Analysis Plan for ISARIC International COVID-19 Data

This document details the plan for an analysis of COVID-19 patient data submitted to the ISARIC database. Details of this cohort can be found in the MedRxiv reports available here: https://www.medrxiv.org/content/10.1101/2020.07.17.20155218v15

All contributors to the ISARIC COVID-19 database are invited to participate in this analysis through review and input on the statistical analysis plan and 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 used in this analysis as cited collaborators on the resulting publication, subject to the ISARIC publication policy.

Title of proposed research
Clinical outcomes among elderly COVID-19 patients: An ecological analysis based on national income levels and healthcare access and quality.

Working Group Lead(s) (name, ORCID ID, email, institution, country) [1]
Jose A. Calvache, MD, PhD ORCID ID 0000-0001-9421-3717 Email jacalvache@unicauca.edu.co; j.calvacheespana@erasmusmc.nl Department of Anesthesiology, Universidad del Cauca, Colombia Department of Anesthesiology, Erasmus MC Rotterdam, The Netherlands

Working group members (name, ORCID ID, email, institution, country)
Daniel Perez, MD ORCID ID 0000-0002-6646-441X Email danielpmu@unicauca.edu.co Department of Anesthesiology, Universidad del Cauca, Colombia

Statistician (name, ORCID ID, email, institution, country)
Henry Laniado Rodas ORCID ID 0000-0002-8389-5385 Email hlaniado@unicauca.edu.co Department of Mathematics, Universidad del Cauca, Colombia Esther Sofia Mamian Email esmamian@unicauca.edu.co Department of Mathematics, Universidad del Cauca, Colombia

---

[1] Either chair and/or co-chair are based in an institution in an LMIC. If you would like to be connected with a potential partner researcher, please let us know at data@isaric.org.

Introduction

The COVID-19 pandemic impacted global health and socioeconomic conditions worldwide specially in vulnerable populations. Within the population infected with SARS-CoV-2, various factors contribute to heightened virus susceptibility and increased COVID-19 severity (1). As comorbidities tend to accumulate with age, older individuals may experience more severe COVID-19 outcomes. Age-related physiological changes, particularly in the respiratory system, are strongly associated with worse outcomes (1, 2).

Current epidemiological data demonstrate that individuals infected with SARS-CoV-2 exhibit higher mortality rates as age increases across a wide range (1, 3-5). Notably, individuals aged over 80 years are at a significantly higher risk of mortality when compared to younger age groups (4). This elevated risk underscores the likelihood of older patients requiring hospitalisation, admission to intensive care units (ICUs), or mechanical ventilation for respiratory support, highlighting critical implications for healthcare systems and potentially life-threatening outcomes for this age group (2, 6). In the United States, a significant proportion of COVID-19 fatalities occurred in individuals aged 65 and older, with this demographic accounting for 80% of deaths attributed to the disease (6).

However, there is a paucity of global data comparing the COVID-19 burden among patients older than 65 years old in low- to middle-income countries (LMICs) versus high-income countries (HICs). Strengthening research efforts, particularly in LMICs with limited healthcare resources or access to them, is essential for bolstering preparedness against future pandemics. Enhanced understanding of COVID-19's impact will better equip these nations to manage future infectious disease surges, ultimately saving lives and fortifying their healthcare systems.

Universal health coverage should ensure that all populations have access to quality health care. In 2018, the Healthcare Access and Quality index (HAQ) was introduced and developed using data from the Global Burden of Diseases, Injuries, and Risk Factors Study 2016. This index aimed to evaluate personal health-care access and quality across 195 countries and territories. HAQ Index performance diverged along the development spectrum, ranging from more than 97 in Iceland to less than 20 in the Central African Republic and Somalia revealing significant disparities and progress trends across countries and regions but also substantial improvements since 2000, particularly in some low- and middle-SDI countries (7).

This document outlines the initial analysis plan for publication regarding a subset of COVID-19 patients within the global cohort of the ISARIC database.

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 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.

Research Plan

Background, Research Objectives, Scientific Value

1) To estimate and describe the incidence of mortality, ICU admission rate, and length of hospital stay in adults older that 65 years with diagnosis of COVID-19 and severe COVID-19 reported in the ISARIC registry and stratify these estimates across low- middle- and high-income countries (according to information of the world bank).

  *Severe COVID-19 patients will be defined as those admitted to an ICU and/or those treated with one of the following treatments: invasive or non-invasive mechanical ventilation, high-flow nasal cannula, inotropes, or vasopressors (3).

2) To investigate the relationship between mortality, ICU admission rates, and hospital stay duration among adults aged 65 and older diagnosed with COVID-19 and severe COVID-19, and the Healthcare Access and Quality index by country at the population level.

  *The Healthcare Access and Quality index is a score derived from the Global Burden of Diseases, Injuries, and Risk Factors Study 2016 to assess personal health-care access and quality for 195 countries and territories. It was reported by Fullman N, GBD 2016 Collaborators (https://doi.org/10.1016/S0140-6736(18)30994-2).

Proposed Target Population

The target population comprises ageing adults (older than 65 years) who are admitted to hospital with confirmed or suspected COVID-19 and severe COVID-19 and are recorded in the ISARIC database. The definition for COVID-19 grade of severity will be based on Reyes et al (3).

Clinical Questions/Descriptive Analyses

1. What is the incidence mortality, ICU admission rate, and length of hospital stay in adults older that 65 years with diagnosis of COVID-19 and severe COVID-19 reported in the ISARIC registry?

2. What is the difference between low-, middle-, and high-income countries data mortality, ICU admission rate, and length of hospital stay in adults older that 65 years with diagnosis of COVID-19 and severe COVID-19?

3. What disparities exist at the population level among elderly adults aged over 65 with COVID-19 and severe COVID-19 in terms of mortality, ICU admission rate, and length of hospital stay, as related to the Healthcare Access and Quality index score?

4. Descriptive analysis: To present graphically the estimates of mortality, ICU admission rate, and length of hospital stay in adults older that 65 years with diagnosis of COVID-19 and severe COVID-19 (by using an animated plot and a shiny app).

Planned Statistical Analyses, Methodology and Representation

1. Demographic and baseline clinical characteristics presented in suitable data distribution.
2. Descriptive summaries of ICU admission rate, and length of hospital stay per country.
3. Survival analysis of mortality to provide descriptive measures.
4. To develop a shiny app to show the outcome estimates in adults older that 65 years with diagnosis of COVID-19 and severe COVID-19?
5. Multivariable analyses (with descriptive purposes) of included outcomes as a function of income country data and the Healthcare Access and Quality index.

Handling of Missing Data

Preliminary analysis would 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 perform on them. Variables with greater than 30% missingness will be excluded from analysis. Where appropriate, imputation will be performed using Multiple Imputation.

Other Information

March 2024: Statistical Analysis Plan (SAP) submission to ISARIC
April - May 2024: Approval of SAP from ISARIC Committees and admitting input from ISARIC investigators.
June - July 2024: Evaluate validity of data and Data analysis.
August - October 2024: Results and Write manuscript.
November - December 2024: Manuscript submission to the journal

References

1. Perrotta F, Corbi G, Mazzeo G, Boccia M, Aronne L, D'Agnano V, et al. COVID-19 and the elderly: insights into pathogenesis and clinical decision-making. Aging Clin Exp Res. 2020 Aug;32(8):1599–608.
2. Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. The Lancet. 2020 Mar 28;395(10229):1054–62.
3. Reyes LF, Murthy S, Garcia-Gallo E, Irvine M, Merson L, Martin-Loeches I, et al. Clinical characteristics, risk factors and outcomes in patients with severe COVID-19 registered in the International Severe Acute Respiratory and Emerging Infection Consortium WHO clinical characterisation protocol: a prospective, multinational, multicentre, observational study. ERJ Open Res 2021;8:00552–2021. https://doi.org/10.1183/23120541.00552-2021.
4. Covid-19 Risks and Information for Older Adults. Centers for Disease Control and Prevention, Centers for Disease Control and Prevention, 22 Feb. 2023, www.cdc.gov/aging/covid19/index.html
5. Christiana Kartsonaki, J Kenneth Baillie, Noelia García Barrio, ISARIC Clinical Characterisation Group, et al. Characteristics and outcomes of an international cohort of 600 000 hospitalized patients with COVID-19, International Journal of Epidemiology, Volume 52, Issue 2, April 2023, Pages 355–376, https://doi.org/10.1093/ije/dyad012
6. Abul Y, Leeder C, Gravenstein S. Epidemiology and Clinical Presentation of COVID-19 in Older Adults. Infect Dis Clin North Am. 2023 Mar;37(1):1–26.
7. Fullman N, Yearwood J, Abay SM, Abbafati C, Abd-Allah F, Abdela J, et al. Measuring performance on the Healthcare Access and Quality Index for 195 countries and territories and selected subnational locations: a systematic analysis from the Global Burden of Disease Study 2016. The Lancet 2018;391:2236–71. https://doi.org/10.1016/s0140-6736(18)30994-2.

Data Required

Use the links and descriptions below to define which types of data are required for your analysis. Please delete the descriptions and any domains you do not require, leaving a list of domains needed for the analysis.

(DI) Device Identifiers Domain

"The identity of a device is established by means of a number of parameters, then assigned an identifier. The parameters used for identification of a device depend on the kind of device and the needs of the study to distinguish among devices."

The Device Identifiers (DI) Domain contains information about the type of device used in the study (this is not a subject-level domain). This is currently only utilized for oxygen therapy for COVID-19 Clinical Data. This information is collected with a single row for each device with the identifier listed in **SPDEVID** and the description of the device in **DIVAL**. This domain acts a reference table to identify the devices located in the Treatments and Interventions (IN) Domain. The key between the two is the **SPDEVID** variable.

(DM) Demographics Domain

"A special purpose domain that includes a set of essential standard variables that describe each subject in a clinical study. It is the parent domain for all other observations for human clinical subjects. One record per subject."

The Demographics (DM) Domain contains demographics (e.g, Age and Sex) and other essential variables (e.g., Study Site, Country of Participation, Randmomized Treatment Arm) that describe the subject. This information is collected with a single row for each subject described in USUBJID and information about that subject recorded on the same row.

(DS) Disposition Domain

"An events domain that contains information encompassing and representing data related to subject disposition. One record per disposition status per subject."

The Disposition (DS) domain contains information about a subject's FINAL disposition status as the end of the study as provided by the data contributor. Studies that provide interim dispositions can also be capture here (e.g., an interim Day 28 status and then a final Day 63 outcome). This information is collected with a single row for each outcome (named in **DSTERM** and standardized into **DSDECOD**) and the timing of that outcome (either a date in **DSSTDTC** or a visit description in **VISIT**).

(ER) Environmental Risk Factor Domain

"The Environmental Risk Factors (ER) domain is an events domain for representing data collected to assess potential exposures to, or risk factors associated with, diseases by way of environmental contact or through participation in activities associated with risk."

The Environmental Risk Factors (ER) Domain contains information about exposure or risk factors related to diseases. This information is collected with a single row for each risk factor or potential exposure described in **ERTERM** and information about that risk factor or exposure on the same row (e.g., whether that exposure was pre-specified and if it occurred (**ERPRESP**=Y, **EROCCUR**), the category of the risk factor (**ERCAT**), and the timing of that factor (**ERSTDTC**, **EREVINTX**, etc.).

(HO) Healthcare Encounters Domain

"A events domain that contains data for inpatient and outpatient healthcare events (e.g., hospitalization, nursing home stay, rehabilitation facility stay, ambulatory surgery). One record per healthcare encounter per subject."

The Healthcare Encounters (HO) Domain contains information about a subject's hospitalizations and other encounters. This information is collected with a single row for each encounter (for example, each period of hospitalization or each stay in an intensive care unit) described in **HOTERM** and information about that encounter recorded on the same row (e.g., whether that encounter occurred (**HOOCCUR**), the reason for the encounter (**HOINDC**), the start and end dates of that encounter (**HOSTDTC** & **HOENDTC**), and the duration of that encounter (**HODUR**)).

(IN) Treatments and Interventions Domain

"An interventions domain that contains concomitant and prior medications used by the subject, such as those given on an as needed basis or condition-appropriate medications (Concomitant and Prior Medications Domain). An interventions domain that contains the details of a subject's exposure to protocol-specified study treatment. Study treatment may be any intervention that is

prospectively defined as a test material within a study, and is typically but not always supplied to the subject (Exposure Domain). An interventions domain that contains information about protocol-specified study treatment administrations, as collected (Exposure as Collected Domain). One record per recorded intervention occurrence per subject."

This is an IDDO-created Custom Domain (non-standard SDTM implementation). All variables found in the Concomitant and Prior Medications (CM), Exposure (EX), and Exposure as Collected (EC) domains are found in this custom domain.

The Treatments and Interventions (IN) Domain contains information on treatments, interventions, and procedures used or given to the subject, both past and present. This information is collected as a single row for each dose/occurrence of a treatment, intervention, or procedure named in **INTRT** with information about that treatment, intervention, or procedure on the same row (e.g., the dose and unit of the intervention (**INDOSE**, **INDOSU**), whether a pre-specified intervention occurred (**INPRESP**=Y, **INOCCUR**), and the timing of that intervention (**INSTDTC**, **INENDTC**, **INSTRF**, etc.).

(LB) Laboratory Test Results Domain
"A findings domain that contains laboratory test data such as hematology, clinical chemistry and urinalysis. This domain does not include microbiology or pharmacokinetic data, which are stored in separate domains. One record per lab test per time point per visit per subject."
The Laboratory Test Results (LB) Domain contains information about laboratory tests collected (e.g, Hemoglobin, Bilirubin, pH, G6PD Activity). This information is collected with a single row for each test (named in **LBTESTCD** and **LBTEST**) and the result for that test (in raw form within **LBORRES** and in standardized form within **LBSTRESC** and **LBSTRESN**).

(MB) Microbiology Domain
"The Microbiology Specimen domain is used for the detection, identification, quantification, and other characterizations of microorganisms in subject samples, except for drug susceptibility testing. All non-host infectious organisms, including bacteria, viruses, fungi, and parasites are appropriate for the microbiology domains. One record per microbiology specimen finding per time point per visit per subject"
The Microbiology Specimen Test Results (MB) Domain contains information about microbiology specimen tests collected (e.g., malaria blood smears, blood cultures, COVID-19 PCR tests). This information is collected with a single row for each test (named in MBTESTCD and MBTEST) and the result for that test (in raw form within MBORRES and in standardized form within MBSTRESC and MBSTRESN).

(RP) Reproductive System Findings Domain
"A findings domain that contains physiological and morphological findings related to the male and female reproductive systems. One record per finding or result per time point per visit per subject."
The Reproductive System Findings (RP) Domain contains information about the subject's reproductive ability (like Last Menstrual Period or Childbearing Potential) and history (like Number of Pregnancies and Pregnant Indicator). This information is collected with a single row for each finding (named in **RPTESTCD** and **RPTEST**) and the result for that finding (in raw form within **RPORRES** and in standardized form within **RPSTRESC** and **RPSTRESN**).

(RS) Disease Response and Clinical Classification Domain
"A findings domain for the assessment of disease response to therapy, or clinical classification based on published criteria. One record per response assessment or clinical classification assessment per time point per visit per subject."
The Disease Response and Clinical Classification (RS) Domain contains two types of information - clinical classifications of a subject's status (like the Glasgow Coma (GCS) Score and the Pediatric Risk of Mortality (PRISM) Score) and determination of disease response to treatment (like the WHO Treatment Response for Malaria). This information is collected with a single row for each test (named in **RSTESTCD** and **RSTEST**) and the result for that test (in raw form within **RSORRES** and in standardized form within **RSSTRESC** and **RSSTRESN**).

(SA) Clinical and Adverse Events Domain
"An events domain that contains data describing untoward medical occurrences in a patient or subjects that are administered a pharmaceutical product and which may not necessarily have a

causal relationship with the treatment (Adverse Events Domain). An events domain that contains clinical events of interest that would not be classified as adverse events (Clinical Events Domain). One record per event per subject."

This is an IDDO-created Custom Domain (non-standard SDTM implementation). All variables found in the Adverse Events (AE) and Clinical Events (CE) domains are found in this custom domain. The Symptoms and Adverse Events (SA) Domain contains information on clinical events that may or may not be defined as Adverse Events by the contributor as well as previous events (i.e., medical history). This information is collected as a single row for each event described in **SATERM** with information about that event on the same row (e.g., a categorization of the event (**SACAT**, **SASCAT**), whether a pre-specified event occurred (**SAPRESP**=Y, **SAOCCUR**), the severity or seriousness of the event (**SASEV**, **SASER**), and the timing of the event (**SASTDTC**, **SAENDTC**, **SADUR**, **SASTRF**, etc.).

[(SC) Subject Characteristics Domain](#)
"A findings domain that contains subject-related data not collected in other domains. One record per characteristic per subject."

The Subject Characteristics (SC) domain contains information about a subject that is not collected in other domains (e.g., marital status, education level, information about children like preterm birth or breastfed status). This information is collected with a single row for each characteristic (described in **SCTESTCD** and **SCTEST**) and the result for that characteristic (in raw form within **SCORRES** and in standardized form in **SCSTRESC/N**).

[(SV) Subject Visits Domain](#)
"A special purpose domain that contains the actual start and end date/time for each visit of an individual subject."

This is a modified domain (non-standard SDTM implementation). This is used to capture the date of enrollment for COVID-19 datasets, as the date of hospitalization is used for the RFSTDTC value. Other visits are not captured for COVID-19 and this domain is not utilized in other IDDO themes. The Subject Visits (SV) Domain contains information about the enrollment data for COVID-19 clinical patients. This information is collected as a single row with a description of the event in **VISIT** and the date of enrollment in **SVSTDTC** for each patient.

[(TI) Trial Inclusion Exclusion Criteria Domain](#)
"A trial design domain that contains one record for each of the inclusion and exclusion criteria for the trial. This domain is not subject oriented. It contains all the inclusion and exclusion criteria for the trial, and thus provides information that may not be present in the subject-level data on inclusion and exclusion criteria. One record per inclusion/exclusion criterion."

The Trial Inclusion/Exclusion Criteria (TI) Domain is a study-level domain (i.e., there is no individual patient-level data in this domain) that contains information on the inclusion and exclusion criteria used for the study. This information is captured with each individual criterion described in **IETEST** with an indication of whether it is inclusion or exclusion in **IECAT**.

[(TS) Trial Summary Domain](#)
"A trial design domain that contains one record for each trial summary characteristic. This domain is not subject oriented. The Trial Summary (TS) dataset allows the sponsor to submit a summary of the trial in a structured format. Each record in the Trial Summary dataset contains the value of a parameter, a characteristic of the trial. For example, Trial Summary is used to record basic information about the study such as trial phase, protocol title, and trial objectives. The Trial Summary dataset contains information about the planned and actual trial characteristics. One record per trial summary parameter."

[(VS) Vital Signs Domain](#)
"A findings domain that contains measurements including but not limited to blood pressure, temperature, respiration, body surface area, body mass index, height and weight. One record per vital sign measurement per time point per visit per subject."

The Vital Signs (VS) Domain contains information about vital signs tests collected (e.g, Height, Blood Pressure, Temperature). This information is collected with a single row for each test (named in **VSTESTCD** and **VSTEST**) and the result for that test (in raw form within **VSORRES** and in standardized form within **VSSTRESC** and **VSSTRESN**).