Analysis Plan for ISARIC International COVID-19 Critically Ill Cohort according to their vaccination status

Title of proposed research
Description of vaccination history and its association with patient characteristics and outcomes in the ISARIC COVID-19 database
Version: (Date: Day/Month/Year)
14.SEP.2022
Working Group Chair (name, ORCID ID, email, institution, country)
Bronner P. Gonçalves, ISARIC, University of Oxford, UK, bronner.goncalves@ndm.ox.ac.uk
[1]Working group co-chair (name, ORCID ID, email, institution, country)
Luis Felipe Reyes, Universidad de La Sabana, Colombia, luis.reyes5@unisabana.edu.co Ignacio Martin-Loeches, Trinity College Dublin, Ireland, imartinl@tcd.ie Waasila Jassat, National Institute of Communicable Diseases, South Africa, WaasilaJ@nicd.ac.za Zeno Bisoffi, IRCCS Sacro Cuore Don Calabria Hospital, Italy, zeno.bisoffi@sacrocuore.it Chiara Piubelli, IRCCS Sacro Cuore Don Calabria Hospital, Italy, chiara.piubelli@sacrocuore.it Gianluigi Li Bassi, University of Queensland, g.libassi@uq.edu.au Sally Shrapnel, University of Queensland, s.shrapnel@uq.edu.au Piero L. Olliaro, ISARIC, University of Oxford,UK, piero.olliaro@ndm.ox.ac.uk Laura Merson, ISARIC, University of Oxford, UK, laura.merson@ndm.ox.ac.uk

[1] ISARIC supports diversity in data users and encourages SAPs to involve chairs from both lower-resourced and higher-resourced settings. Please contact ncov@isaric.org if you would like support to connect with experts to work with.

Barbara Citarella, ISARIC, University of Oxford, UK, barbara.citarella@ndm.ox.ac.uk

Joaquín Baruch, ISARIC, University of Oxford, UK, joaquinbaruch2@gmail.com

Amanda Rojek, ISARIC, University of Oxford, UK, amanda.rojek@ndm.ox.ac.uk

Christiana Kartsonaki, ISARIC, University of Oxford, UK, christiana.kartsonaki@dph.ox.ac.uk

Benjamin Lefèvre, Université de Lorraine, Nancy, France, b.lefevre@chru-nancy.fr

Jose W Lopez, Insituto Nacional de Salud del Niño San Borja, jlopez@insnsb.gob.pe

Ewen Harrison, University of Edinburgh, UK, ewen.harrison@ed.ac.uk

Introduction

Scope of document
This document details the initial plan for analysis of a subset of patients in the cohort in the ISARIC database with information on vaccination history.

Rationale for project
In this analysis, in addition to describing the frequency of history of COVID-19 vaccination (henceforth referred to simply as 'vaccinated') in patients in the ISARIC database over time and in different countries, we aim to address the following research questions:

(i) In the subset of patients who were hospitalised when vaccines were available, do vaccinated patients differ from patients without history of vaccination in terms of clinical presentation, history of comorbidities and demographics?
(ii) Are there differences in the frequency of clinical outcomes of vaccinated patients versus concurrently hospitalized patients without history of COVID-19 vaccination?

We will combine clinical information in the ISARIC database with national-level, age-specific whenever available, vaccination coverage information. Note that our objective is not to estimate vaccine effectiveness because this is not feasible with the data available, as it would require either information on non-hospitalised patients or patients hospitalised with conditions other than COVID-19.

As vaccination data might not be available for a non-negligible fraction of the study population, we will also use a second approach that relies on population-level data on vaccination coverage and variant frequency. The rationale and details of this approach are described in the *Research questions* sub-section.

Overall, we aim to advance understanding of the role that observational data can serve in assessing vaccine impact and the methods to evaluate it.

Project aims

The primary objectives of this analysis are:

1) To assess the frequency of vaccination history in patients admitted to hospital with COVID-19 across time during the pandemic and in different countries after the initiation of each country's vaccination campaign.
2) To describe the demographic composition and frequency of comorbidities in patients with a history of vaccination versus patients without a history of vaccination. To ensure that patients with different vaccination status are comparable in relation to factors other than vaccination, we will restrict the analysis to time periods when country-specific vaccination coverage is above pre-defined thresholds (see below).
3) To describe clinical outcomes (e.g., fatality risk and treatment with advanced respiratory support [i.e., IMV, NIV, or HFNC]) of patients depending on their vaccination status. To ensure that patients with different vaccination status are comparable in relation to other factors, we will restrict the analysis to time periods when country-specific vaccination coverage is above predefined thresholds.
4) To qualitatively compare analyses that use individual-level data on vaccination with analyses informed by population-level data on both vaccination coverage and variant relative frequencies.

Participatory approach

All contributors to the ISARIC database are invited to participate in this analysis through review and input on the statistical analysis plan and the resulting publication. The outputs of this work will be disseminated as widely as possible to inform patient care and public health policy; this will include submission for publication in an international, peer-reviewed journal. ISARIC aims to include the names of all those who contribute data in the cited authorship of this publication, subject to the submission of contact details and confirmation of acceptance of the final manuscript within the required timelines, per ICMJE policies and the ISARIC publication policy.

Data

Datasets from all sites with data on vaccination history are eligible for inclusion in the analysis that uses individual-level vaccination information.

Research questions

A general approach to improving comparability of vaccinated and unvaccinated individuals

For analyses that require individual-level vaccination information, the primary comparisons will use vaccination status as a binary variable since data are missing for a non-negligible number of vaccinated patients on the number of previous doses received and the vaccination date. However, exploratory analyses will also be performed and presented for the Clinical questions (1 – 3) below by vaccination dose, date, and type if sufficient data are available.

An important aspect of the analysis is the definition of the time periods that will be used for the inclusion of the study population. Information on vaccination history was not collected in the database before March 2021, hence any records from before that month will be automatically excluded from the analysis. Furthermore, vaccination coverage remained low in many countries after data on vaccination history started to be recorded in the ISARIC database. To account for this and ensure that non-vaccinated individuals included in the analysis are those recruited during the same time period when vaccinated individuals were (or could have been) recruited, data from https://ourworldindata.org/ will be used to define when vaccination started in countries included in this analysis. Comparisons will be repeated, as secondary analyses, for different periods based on national vaccine coverage: for example, vaccination coverage above 30% and above 50%. If vaccination coverage in a country did not reach the coverage threshold used, data from the country will not be included in that particular secondary analysis. If age-specific vaccine coverage data at the country level are available, we will investigate the possibility (and need) of performing analyses that account for age-specific, rather than all-ages-combined vaccination coverage. The rationale for performing these secondary analyses is that, even assuming constant vaccine effectivenesss that prevent hospitalization given infection, changes in country level coverage imply that changes in frequency of vaccination history in the hospital population will also be observed, as might changes in patient profile, in terms of comorbidities and demographic factors.

Analysis that relies on both population-level vaccination and SARS-CoV-2 variant data

A non-negligible proportion of the study population might have missing data on history of COVID-19 vaccination. For this reason, we will also use a second approach that will involve population-level data. The steps of this analysis will include:

- Country-level data on variant relative frequency will be obtained from GISAID to identify country-specific periods when most local infections were caused by a specific variant (e.g. periods when more than 90% of infections in the country were caused by a specific variant). The periods identified using this approach will be used for within-country comparisons. The rationale for using variant frequency-defined periods for this second analytical approach is to reduce variability in patient profile and clinical outcomes due to

differences in variant severity; furthermore, by limiting analysis to relatively short time periods compared to the study duration we aim to minimise potential confounding by naturally acquired immunity (i.e. previous infection).

- For each variant-specific period, in countries with sufficient data, we will use age-specific vaccination coverage data to assess whether there were changes in coverage during the period; two sub-periods will be defined, “lower coverage” and “higher coverage”, at the beginning and end of each variant period, respectively.
- Descriptive analyses will be performed comparing “lower coverage” and “higher coverage” groups (see Clinical questions 2 and 3 below), for each country-, variant-specific period, by age.

Due to the variable vaccination coverage in different countries, and changing coverage within countries in different variant-specific periods, and the likely variable increase in coverage during specific periods, analyses using this approach will be only descriptive, and restricted to country-specific periods with sufficient data (e.g. at least 50 participants in the “lower coverage” and “higher coverage” groups).

Clinical questions and planned analyses

The following questions will be addressed by analyses that use individual-level vaccination data:
Clinical question (1): How did the frequency of vaccination history in patients admitted to hospital with COVID-19 varied by country group and across temporal phases of the pandemic?

Planned statistical analyses (1): For this analysis, we will restrict the analytical dataset to the calendar time periods after vaccination campaigns started in different countries. Frequencies of vaccination history and distribution of number of doses will be presented by age categories, sex and time periods (monthly or bimonthly or longer periods). Data will be analysed by country, where sufficient data are available.

Clinical question (2): What are the demographic and clinical (symptoms and comorbidities) characteristics of patients stratified by COVID-19 vaccination history ?

Planned statistical analyses (2): As for Clinical question (1), we will restrict the analytical dataset to the calendar time periods after vaccination campaigns started in different countries. Overall frequencies of key demographic variables will be stratified by vaccination status, and countries (only those with considerable data) as well as over time.

Clinical question (3): Are the clinical outcomes (e.g., fatality risk or advanced respiratory support (i.e., IMV, NIV, or HFNC]) of patients different depending on their vaccination status?

Planned statistical analyses (3): To ensure that patients with different vaccination status are comparable in relation to other factors, we will restrict analysis to time periods when country-specific vaccination coverage is above pre-defined thresholds. Descriptive analyses will

be performed, where fatality risk will be plotted by vaccination status, and individual country (where sufficient data are available). Survival analysis methods will be used to quantify the association between vaccination status and clinical outcomes; as mentioned above, the observational association does not need to reflect the vaccine effectiveness as measured in the community. We will also perform an exploratory evaluation of the impact of the type of vaccine in the clinical outcomes, and investigate the possibility of using propensity score matching in comparisons by vaccination status.

Descriptive results using the approach described above that relies on population-level vaccination and variant data will also be presented for Clinical questions 2 and 3; analyses will be stratified by country, dominant variant, age and relative increase in vaccine coverage during the “higher coverage” versus “lower coverage” subperiods.

Statistical procedures

Descriptive statistics including tables with percentages/proportions in different categories, median/IQR, or mean/SD will be used. Cox proportional hazards models will be used to assess the association between fatality risk and the history of COVID-19 vaccination; the proportional hazards assumption will be assessed using Schoenfeld residuals. Cox models will be stratified by country. We will also investigate the feasibility of fitting survival models that account for competing events (here, discharge from hospital and death during hospital).

Handling of missing data

Preliminary analysis will be performed to ascertain a detailed overview of the extent of missingness in the data. This should enable the identification of variables which lack sufficient data to allow for any useful analysis to performed on them. Type of missingness shall be considered including whether data are not missing at random and follow-up with sites will be conducted if appropriate.