ISARIC (International Severe Acute Respiratory and Emerging Infections Consortium)

A global federation of clinical research networks, providing a proficient, coordinated, and agile research response to outbreak-prone infectious disease

Analysis Plan for ISARIC International COVID-19 Patients

Please complete the following sections:

Title of proposed research
Applicability of commonly used COVID-19 clinical case definitions and severity score amongst patients hospitalized with COVID-19 in low-and middle-income countries
Version: (Date: Day/Month/Year)
24 March 2021
Working Group Chair (name, ORCID ID, email, institution, country)
Dr. Md. Zakiul Hassan, MBBS Institution: MSc candidate, International Health and Tropical Medicine, University of Oxford, United Kingdom Email: zakiul.hassan@gtc.ox.ac.uk
[1]**Working group co-chair (name, ORCID ID, email, institution, country)**
Professor Dr Buddha Basnyat, MD, MSc, FACP, FRCP Director, Oxford University Clinical Research Unit-Nepal, Email: buddha.basnyat@ndm.ox.ac.uk
Statistician (name, ORCID ID, email, institution, country)
Dr. Aronrag Meeyai, PhD Institution: Centre for Tropical Medicine and Global Health, University of Oxford, United Kingdom Email: aronrag.meeyai@ndm.ox.ac.uk

---

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

Final draft SAPs will be circulated to all ISARIC partners for their input with an invitation to participate. ISARIC can help to set up collaborator meetings; form a working group; support communications; and accessing data. Please note that the details of all approved applications will be made publicly available on the ISARIC website. Please complete all sections of this form fully and return to ncov@isaric.org

Introduction

This document details the initial analysis plan for publication on a subset of COVID-19 patients in the global cohort in the ISARIC database, as of 30 April 2021. There are currently 64 countries (as of 10 February 2021) contributing data and these have so far contributed data on 305, 241patients. This data will represent the global experience of ~ 12 months of this pandemic.

Recent studies have shown that a substantial number of COVID-19 patients have presented with less-typical symptoms, not captured in commonly used or adopted COVID-19 clinical case definitions (e.g. case definition by the World Health Organization, the US Centers for Disease Control and Prevention, the European Centers for Disease Control and prevention and, Public Health England) (1-3). Atypical symptoms are more frequently reported in children (≤18 years), older adults (≥70 years) and women (2-4). A sensitive clinical case definition is vital for hospital admission policy, patient management decisions, disease surveillance and prevention activities. A precise case definition is also helpful for avoiding unintentional cohorting/exposure to uninfected patients, as well as diagnosing co-circulating pathogens/other illnesses. Due to lack of availability and limited resources for testing in resource-limited settings, clinical decision-making is more reliant on clinical case definition (5). Hence, understanding the sensitivity of commonly used clinical case definition for COVID-19 in low and middle-income countries (LMICs) and assessing the need for alternative case definitions is essential.

Similarly, prognostic scores are also vital for such settings to support the optimal use of limited hospital and ICU beds (6). A predictive score allows for the identification of patients who need advanced care or can be discharged safely for home care (7). The currently available risk stratification tools for COVID-19 have limitations for use in LMICs. Most models were derived and validated in high-income countries. They might not predict the correct risk in other populations such as those in LMICs. Additionally, most tools have many parameters that might not be available in a low-income setting hospital. Many models were derived from an elderly patient cohort (e.g. median age of 73 years) with a high mortality rate (e.g. 32.2%) which might not be applicable in LMIC-settings with a younger age population (8, 9). Additionally, other population characteristics, disease prevalence (such as a higher burden of infectious diseases in LMICs compared to high-income settings) and risk factors (such as malnutrition) vary by settings so as the admission criteria.

We propose therefore to assess the applicability of commonly used clinical case definitions for COVID-19 and the adequacy of existing prognostic scores to the hospitalized patient in LMIC using the ISARIC COVID-19 cohort.

Participatory Approach

All LMIC 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

Summary of Research Objectives
The overall aim for this project is to describe the ISARIC-LMIC data in terms of completeness and conduct following analysis given the availability of required data: 1. To assess the adequacy of common covid-19 clinical case definitions to covid-19 hospitalized patients in LMICs with reference to sensitivity 2. To assess applicability/adequacy of common severity scores to covid-19 hospitalized patients in LMICs

Proposed Target Population
● Patients enrolled from LMICs (World Bank income categories) ● All suspected cases on admission ● Lab-confirmed cases ● Patients with the known outcome (for objective 2)

Clinical Questions/Descriptive Analyses
Objective 1: 1.1 What are the commonly missing/unknown presenting symptoms in the ISARIC dataset? 1.2 What proportion of patients are missing all symptoms? 1.3 What are the most commonly reported symptoms? 1.4 What proportion of patients meet each set of clinical case definitions (sensitivity)? 1.5 What proportion of patients do not meet any of the set clinical case definitions? 1.6 What proportion of patients present with each symptom stratified by age group and sex? 1.7 What proportion of patients meet each set of clinical case definition stratified by age group and sex?

1.8   What are the frequently reported atypical symptoms not captured by case definitions?

1.9   What are the symptoms which frequently co-occur as clusters?

Objective 2:

2.1   What proportion of patients are missing the required variables for each of the candidate risk scores?

2.2   Which candidate score is more applicable for LMIC settings in terms of availability of data?

2.3   What proportion of patients has a complete set of data for each set of risk score?

2.4   What is the predictive performance of each candidate score among the study population?

Planned Statistical Analyses, Methodology and Representation

Methods for objective 1 (adequacy of case-definition):

Study participants:

- All hospitalized lab-confirmed COVID-19 cases reported from LMICs (according to World Bank income category)

Variables of interest:

- country
- age
- sex at birth
- symptoms (as indicated in the ISARIC case report form)
- Date of symptom onset (for a possible temporal analysis as case-definition may have been updated)
- Date of hospitalization
- SARS-CoV-2 confirmation

Analysis plan:

- We will collate symptoms according to four clinical case definitions. We will calculate the proportion of patients meeting each set of case definitions with a 95% confidence interval and in feasible will be stratified by age group, sex, and WHO region.
  1. World Health Organization: fever plus cough; or any three of fever, cough, fatigue, headache, myalgia, sore throat, rhinorrhea, shortness of breath, nausea and vomiting, diarrhoea, and confusion (10);
  2. US Centers for Disease Control and Prevention: At least two of fever (measured or subjective), chills, rigors, myalgia, headache, sore throat, new olfactory and taste disorder(s); or one of cough, shortness of breath, or difficulty breathing (11);
  3. Public health England: new cough, or temperature ≥37.8 °C, or a loss or change in sense of smell or taste (12);
  4. European Centre for Disease Prevention and Control: At least one of cough, fever, shortness of breath and sudden onset anosmia, ageusia or dysgeusia (13).

Questions	Planned statistical analysis	Planned representation in the manuscript
1.1 What are the commonly missing/unknown presenting symptoms in the ISARIC dataset?	frequency distribution of missing symptoms as per the ISARIC case report form (stratified by age, sex and country)	summary table
1.2 What proportion of patients are missing all symptoms?	frequency and percentages	summary table
1.3 What are the most commonly reported symptoms?	distribution of symptoms by age, sex and country	summary table/bar plot
1.4 What proportion of patients meet each set of clinical case definitions?	We will create a variable for each case definition and will calculate the percentage with 95% confidence interval	bar plot
1.5 What proportion of patients do not meet any of the set clinical case definitions?	frequency and percentages	summary table
1.6 What proportion of patients present with each symptom stratified by age group and sex?	percentage with 95% confidence interval	bar plot
1.7 What proportion of patients meet each set of clinical case definitions stratified by age group and sex?	percentage with 95% confidence interval	bar plot
1.8 What are the frequently reported atypical symptoms not captured by case-definition?	frequency distribution of atypical symptoms by age, sex, and country	summary table

1.9 What are the symptoms that frequently co-occur as clusters?	the distribution of combinations of 3/5 most common symptoms, amongst all patients	summary table

Methods for objective 2 (applicability of prognostic score):
Study population/Participants:
- Patients enrolled from LMICs (world Bank income category)
- All suspected cases on admission
- Lab-confirmed cases
- Patients with known outcome

Variables of interest (not exhaustive):
Age, sex at birth, BMI, Ethnicity, temperature, pulse rate, systolic BP, diastolic BP, comorbidities, acute respiratory distress Syndrome, acute kidney injury, shock, radiographic chest infiltrates, pleural effusion, receiving oxygen (when oxygen saturation measured), Glasgow Coma Scale, confusion/altered mental state, respiratory rate, admission oxygen saturation, arterial Ph, arterial oxygen tension (PaO2), hematocrit, blood glucose, sodium, urea, CRP, lymphocytes, D-dimer, LDH, ICU admission, mortality

Outcomes:
- In-hospital mortality

Candidate scores:
- COVID-19 specific risk score: At the time of analysis, we will select 10 candidate scores from an ongoing living systematic review (14) predicting in-hospital mortality among hospitalized patients with reasonable predictive performance, adequate sample size, derived and validated using real-patient data.
- Common risk scores for respiratory infections and acute illness: scores that will be included are CURB65, PSI, QSOFA, PMEWS, ATS/IDSA and NEWS2)

Questions	Planned statistical analysis	Planned representation in the manuscript
2.1 What proportion of patients are missing the required variables for each of the candidate risk scores?	frequency and percentages (stratified by country)	summary table

2.2 Which score is more applicable in terms of availability of data?	missing variable (stratified by each set of prognostic score)	summary table
2.3 What is the proportion of patients that have a complete set of data for each candidate risk score?	patient frequency (stratified by candidate score)	summary table
2.4 What is the predictive performance of each candidate score among the study population?	Calibration: For each set of risk scores we will plot the observed outcomes in the ISARIC dataset against the predicted outcomes across deciles of mortality risk to determine over or under prediction. Discrimination: We will plot the receiver operating characteristic (ROC) curves for each score and will calculate the area under the curve. We will calculate the sensitivity, specificity, positive predictive value, negative predictive value, and 95% CIs with various thresholds. Additionally, considering the impact of time, we will measure the predictive performance of selected scores at different time points such as during early months of the pandemic and more recent data.	ROC curve/summary table

Limitations and further plan:
We assume that data reported to ISARIC from LMICs may not have an adequate number of patients with required parameters due to completeness of data, limiting our ability for external validation of the candidate risk scores. If the external

validation of the existing score is not feasible due to insufficient numbers, we propose to adapt/develop an alternative score suitable for LMIC-settings. If neither validating existing scores nor generating a new score is possible, we will critically evaluate and discuss the importance and challenges of data sharing from low-and-middle-income countries with a particular focus on quality and missingness of data. We will also highlight the potential role of different stakeholders including ISARIC to address the challenges and promote more quality data contribution in a timely manner from LMICs in global data-sharing platforms.

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 be 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. Variables with greater than 30% missingness will be excluded from analysis. Where appropriate, imputation will be performed using Multiple Imputation by Chained Equations (MICE).

References

1. Manoharan L, Cattrall JW, Harris C, Newell K, Thomson B, Pritchard MG, et al. Early clinical characteristics of Covid-19: scoping review. MedRxiv. 2020.
2. Gómez-Belda AB, Fernández-Garcés M, Mateo-Sanchis E, Madrazo M, Carmona M, Piles-Roger L, et al. COVID-19 in older adults: What are the differences with younger patients? Geriatrics & gerontology international. 2021;21(1):60-5.
3. Chuang DT, Aydemir S, Magda P, Thomas C, Zarnegar R. Neurological manifestations as primary presentation of COVID-19 in hospitalized patients. Acta Neurologica Scandinavica. 2021.
4. Pritchard MG. COVID-19 symptoms at hospital admission vary with age and sex: ISARIC multinational study. medRxiv. 2020.
5. Siow WT, Liew MF, Shrestha BR, Muchtar F, See KC. Managing COVID-19 in resource-limited settings: critical care considerations. Springer; 2020.
6. Phua J, Weng L, Ling L, Egi M, Lim C-M, Divatia JV, et al. Intensive care management of coronavirus disease 2019 (COVID-19): challenges and recommendations. The Lancet Respiratory Medicine. 2020;8(5):506-17.
7. Haegdorens F, Monsieurs KG, De Meester K, Van Bogaert P. An intervention including the national early warning score improves patient monitoring practice and reduces mortality: A cluster randomized controlled trial. Journal of advanced nursing. 2019;75(9):1996-2005.
8. Knight SR, Ho A, Pius R, Buchan I, Carson G, Drake TM, et al. Risk stratification of patients admitted to hospital in the United Kingdom with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: development and validation of a multivariable prediction model for mortality. British Medical Journal. 2020.

9.   Dowd JB, Andriano L, Brazel DM, Rotondi V, Block P, Ding X, et al. Demographic science aids in understanding the spread and fatality rates of COVID-19. Proceedings of the National Academy of Sciences. 2020;117(18):9696-8.

10.   Organization WH. WHO COVID-19 Case definitions: World Health Organization; 2021 [Available from: https://www.who.int/publications/i/item/WHO-2019-nCoV-Surveillance_Case_Definition-2020.2.

11.   Prevention CfDCa. Coronavirus Disease 2019 (COVID-19) 2020 Interim Case Definition, Approved April 5, 2020: Centers for Disease Control and Prevention; 2020 [Available from: https://wwwn.cdc.gov/nndss/conditions/coronavirus-disease-2019-covid-19/case-definition/2020/.

12.   England Ph. COVID-19: investigation and initial clinical management of possible cases: Public health England; 2021 [cited 2021. Available from: https://www.gov.uk/government/publications/wuhan-novel-coronavirus-initial-investigation-of-possible-cases/investigation-and-initial-clinical-management-of-possible-cases-of-wuhan-novel-coronavirus-wn-cov-infection.

13.   Control ECfDPa. Case definition for coronavirus disease 2019 (COVID-19), as of 3 December 2020: European Centre for Disease Prevention and Control; 2021 [cited 2021. Available from: https://www.ecdc.europa.eu/en/covid-19/surveillance/case-definition.

14.   Wynants. Update to living systematic review on prediction models for diagnosis and prognosis of covid-19. BMJ-BRITISH MEDICAL JOURNAL. 2020;370.

Working Group Members

FIRST NAME	LAST NAME	EMAIL
Luis Felipe	Reyes	luisreyve@unisabana.edu.co felipereye@gmail.com
Arie Zainul	Fatoni	ariezainulfatoni@ub.ac.id
Anna	Beltrame	beltrame.anna@gmail.com
Christiana	Kartsonaki	christiana.kartsonaki@dph.ox.ac.uk
Jan Cato	Holter	jcholter@gmail.com
Andrea	Cortegiani	andrea.cortegiani@unipa.it
Pedro	Povoa	pedrorpovoa@gmail.com
James	Lee	james.lee@ndm.ox.ac.uk
Barbara	Citarella	barbara.citarella@ndm.ox.ac.uk

Marina	Wainstein	m.wainstein@uq.edu.au
Stephen	Knight	stephenknight@doctors.org.uk
Daniel	Munblit	daniel.munblit08@imperial.ac.uk
Jose Andres	Calvache	jacalvache@unicauca.edu.co