

## PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<http://bmjopen.bmj.com/site/about/resources/checklist.pdf>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

### ARTICLE DETAILS

<b>TITLE (PROVISIONAL)</b>	Characteristics, outcomes and risk factors for mortality of 522,167 patients hospitalized with COVID-19 in Brazil: a retrospective cohort study
<b>AUTHORS</b>	Castro, Marcia; Gurzenda, Susie; Macário, Eduardo; Araújo de França, Giovanni

### VERSION 1 – REVIEW

<b>REVIEWER</b>	Zaid Imam William Beaumont Hospital-Royal Oak United States of America
<b>REVIEW RETURNED</b>	19-Nov-2020

<b>GENERAL COMMENTS</b>	<p>We thank the authors for their submission that is certainly of significant interest during the COVID-19 pandemic. This represents the largest data cohort from Brazil which certainly helps to compare with how the pandemic is behaving compared to other countries and the impact of social determinants of health (SDH) and health inequalities in shaping patient outcomes. Well written manuscript, and appropriate analyses have been conducted. Only minor comments are:</p> <p>1-Would change pneumatopathy to chronic lung disease in the tables (labelled as one of the comorbidities).</p> <p>2-Given that the study evaluated patients in June 2020, likely a final outcome for these patients is now determined if the database is prospectively updated; it would be of interest to see if the patients labelled with no outcomes reported could have their outcomes evaluated and added to the analysis as these represent &gt;20,000 of the cohort. This should not significantly affect the results as strong trends with narrow CIs have been identified from the cohort but would allow for clearer and more updated data.</p>
-------------------------	--

<b>REVIEWER</b>	Sridhar Chilimuri Bronxcare Health System, Bronx USA
<b>REVIEW RETURNED</b>	20-Nov-2020

<b>GENERAL COMMENTS</b>	<p>The authors present an interesting descriptive study with an impressive number of parameters from a large study population, derived from a national database. The following are our recommendations that should be considered by the authors prior to publication of this manuscript:</p> <p>Major Recommendations:</p> <ul style="list-style-type: none"><li>• The paper contains multiple grammatical errors throughout the paper which will require changes and further proofreading.</li></ul>
-------------------------	---

	<ul style="list-style-type: none"> <li>• As the study derives its data from a national database, we recommend that the authors describe the database and its methods of data acquisition in detail. To our knowledge, there was a change of COVID-19 reporting systems in Brazil on 03/27. The REDCAP system was changed to the SIVE-Gripe database in late March 2020. (de Souza, W.M., Buss, L.F., Candido, D.d.S. et al. Epidemiological and clinical characteristics of the COVID-19 epidemic in Brazil. Nat Hum Behav 4, 856–865 (2020))</li> <li>• Line 114: We recommend that they explain the rationale behind their technique of dealing with missing data prior to statistical analysis. We refer to de Souza et al where large proportions of comorbidity data were reported as missing in the SIVE-Gripe database</li> </ul> <p>Minor Recommendations:</p> <ul style="list-style-type: none"> <li>• Line 100: The authors should define duplicate records, and the rationale for their removal. Our concern is that duplicate records could be readmissions. Deleting them would skew results. If they were deemed to be clerical errors it would be reasonable to delete them.</li> <li>• Line 86, 99: The authors should define SARS and COVID-19 clearly to avoid any confusion to readers. We believe the authors were referring to SARI (severe acute respiratory illness) instead of SARS. As per CDC: Severe acute respiratory syndrome (SARS) is a viral respiratory illness caused by a coronavirus, called SARS-associated coronavirus (SARS-CoV)</li> </ul> <p>Finally, since this for a international audience, we recommend that the authors elaborate the differences in outcomes based on different regions. They do address it but perhaps some elaboration is needed for great clarity.</p>
--	--

## VERSION 1 – AUTHOR RESPONSE

### **Reviewer: 1**

Reviewer Name

Zaid Imam

#### Comments to the Author

We thank the authors for their submission that is certainly of significant interest during the COVID-19 pandemic. This represents the largest data cohort from Brazil which certainly helps to compare with how the pandemic is behaving compared to other countries and the impact of social determinants of health (SDH) and health inequalities in shaping patient outcomes. Well written manuscript, and appropriate analyses have been conducted.

Thank you very much!

Only minor comments are:

1-Would change pneumatopathy to chronic lung disease in the tables (labelled as one of the comorbidities).

Thank you for the suggestion. Completed as requested.

2-Given that the study evaluated patients in June 2020, likely a final outcome for these patients is now determined if the database is prospectively updated; it would be of interest to see if the patients labelled with no outcomes reported could have their outcomes evaluated and added to the analysis as these represent >20,000 of the cohort. This should not significantly affect the results as strong trends with narrow CIs have been identified from the cohort but would allow for clearer and more updated data.

Thank you. The review did take a long time, so we do agree that an update was in order. We used a dataset with data until Dec 14 (N=522,167) and updated all analysis, tables, and graphs. While the number of patients with unknown clinical outcome increased (from 4,688 to 11,126), the proportion decreased (from 4.1% to 2.1%). Similarly, the number of patients still in hospital increased from 24,223 to 53,503, however the proportion decreased by half (from 21.1% to 10.2%).

**Reviewer: 2**

Reviewer Name  
Sridhar Chilimuri

Comments to the Author

The authors present an interesting descriptive study with an impressive number of parameters from a large study population, derived from a national database. The following are our recommendations that should be considered by the authors prior to publication of this manuscript:

Major Recommendations:

- The paper contains multiple grammatical errors throughout the paper which will require changes and further proofreading.

As a member of editorial boards and as a reviewer, I never accept/make such comments without providing clear examples. I urge the editor and the reviewer to consider a similar approach. It is respectful and collegial.

We did full proofreading before and did it again in this revised version. If anything missed our eyes, we will certainly welcome constructive suggestions.

- As the study derives its data from a national database, we recommend that the authors describe the database and its methods of data acquisition in detail. To our knowledge, there was a change of COVID-19 reporting systems in Brazil on 03/27. The REDCAP system was changed to the SIVE-Gripe database in late March 2020. (de Souza, W.M., Buss, L.F., Candido, D.d.S. et al. Epidemiological and clinical characteristics of the COVID-19 epidemic in Brazil. *Nat Hum Behav* 4, 856–865 (2020))

Thank you. The data source is described in detail under the Methods/Data source section. We have added text to clarify that the SIVEP-Gripe database is publicly available and provided the website address in the manuscript. Indeed the initial REDCAP system could not support the volume of records. SIVEP-Gripe already

existed, and the migration of records was done without problems. The main issues were (and still are) on the reporting of all cases. Data are available, but not with all epidemiological information it could have. However, the quality of the hospitalization data (which we use in the paper) is very good. That is exactly why any paper analyzing mortality in detail uses SIVEP-Gripe. We assure you this is the best data available. Indeed, two of the authors of the paper are from the Ministry of Health in Brazil, responsible for the management of the dataset.

- Line 114: We recommend that they explain the rationale behind their technique of dealing with missing data prior to statistical analysis. We refer to de Souza et al where large proportions of comorbidity data were reported as missing in the SIVE-Gripe database

Thank you. We added an online supplement with a table reporting the missingness in all variables listed in our tables. In the case of the variables that are critical for the analysis, such as ethnoracial self-classification, we added a “not reported” category, a procedure commonly used in statistics.

#### Minor Recommendations:

- Line 100: The authors should define duplicate records, and the rationale for their removal. Our concern is that duplicate records could be readmissions. Deleting them would skew results. If they were deemed to be clerical errors it would be reasonable to delete them.

Thank you. The Ministry releases a new dataset every two weeks (most often). As new data are added, there is a routine to remove duplicate records. The variables used to detect and verify duplicate records are: name of the patient, name of the mother of the patient, date of birth of the patient, and date of first symptoms. Therefore, those that had another infection would have a different date of first symptoms. Thus, duplicates are not readmissions, but records that have exactly the same information, that were entered twice due to problems in hospital data entry. This can happen in other administrative hospital systems as well, and the checking procedure is always necessary.

- Line 86, 99: The authors should define SARS and COVID-19 clearly to avoid any confusion to readers. We believe the authors were referring to SARI (severe acute respiratory illness) instead of SARS. As per CDC: Severe acute respiratory syndrome (SARS) is a viral respiratory illness caused by a coronavirus, called SARS-associated coronavirus (SARS-CoV).

Thank you. We changed as requested.

Finally, since this is for an international audience, we recommend that the authors elaborate the differences in outcomes based on different regions. They do address it but perhaps some elaboration is needed for great clarity.

Lines 242-266 are focused on putting our results in an international context. Since the main focus of the paper is Brazil, we did not have to do that, as is the case of other

studies published in journals of wide audience (such as The Lancet and JAMA) for China, Germany, and Italy, just to mention a few examples. We wanted to compare to other countries as we consider this of utmost importance, and we compared based on what is available and possible to compare. If the reviewer is referring to regions within Brazil, lines 266-275 discuss the context of the inequalities and differences among the five regions, specifically identifying the North and Northeast as having the worst health. We are not certain what additional comparisons the reviewer would like to see.