



---

## *ANALYSIS PLAN FOR PUBLICATION ON TIME VARIATION IN THE 'INPATIENT JOURNEY'*

---

**July 2020**

### Introduction

This document details the analysis plan for a proposed project exploring temporal changes in the 'inpatient journey' of hospitalised COVID-19 patients over the course of the pandemic, following the gradual accumulation of data.

### 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.

### Research questions/ Hypotheses

- 1) Research questions:
  - a. Does the 'inpatient journey' change over time in the course of the COVID-19 pandemic?
  - b. Do patient's presenting characteristics change in the course of the COVID-19 pandemic?
  - c. Do patient's presenting characteristics explain changes in 'inpatient journey'?
  
- 2) Explore variation in:
  - a. The overall 'inpatient journey': This will be done using both interactive visualization (e.g. Sankey plots) and by taking snapshots (e.g. for the first 100, 1,000, 10,000 patients, and for every 10,000-patient increment in the global dataset).
  - b. Time from symptom onset to hospital admission
  - c. Time from hospital admission to ICU admission (if any)
  - d. Length of ICU stay
  - e. Time from hospital admission to oxygen supplementation (subdivided by modality)
  - f. Time on oxygen supplementation (subdivided by modality)
  - g. Length of hospital stay (from admission to outcome – death or discharge)
  
- 3) Investigate if such trends can be accounted for by other covariates (uni- and multivariable analysis):

- a. Location
    - i. NB: subgroup analyses will be explored for e.g. UK-only, high-income, non-high-income countries
  - b. Patient age
  - c. Patient sex
  - d. Presenting symptoms
  - e. Comorbidities
  - f. Differences in the frequency of particular medications and modalities of supportive care over time
  - g. Any identified individual circumstances within particular sites (for example resource rationing or a change in practice), in consultation with staff at those sites
- 4) Investigate and describe the overall distribution of these variables (2.b – 2.e) at fixed time-points (see 2.a) and over time. Understand what is driving their apparent overdispersion and identify optimal way of summarizing data.
  - 5) Produce (interactive) graphical outputs to visualise the data for these outputs.

## Exclusion criteria

- 1) Patients who have no positive confirmation of SARS-CoV-2 infection shall be excluded. We shall explore the consequences on requiring a positive laboratory confirmation.
- 2) Patients whose clinical journey records show clear inconsistencies that cannot be satisfactorily resolved shall be omitted.
- 3) Patients with an apparent hospital-acquired infection (SARS-CoV2 infection detected subsequent to hospital admission) will be excluded from analysis of time from onset to admission.

## Univariable analyses

- 1) Time from symptom onset to hospitalisation, by position in the ranking of patients by symptom onset
- 2) Proportion of patients with ICU admission, by position in the ranking of patients by admission date
- 3) Proportion of patients who received oxygen supplementation (subdivided by modality), by position in the ranking of patients by admission date
- 4) Time from hospitalisation to ICU admission, by position in the ranking of patients by admission date
- 5) Time from symptom onset to hospitalisation, by position in the ranking of patients by ICU admission
- 6) Time spent on oxygen supplementation (subdivided by modality)
- 7) Time from hospitalization to reaching an outcome (dead or discharged)
- 8) Investigation of appropriate probability distributions to model the individual elements of the journey

## Multivariable analyses

- 1) Survival analysis, in some cases with competing outcomes (death and discharge, or ICU admission and discharge - see [1]), for the time variables in univariate analyses 1, 4, 5, 6 and 7 adjusting for (potentially) age, sex, location, comorbidities, presenting signs/symptoms.
- 2) Logistic regression for the variables in univariate analyses 2 and 3, with the same adjustments

## Statistical Methods

- 1) T-tests
- 2) Linear regression
- 3) Survival analysis
- 4) Logistic regression  
Model-fitting of parametric distributions to the time periods making up the 'inpatient journey' (onset to admission, admission to ICU, ICU stay, length of oxygen supplementation).

## Figures

- Line plots for trajectories of time periods of interest over the course of the epidemic, potentially split by age/country/comorbidities/outcome.
- Cumulative risk/Kaplan-Meier plots for visualisation of survival analysis
- Sankey plots of the 'inpatient journey'; multiple "snapshots" for patients infected/admitted/ICU admitted at different points in the epidemic.

## Software

All analyses will be performed in R [2]. Interactive visualisations will be made using shiny [3].

## References

1. Wolkewitz, Martin, Ben S. Cooper, Marc J. M. Bonten, Adrian G. Barnett, and Martin Schumacher. 2014. "Interpreting and Comparing Risks in the Presence of Competing Events." *BMJ* 349 (August): g5060.
2. R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
3. Winston Chang, Joe Cheng, JJ Allaire, Yihui Xie and Jonathan McPherson (2020). shiny: Web Application Framework for R. R package version 1.4.0.2. URL <http://CRAN.R-project.org/package=shiny>