Non-interventional Study Protocol

Real-world characteristics, management and outcomes of subjects screened or diagnosed with COVID-19 in Spain

Version	25th March 2020
Sponsor(s) and contact person	Dra. Yolima Cossio (principal investigator) Information Systems Unit Collaborators: Dra. Soledad Romea Zaira Benitez Andres Antón Eva Aurín Hospital Vall d´Hebron Passeig de la Vall d'Hebron, 119 ES-08035 Barcelona (Spain) Dr. Juan Pablo Horcajada (principal investigator) Dr. Jordi Martinez (collaborator) Infectious Diseases Department Hospital del Mar, Barcelona Passeig Marítim de la Barceloneta, 25-29 ES-08003 Barcelona (Spain)
Responsible Research Institutions	Study design and data ownership: - Hospital Vall d´Hebron, Barcelona - Hospital del Mar, Barcelona Data extraction and codification: - Iomed Medical writing and statistical analysis: - TFS Observational Health Data Sciences and Informatics (OHDSI) community: - OMOP-based phenotype definition - Analysis packages for characterization, patient-level prediction and population-level estimation (SQL, R) - Technical support for data partners executing OHDSI study packages
Countries	Spain
Author	Neus Valveny

	TFS (Trial Form Support SL) Consell de Cent, 334-336, 4ª floor ES-08009 Barcelona, Spain Phone: +34 617 414 124 E-mail address: neus.valveny@tfscro.com

Background

Current coronavirus disease 2019 (COVID-19) pandemic has an urgent need for answering many questions regarding epidemiology and natural history of the disease, similarities/differences with other viral infections and effectiveness/safety of current preventive/therapeutic measures.

Models assessing most relevant prognostic factors in the daily practice will be also crucial to help healthcare providers deciding which subjects should be prioritized and given more intensive therapies.

The recent report from the Chinese Center for Disease Control and Prevention, including 72 314 confirmed/suspected cases (up to February 11, 2020) gives preliminary insights to some of these questions: 87% of cases were aged 30 to 79 years; 81% were mild (non-pneumonia and mild pneumonia) and 14% severe (defined by dyspnea, respiratory frequency ≥30/min, blood oxygen saturation ≤93%, partial pressure of arterial oxygen to fraction of inspired oxygen ratio <300 and/or lung infiltrates > 50% within 24 to 48 hours). The overall case-fatality rate (CFR) was 2.3% (1023 deaths among 44 672 confirmed cases). The CFR was higher in subjects with comorbidities (10.5% for cardiovascular disease, 7.3% for diabetes, 6.3% for chronic respiratory disease, 6.0% for hypertension, and 5.6% for cancer)(1).

Some prognostic factors, previously suggested in the literature, currently being discussed are: gender, age, current or former smoking (2), treatment with angiotensin converting enzyme inhibitors (ACEI)/angiotensin II receptor antagonists (ARA), nonsteroidal anti-inflammatory drugs (NSAIDs) or immunosuppressive agents, prior cardiovascular disease and diabetes. Some of these factors seem to be associated with the expression of the angiotensin-converting enzyme 2 (ACE2), which is very likely to serve as the binding site for the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the strain implicated in the current COVID-19 epidemic(3).

Treatments that may assist to contain its spread and reduce the high mortality rates are urgently needed (4). Due to the amount of time needed to develop an effective vaccines, already existing therapeutics are being assessed. In this context, chloroquine, a broadly used antimalarial drug, could benefit the treatment of patients infected by SARS-CoV-2 (5), with or without other agents like azithromycin. Moreover, the potential of combining direct-acting antiviral (e.g. lopinavir, ritonavir, darunavir or remdesivir) with an anti-inflammatory agent (Interleukin-6 receptor inhibitors, anti-tumor necrosis factor alpha (TNFα) agents, Janus kinase (JAK) inhibitors) such as tolicizumab or baricitinib has also been suggested (6).

With the increasing availability of real-world COVID-19 data worldwide, the OHDSI community has proposed to conduct a large, world-wide, retrospective study to answer the main questions described above. This community composed of researchers from many countries will provide OMOP-based phenotype definitions for viral infection and viral pneumonia, together with analysis packages for characterization, patient-level prediction and population-level estimation in the data sets available at each participating country.

The present study aims to participate in this international collaboration by analyzing data from Spanish hospitals with the standardized tools provided by the OHDSI collaboration. Further analyses to answer specific, local COVID-19 research questions that may arise during the ongoing crisis might be also considered.

Research objectives

The general aim is to help addressing current gaps in the management of COVID-19 pandemic by:

1) Describing the characteristics of subjects screened and/or diagnosed with COVID-19 in Spain
2) Describing therapeutic management in the clinical practice
3) Describing clinical outcomes for the infected subjects and compare them with those from other viral infections and pneumonias
4) Describing healthcare resource use
5) Assessing predictor factors for infection, hospitalization, intensive care admission and death
6) Exploring effectiveness and safety of administered therapies and compare them with same use in other viral infections and pneumonias

Specific questions linked to the aforementioned objectives include:

- Epidemiological (population-level) questions:
  - Number of daily identified COVID-19 cases, deaths and location
  - Age/gender standardized mortality rate
  - Time between exposure and symptom initiation
  - Time between exposure and positive test
  - Time between recovery and negative test
  - Which are the subgroups in which the test is most likely to test positive? (percentage of positive COVID-19 test results among: subjects without symptoms [contact with infected only], upper respiratory infection (URI) symptoms without pneumonia and confirmed pneumonia on chest-radiography)
  - Behavioral and socio-economic risk factors for symptomatic/asymptomatic infection: households / institutions vs community; health-care workers vs general population; other active workers vs confined population

- Patient-level prediction questions:
  - Does chloroquine/hydroxychloroquine (HCQ) confer a lower risk of viral infection/s, viral pneumonia or worse pneumonia outcomes?
  - Do IL6-R/JAK inhibitors (e.g. baricitinib) or anti-TNFα confer a lower risk of viral infection/s or of viral pneumonia or worse pneumonia outcomes?
  - Do antiretrovirals (e.g. remdesivir) confer a lower risk of viral infection/s or of viral pneumonia or worse pneumonia outcomes?
  - Do the drugs above reduce morbi-mortality amongst viral pneumonia sufferers?
  - Do systemic steroids reduce morbi-mortality amongst viral pneumonia sufferers?
  - Is ACEI/ARB use associated with a higher or lower risk of viral infection/s or of viral pneumonia or worse pneumonia outcomes? (overall and separately in hypertensive, diabetics and both conditions at the same time)
  - Are other concomitant medications (e.g. statins) associated with a higher risk of viral infection/s or of viral pneumonia or worse pneumonia outcomes?
  - Are subjects who are immunosuppressed at higher risk of viral pneumonia or worse pneumonia outcomes)?
  - Are hypertension or diabetes associated to poorer prognosis relative to immunosuppression status?

Methods

Study design

The study is a non-interventional, retrospective, database, cohort study based on anonymized and routinely-collected health care data from several Spanish Hospitals which has been mapped to the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM). The study period will be:

1. COHORT 1 = COVID-19 cohort (screened and/or infected): starting from November-2019 until the latest available data collection, with sufficient data quality, for each participating hospital.

2. COHORT 2 = other viral infections and/or pneumonias: starting from January -2010 until the latest available data collection (up to November 2019), with sufficient data quality, for each participating hospital.

For each subject, retrospective follow-up period with extend from the index date (COHORT 1: date of first COVID-19 test and/or confirmed infection; COHORT 2: date of diagnosis of viral infection and/or pneumonia) until complete resolution. Data about prior comorbidities and/or medications will have a longer look-back period (as long as data is available in the data source).

Study participants

COHORT 1:
Study participants will

- COHORT 1 A (screened): Have a recorded COVID-19 test with available result
  And/or
- COHORT 1 B (infected): Have a confirmed/suspected COVID-19 infection (either through a positive test or through physician´s criterion based on a combination of symptoms and exposure)

COHORT 2:
Study participants will

- COHORT 2 A (viral infection): Have a recorded prior viral infection
  And/or
- COHORT 2 B (viral pneumonia): Have a recorded prior viral pneumonia

Exposures

See Table *1*. Exposures will be identified on the basis of records in the procedure occurrence table in the OMOP CDM. Eligibility criteria will be assessed using the condition and procedure occurrence tables. OMOP CDM 'standard codes' will be specified, with conditions identified using SNOMED codes and procedures mapped to a number of different codes, including SNOMED, ICD10PCS, ICD9Proc, and CPT4.

Covariates

See Table *1*.

Outcomes of interest and time-at-risk

All-cause mortality will be assessed from index date to 1) 30 days, 2) 90 days and 2) 180 days. Safety events/complications (e.g. opportunistic infections, etc.) will be assessed from index date to 30 days post-infection resolution.

Analytic methods

Since this is a descriptive and exploratory study with no formal hypothesis testing, no sample size calculation has been done.

All subjects fulfilling the selection criteria for cohorts 1 or 2 during the respective study periods will be included and analyzed using available data.

A descriptive summary of the characteristics of subjects from each cohort and relevant subgroup of interest will be provided. Both their characteristics at index date and their prior medical history will be summarized.
Event counts and incidence rates will be provided for each of the outcomes of interest.
Prediction models will be developed for mortality and other adverse outcomes using multivariable Cox or logistic regression, as applicable. Internal validation (using data extracted from the same database) and external validation (using data extracted from a different database) will be used to evaluate the models. We will assess model discrimination using the area under the receiver operator curve (AUROC) and area under the precision recall curve (AUPRC). We will assess the calibration of the models using a graphical calibration plot, where patients are partitioned into deciles based on the predicted risk with the mean predicted risk calculated and the fraction of the patients who experienced the outcome during the time at risk calculated for each group. External validation will be performed using other databases mapped to the OMOP CDM.

Table 1. Exposures, covariates and outcomes

Objective	Independent Variables (including exposures and covariates)	Outcome(s)	Statistical Analysis
1-Characteristics of study cohorts	Sociodemographic and lifestyle (age, gender, smoking status and history, socioeconomic status, living situation, etc.) Geographical/Ethnic origin History of exposure to COVID-19 and dates; test results and viral load, if available Current and past concomitant medications (including, but not limited, all those listed in "specific questions" above) Current and past comorbidities Diagnosis, clinical symptoms and signs (confusion, multilobar infiltrates, pleural effusion) and date of findings Vital Signs (Tº, heart rate, respiratory rate, blood pressure, oxygen saturation, arterial oxygen pressure/fraction of inspired oxygen (PaO2/FiO2) ratio) Laboratory values at diagnosis (hematology and biochemistry, including blood urea nitrogen , CRP, procalcitonin, blood count, etc.) Severity scores (such as the Pneumonia Severity Index [PSI]; the confusion, elevated blood urea nitrogen level, respiratory rate, and blood pressure [CURB] score; and the CURB plus age ≥65 years [CURB 65] score)		Descriptive analysis in the overall sample and by subgroups of interest (e.g. Hospital, gender, age subgroups, smoking history, diabetes status, cardiovascular history etc.)
2-Describing therapeutic management in the clinical practice	Therapeutic management for COVID-19: - Steroids - Vasopressors - Antivirals - Antibiotics - NSAIDs / analgesics - Other drugs (e.g. intravenous immunoglobulin/IVIg, hydroxychloroquine, etc.) - Non-drug therapies: Oxygen therapy, mechanical ventilation, intravenous fluid rehydration		

3-Describing clinical outcomes for the infected subjects and compare them with those from other viral infections and pneumonias		Infection status Pneumonia status Recovery status and time to recovery Hospitalization and length of stay ICU admission and length of stay Overall survival Complications during infection (bacterial pneumonia, sepsis…) Recurrence/reinfection	Descriptive analysis in the overall sample and by subgroups of interest Chi-Square, Student´s T test and/or Mann-Whitney tests, as applicable
4-Describing healthcare resource use	Home visits, specialized care visits, emergency ward visits, hospitalizations, ICU admission, laboratory tests, diagnostic tests, medications, non-drug therapies		Descriptive analysis in the overall sample and by subgroups of interest
5-Assessing predictor factors for infection, hospitalization, intensive care admission and death	All variables from objectives 1 & 2	All outcomes from objective 3	Multivariable Cox Regression Multivariable Logistic Regression
6-Exploring effectiveness and safety of administered therapies and compare them with same use in other viral infections and pneumonias	All variables from objectives 1 & 2	All outcomes from objective 3	Multivariable Cox Regression Multivariable Logistic Regression Chi-Square, Student´s T test and/or Mann-Whitney tests, as applicable

Protection of human subjects

Regulatory and ethical approvals will be obtained before starting the study, as required by local regulation. Since data will be fully anonymized, patients will not need to provide written informed consent.

Limitation of Research Methods

The main limitation is that both positive and/or infected patients will present a bias towards the most severe cases. Mild cases will stay at home and will not be tested, so epidemiological estimations will be biased towards underestimated incidence/prevalence of infections and overestimated morbi-mortality in those infected. Population will be biased towards the severe, especially when moving towards the right of the exponential curve of the pandemics.

Management and reporting of adverse events/adverse reactions

Not applicable (analysis of secondary data).

References

1. Wu Z, McGoogan JM. Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72 314 Cases From the Chinese Center for Disease Control and Prevention. DOI:10.1001/jama.2020.2648. JAMA. 2020;

2. Cai H. Sex difference and smoking predisposition in patients with COVID-19. DOI: 10.1016/S2213-2600(20)30117-X. Lancet Respir Med. 2020;

3. Gurwitz D. Angiotensin receptor blockers as tentative SARS-CoV-2 therapeutics. DOI: 10.1002/ddr.21656. Drug Dev Res. 2020;

4. Watkins J. Preventing a covid-19 pandemic. BMJ. 2020;368:m810.

5. Touret F, de Lamballerie X. Of chloroquine and COVID-19. Antiviral Res. 2020;177:104762.

6. Stebbing J, Phelan A, Griffin I, Tucker C, Oechsle O, Smith D, et al. COVID-19: combining antiviral and anti-inflammatory treatments. DOI: 10.1016/S1473-3099(20)30132-8. Lancet Infect Dis. 2020;