hydrocortisone
IB	Ibrutinib
IB	IL6 inhibitors (tocilizumab, sarilumab, siltuximab)
IB	infliximab
IB	JAK inhibitors (ruxolitinib, tofacitinib, oclacitinib, baricitinib, fedratinib, upadacitinib)
IB	prednisone or prednisolone
IB	Ruxolitinib
IB	sarilumab
IB	siltuximab
IB	tacrolimus
IB	TNF inhibitors (infliximab, golimumab, etanercept, certolizumab pegol, adalimumab)
IB	tocilizumab
IB	tofacitinib
IB	ustekinumab

APPENDIX 2: OUTCOMES TO EVALUATE

Outcome	Estimation design for patients after COVID-19 in outpatient setting without prior hospitalization	Estimation design for patients treated in hospital for COVID-19
Effectiveness outcomes		
Admission to Hospitalization	Y	
Discharge from hospitalization		Y
Pneumonia	Y	
Pneumonia during hospitalization	Y	Y
Acute respiratory distress syndrome (ARDS) during hospitalization	Y	Y
Acute kidney injury (AKI) using diagnosis codes only during hospitalization	Y	Y
Acute kidney injury (AKI) using diagnosis codes and dialysis and change in measurements during hospitalization	Y	Y
Sepsis during hospitalization	Y	Y
Venous thromboembolic (pulmonary embolism and deep vein thrombosis) during hospitalization	Y	Y
Pulmonary embolism during hospitalization	Y	Y
Deep vein thrombosis during hospitalization	Y	Y
Heart failure during hospitalization	Y	Y
Arrhythmia (heart block/bradycardia, supraventricular arrythmia, ventricular tachycardia/ventricular fibrillation/cardiac arrest/sudden cardiac death) during hospitalization	Y	Y
Heart block/bradycardia during hospitalization	Y	Y
Supraventricular arrythmia during hospitalization	Y	Y
Ventricular tachycardia/ventricular fibrillation/cardiac arrest/sudden cardiac death during hospitalization	Y	Y
Death	Y	Y
Hospitalization requiring intensive services (ventilation OR tracheostomy OR ECMO)	Y	Y
Hospitalization requiring mechanical ventilation	Y	Y
Hospitalization requiring tracheostomy	Y	Y
Hospitalization requiring ECMO	Y	Y

Outcome	Estimation design for patients after COVID-19 in outpatient setting without prior hospitalization	Estimation design for patients treated in hospital for COVID-19
Hospitalization requiring haemodialysis	Y	Y
Safety outcomes		
Hospitalization for asthma exacerbation	Y	
Hospitalization for COPD exacerbation	Y	
Haemorrhage	Y	Y
Chest pain or angina	Y	Y
Angina	Y	Y
Hepatic failure	Y	Y
Acute pancreatitis	Y	Y
Total cardiovascular disease events (ischemic stroke, hemorrhagic stroke, heart failure, acute myocardial infarction or sudden cardiac death)	Y	Y
Gastrointestinal bleeding	Y	Y
Cardiovascular-related mortality	Y	Y
Transient ischemic attack	Y	Y
Stroke (ischemic or hemorrhagic)	Y	Y
Ischemic stroke	Y	Y
Hemorrhagic stroke	Y	Y
Acute myocardial infarction	Y	Y