DATA NOTE

Sociodemographic and clinical characteristics of hospital admissions for COVID-19: A retrospective cohort of patients in two hospitals in the south of Brazil [version 1; peer review: 1 approved with reservations]

Edna Ribeiro de Jesus, Julia Estela Willrich Boell, Juliana Cristina Lessmann Reckziegel, Rafael Sittoni Vaz, Marco Aurélio Goulart, Flávia Marin Peluso, Tiago da Cruz Nogueira, Márcio Costa Silveira de Ávila, Michelle Mariah Malkiewicz, Catiele Raquel Schmidt, Vanessa Cruz Corrêa Weissenberg, Millena Maria Piccolin, Walmiro Martins Charão Junior, Elisiane Lorenzini

1 Federal University of Santa Catarina, Florianópolis, Santa Catarina, Brazil
2 Tereza Ramos General Hospital and Maternity, Lages, Santa Catarina, Brazil
3 Federal University of Planalto Catarinense, Lages, Santa Catarina, Brazil
4 Faculty of Santa Catarina, Florianópolis, Santa Catarina, Brazil
5 Maria Schmitt Institute, Florianópolis, Santa Catarina, Brazil

Abstract

Background: This database aims to present the sociodemographic and clinical profile of a cohort of 799 patients hospitalized with coronavirus disease 2019 (COVID-19) in two hospitals in southern Brazil.

Methods: Data were collected, retrospectively, from November 2020 to January 2021, from the medical records of all hospital admissions that occurred from 1 April 2020 to 31 December 2020. The analysis of these data can contribute to the definition of the clinical and sociodemographic profile of patients with COVID-19.

Data description: This dataset covers 799 patients hospitalized for COVID-19, characterized by the following sociodemographic variables: sex, age group, race, marital status and paid work. The sex variable was collected as sex assigned at birth from medical records data. Clinical variables included: admission to clinical ward, hospitalization in the Intensive Care Unit, COVID-19 diagnosis, number of times hospitalized due to COVID, hospitalization time in days and risk classification protocol. Other clinical variables include: pulmonary impairment; patients ventilation pattern; high-flow oxygen mask; pulmonary thromboembolism; cardiovascular disease; pulmonary.

**Conclusions**: The analysis of these data can contribute to the definition of the clinical and sociodemographic profile of patients with COVID-19. Thus, a great social impact is demonstrated when databases are published. Open data accelerates the research process, facilitates reuse and enriches datasets, in addition to optimizing the application of public resources, that is, enabling more use of the same investment.

**Keywords**
Corresponding author: Edna Ribeiro de Jesus (ednaenfa25@gmail.com)

Author roles: Jesus ERd: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; Boell JEW: Conceptualization, Formal Analysis, Methodology, Supervision, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; Reckziegel JCL: Project Administration, Supervision, Visualization, Writing – Review & Editing; Vaz RS: Investigation, Visualization; Goulart MA: Investigation, Visualization; Peluso FM: Investigation, Visualization; Nogueira TdC: Investigation, Visualization; Ávila MCsd: Investigation, Visualization; Malkiewiez MM: Visualization, Writing – Review & Editing; Schmidt CR: Visualization, Writing – Review & Editing; Weissenberg VCC: Investigation, Visualization; Piccolin MM: Investigation, Supervision, Visualization, Writing – Original Draft Preparation; Charão Junior WM: Project Administration, Visualization; Lorenzini E: Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Methodology, Project Administration, Resources, Software, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: The project is funded by the Foundation for Research Support of Santa Catarina (FAPESC). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2023 Jesus ERd et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Jesus ERd, Boell JEW, Reckziegel JCL et al. Sociodemographic and clinical characteristics of hospital admissions for COVID-19: A retrospective cohort of patients in two hospitals in the south of Brazil [version 1; peer review: 1 approved with reservations] F1000Research 2023, 12:627 https://doi.org/10.12688/f1000research.130532.1

First published: 07 Jun 2023, 12:627 https://doi.org/10.12688/f1000research.130532.1
Introduction

The coronavirus disease 2019 (COVID-19) pandemic has been considered the greatest challenge of the present time, associated with the unprecedented crisis in the health area, due to the expressive demand for hospital beds by patients with severe coronavirus conditions, which resulted in the collapse of health systems worldwide.\(^1\)–\(^3\) Patients affected by COVID-19 have shown clinical and sociodemographic variations, with a mortality rate around 2% in cases where there is massive alveolar damage and progressive respiratory failure.\(^4\)–\(^7\) Sex and gender variables also influenced COVID-19 epidemiology.\(^8\) Its lethality varies, above all, according to age group, clinical conditions and pre-existing comorbidities, such as arterial hypertension, diabetes, previous pulmonary disease, cardiovascular disease, cerebrovascular disease, immunosuppression and cancer.\(^9\)–\(^11\) Although there are disparities with regard to clinical variables and comorbidities associated with increased risk of hospitalization and mortality from COVID-19, growing evidence shows that patients with pre-existing diseases, and advanced age, are especially at risk of death due to viral infection.\(^12\)–\(^15\) Therefore, future analyses of this database can contribute to the analysis of characteristics of hospital admissions of patients affected by COVID-19. This database contains relevant information on the sociodemographic and clinical characteristics of patients hospitalized by COVID-19. The publication of the database promotes open science, the integrity and quality of scientific production and the reuse of data.

Methods

Ethical approval and consent to participate

This research was approved by the Research Ethics Committee of the Federal University of Santa Catarina (UFSC), (opinion No. 4.323.917/2020) Santa Catarina, Brazil. Patients provided written informed consent for data collection and publication.

Research design and method of data collection

This database comes from a cohort of patients who were admitted with a diagnosis of COVID-19 in two hospitals in southern Brazil. Retrospectively, from November 2020 to January 2021, data were collected from medical records of all hospital admissions that occurred from 1 April 2020 to 31 December 2020. Data related to the sex of the patients refer to the biological characteristics at birth from the patients’ medical records. All patients aged 18 years or older were included. This dataset covers 799 hospitalized patients.

Questionnaires hosted in the Survey Monkey platform were used, which contained questions about sociodemographic data, health conditions, and clinical, therapeutic, and outcome data. The variables considered for this study were: sex, age, age group, race, marital status, years of education, number of hospitalizations, hospitalization units, length of hospitalizations, risk classification, whether a COVID-19 test was taken, test used to detect COVID-19, respiratory compromise, ventilatory pattern, evolution, and previous diseases.

The inclusion criteria were: hospital admissions with a medical diagnosis of COVID-19; and being 18 years old or older. Individuals under 18 years of age and those who were not hospitalized due to COVID-19 were excluded.

Data description

Data were characterized by the following variables: sex, age group, race, marital status and paid work. The following clinical variables are included: admission to clinical ward, hospitalization in the Intensive Care Unit (ICU), COVID-19 diagnosis, number of times hospitalized by COVID, hospitalization time in days and risk classification protocol (green, yellow, red and not informed). Other clinical variables included: pulmonary impairment: <50%, between 50 and 75% or >75%; patients ventilation pattern (presented dyspnea with respiratory effort, dyspnea without effort or without dyspnea); high-flow oxygen mask; pulmonary thromboembolism (PE); cardiovascular disease; pulmonary sepsis; influenza exam results. Other health problems (if yes or no): diabetes, systemic arterial hypertension, chronic obstructive pulmonary disease (COPD), obesity, tabaco smoking, asthma, chronic kidney disease, overweight, vascular accident (Stroke), sedentary lifestyle, human immunodeficiency virus (HIV/AIDS), cancer, Alzheimer’s disease, Parkinson’s disease. The description of these characteristics is provided in Table 1. The analysis and reuse of sociodemographic and clinical profile data can be performed using descriptive statistics and measures of central tendency (mean and median) and

<table>
<thead>
<tr>
<th>Label</th>
<th>Name of data file/data set</th>
<th>File types (file extension)</th>
<th>Data repository and identifier (DOI or accession number)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data file 1</td>
<td>COVID-19 Hospital Admissions database</td>
<td>COVID-19 Hospital Admissions Database.xlsx</td>
<td><a href="https://doi.org/10.6084/m9.figshare.16746073.v4">https://doi.org/10.6084/m9.figshare.16746073.v4</a></td>
</tr>
</tbody>
</table>
variability (standard deviation and interquartile range), as well as absolute and relative distributions (n-%). The symmetry of the continuous distribution can be assessed using the Kolmogorov-Smirnov test. The predictive power of the variables can be analyzed using logistic regression. The opening of data from research projects is one of the most important elements of the research lifecycle for the success of Open Science. This is a sine qua non for reproducibility and scientific progress. Open Data speeds up the research process, facilitates reuse and enriches data sets, in addition to optimizing the application of public resources, in other words, enabling more use of the same investment. Opening data also allows detecting false, biased and inaccurate conclusions, as they are subject to replicability tests. Thus, great social impact is demonstrated when databases are published.16

Dataset validation

Limitations

This dataset is limited to a retrospective cohort of patients from two hospitals in southern Brazil. This can be considered a limitation. However, the data are very relevant, as there are few published studies and databases available on COVID-19 in Brazil. Researchers interested in the sociodemographic and clinical profile of patients hospitalized for COVID-19 can extensively explore the variables described here.

Ethical considerations

The present study was approved by the Research Ethics Committee of the Federal University of Santa Catarina (UFSC), (opinion No. 4.323.917/2020) Santa Catarina, Brazil. The basis and necessary information about the study objectives and method were given to all participants before the commencement of the study, and written informed consent was obtained from them. Participants consented to data publication. Participants were assured of the confidentiality of data and that only general statistics would be presented.

Authors’ contributions

ERJ, EL and JEWB made substantial contributions to the conception and design of the work, or the acquisition, analysis or interpretation of data, and to the writing of the work or critically reviewing important intellectual content. They have given final approval to the version and have agreed to be responsible for all aspects of the work, ensuring that issues relating to the accuracy or completeness of any part of the work are properly investigated and resolved. JCRL, FMP, RSV, MAG, TCN, MCSA, MMP, MMM, CRS, WMCI, VCCW, made substantial contributions to the acquisition of the data and the writing of the paper. They have approved the final version and agree to be responsible for all aspects of the work, ensuring that issues relating to the accuracy or completeness of any part of the work are properly investigated and resolved. All authors read and approved the final manuscript.

Data availability

Underlying data


Data are available under the terms of the Creative Commons Attribution 4.0 International license (CC-BY 4.0).

This is an open access database distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution and reproduction, provided the original work is properly cited.

Acknowledgments

The authors gratefully acknowledge the Foundation for Research Support of Santa Catarina (FAPESC).

References


Open Peer Review

Current Peer Review Status:  

Version 1

Reviewer Report 14 June 2023

https://doi.org/10.5256/f1000research.143299.r178045

© 2023 Defante Ferreto L. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Lirane Elize Defante Ferreto
Health Sciences Center, Postgraduate Program in Applied Health Sciences, Public Health Lab and Biosciences and Health Lab, Universidade Estadual do Oeste do Parana, Cascavel, State of Paraná, Brazil

The information available can be accessed by researchers who are interested in better understanding health policies, planning health actions, and estimating costs and demands in emergency situations. Database studies of this nature are important as they contribute to improving the effectiveness of health promotion and prevention actions. The results of data analysis provide information to managers who can assess the situation and identify critical points and propose strategies in situations that require immediate solutions.

Review the inclusion and exclusion criteria. The author informs, “The inclusion criteria were: hospital admissions with a medical diagnosis of COVID-19; and be 18 years or older. Individuals under 18 years of age and those who were not hospitalized due to COVID-19 were excluded”, so if being 18 years of age or older is an inclusion criterion and hospital admissions with a medical diagnosis of COVID-19, the exclusion criteria must be those within this population who were included, not being 18 years old and being hospitalized for COVID-19 was already a characteristic or circumstance that prevented the inclusion of the subject in the study.

In the variables, the inclusion criterion is having a diagnosis of COVID-19 and being hospitalized, the diagnostic variable for COVID-19 is not justified since everyone in the study has this outcome. Inform how the risk classification (green, yellow, red and not informed) was defined in the protocol in the study patients. Is it relevant that the author informs the distinction between obesity and overweight, would it not be possible to present this information in a single variable? It is not clear in the text how the authors defined a sedentary lifestyle. It is suggested that age in the bank be presented as a continuous variable, which allows the application of numerous statistical methods with better power for inference and modeling. As the authors inform that they collected data regarding income, the information in the database would be interesting.

As argued by the authors, “open data streamline the research process, facilitate reuse and enrich

---

Page 7 of 9
datasets, in addition to optimizing the application of public resources, that is, enabling greater use of the same investment”, it would be recommended that the authors inform the geographical location of the two hospitals in southern Brazil, as well as better describing the population from which the sample originated.

I understand that this information described by the author “Researchers interested in the sociodemographic and clinical profile of patients hospitalized for COVID-19 can widely explore the variables described here”, is not a limitation.

**Is the rationale for creating the dataset(s) clearly described?**
Yes

**Are the protocols appropriate and is the work technically sound?**
Partly

**Are sufficient details of methods and materials provided to allow replication by others?**
Partly

**Are the datasets clearly presented in a useable and accessible format?**
Yes

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Research area in infectious and contagious diseases with a project similar to the one presented with the collection of information from patients hospitalized with COVID-19.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.
The benefits of publishing with F1000Research:

• Your article is published within days, with no editorial bias
• You can publish traditional articles, null/negative results, case reports, data notes and more
• The peer review process is transparent and collaborative
• Your article is indexed in PubMed after passing peer review
• Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com