

1 **Cohort Study Protocol of the Brazilian Collaborative Research Network on**
2 **COVID-19: strengthening WHO global data**

3 **Authors:**

4 Fernando Anschau¹, Natália Del' Angelo Aredes², Ludovic Reveiz³; Monica Padilla³;
5 Rosane de Mendonça Gomes⁴; Wellington Mendes Carvalho⁴; Fernando Antonio
6 Gomes Leles⁴; Fernanda Baeumle Reese⁵; André Hostílio Hubert⁵; Elisandrea
7 Sguario Kemper⁶; Renilson Rehem de Souza⁶; Cristiane Feitosa Salviano⁶; Hevelin
8 Silveira e Silva⁶; Eduardo Barbosa Coelho⁷; Giuseppe Cesare Gatto⁷; Rafael Freitas
9 de Moraes⁷; Leonardo Nunes Alegre⁷, Rodrigo Citton Padilha dos Reis⁸, Joaquim
10 Francisco dos Santos Neto⁸, Andresa Fontoura Garbini¹; Cesar Perdomo Purper¹,
11 Veridiana Baldon dos Santos¹; Rafaela dos Santos Charão de Almeida¹; Bruna
12 Donida¹; Rogério Farias Bitencourt¹; Luciane Kopittke¹; Fernanda Costa dos
13 Santos¹; Raquel Lutkmeier¹; Daniela dos Reis Carazai¹; Virgínia Angélica Silveira
14 Reis⁹; Flávio Clemente Deulefeu⁹; Fernanda Gadelha Severino⁹; José Gustavo da
15 Costa Neto⁹; Nirvania do Vale Carvalho¹⁰, André Jamson Rocha de Andrade¹⁰;
16 Adriana Melo Teixeira¹¹; Olavo Braga Neto¹¹, Gabriel Cardozo Muller⁸, Ricardo de
17 Souza Kuchenbecker⁸

18
19 ¹ Hospital Nossa Senhora da Conceição – Grupo Hospitalar Conceição

20 ² Nursing School, Universidade Federal de Goiás, member of the EBSEPH research
21 team

22 ³ Pan American Health Organization (PAHO) - World Health Organization
23 (WHO)/Washington, DC, US

24 ⁴ PAHO/WHO in Brazil

25 ⁵ Complexo Hospitalar do Trabalhador

26 ⁶ Hospital da Criança de Brasília

27 ⁷ Brazilian Company for Hospital Services (Empresa Brasileira de Serviços
28 Hospitalares - EBSEPH)

29 ⁸ Graduate Program in Epidemiology UFRGS, Hospital de Clínicas de Porto Alegre

30 ⁹ Instituto de Saúde e Gestão Hospitalar

31 ¹⁰ Hospital Getúlio Vargas- Teresina, Piauí - Health Department of the State of Piauí

32 ¹¹ Department of Hospital, Home, and Emergency Care (DAHU) of the Specialized
33 Health Care Office (SAES) of the Brazilian Ministry of Health

1 **Cohort Study Protocol of the Brazilian Collaborative Research Network on COVID-**
2 **19: strengthening WHO global data**

3 **Corresponding author:**

4

5 Fernando Anschau

6 Av. Francisco Trein n.:326

7 Bairro Cristo Redentor – Porto Alegre, Rio Grande do Sul,

8 Brazil Zip Code: 91350-200

9 e-mail: afernando@ghc.com.br

10 telephone number: +55 (51) 3357 2805

11

12 **Word Count: 4975**

13

14

15

1 **Cohort Study Protocol of the Brazilian Collaborative Research Network on**
2 **COVID-19: strengthening WHO global data**

3
4 **Article Summary**

5 This is a multicenter project among Brazilian hospitals to provide data on
6 COVID-19 through the WHO global platform.

7 It is expected to deepen knowledge about the pandemic scenario and help
8 hospital institutions to develop preventive measures, health service protocols and
9 strengthen the training of teams in the existing complications.

10

1 **Cohort Study Protocol of the Brazilian Collaborative Research Network on COVID-** 2 **19: strengthening WHO global data**

3 **Abstract**

4 **Introduction:** With the COVID-19 pandemic, hospitals in low-income countries were
5 faced with a triple challenge. First, a large number of patients required hospitalization
6 because of the infection's more severe symptoms. Second, there was a lack of
7 systematic and broad testing policies for early identification of cases. Third, there
8 were weaknesses in the integration of information systems, which led to the need to
9 search for available information from the hospital information systems. Accordingly, it
10 is also important to state that relevant aspects of COVID-19's natural history had not
11 yet been fully clarified. The aim of this research protocol is to present the strategies
12 of a Brazilian network of hospitals to perform systematized data collection on
13 COVID-19 through the World Health Organization (WHO) Platform. **Methods and**
14 **Analysis:** This is a multicenter project among Brazilian hospitals to provide data on
15 COVID-19 through the WHO global platform, which integrates patient care
16 information from different countries. From October 2020 to March 2021, a committee
17 worked on defining a flowchart for this platform, specifying the variables of interest,
18 data extraction standardization and analysis. **Ethics and Dissemination:** This
19 protocol was approved by the Research Ethics Committee (CEP) of the Research
20 Coordinating Center of Brazil (CEP of the Hospital Nossa Senhora da Conceição),
21 on January 29, 2021, under approval No. 4.515.519 and by the National Research
22 Ethics Commission (CONEP), on February 5, 2021, under approval No. 4.526.456.
23 The project results will be explained in WHO reports and published in international
24 peer-reviewed journals, and summaries will be provided to the funders of the study..

25 **Keywords:** COVID-19, clinical data, morbidity, mortality

26
27 **Strengths and limitations of this study:** As the study involves a convenience and
28 non-probabilistic sample of patients hospitalized in health units, it may not represent the
29 population of patients with COVID-19 hospitalized in the country. However, the
30 information generated by this research can serve as a basis for the development of
31 maps of the evolution of SARS-CoV-2 infection and public policies to face pandemics. It
32 is a study that uses secondary data, and therefore, information bias may occur, but on
33 the other hand, it has a low cost and facilitates a population-based study with national
34 coverage.

1 Cohort Study Protocol of the Brazilian Collaborative Research Network on 2 COVID-19: strengthening WHO global data

3 4 Introduction

5 COVID-19 is a disease caused by SARS-CoV-2, which started in Wuhan
6 (Hubei province, China) in December 2019 (1), and it was characterized as a
7 pandemic by the World Health Organization (WHO) on March 11, 2020 (2). The
8 disease caused by SARS-CoV-2 has a lower lethality when compared to that of two
9 other coronaviruses that have already caused epidemics, namely SARS-CoV and
10 MERS-CoV. However, its high transmission rate places a great demand on health
11 systems, with an increase in the number of hospitalizations, need for oxygen by
12 assisted ventilation, and other forms of clinical support (3-5).

13 Brazil ranks third worldwide in the total number of COVID-19 cases, with
14 all regions affected (6,7), and demographic and social inequalities impacted the
15 results obtained in transmission control and health care. Low- and middle-income
16 countries, such as Brazil, had great difficulty adopting pandemic mitigating measures
17 such as non-pharmacological interventions (8). Along with the high transmissibility of
18 variants of interest, so-called *variants of concern* (VOC), the pandemic's increased
19 extent was understandable (9,10).

20 Brazil is a country of continental dimensions, with 6,820 hospitals, of which
21 2,423 are public hospitals and 4,397 are private hospitals (11). The country has
22 536,474 hospital beds (1.96 and 0.8 beds per 1,000 inhabitants in private and public
23 hospitals, respectively) (11,12). There are disparities in the availability of hospital
24 beds in the different regions, including those in the intensive care unit (ICU), which is
25 of utmost importance for the care of COVID-19 patients. For every 100,000
26 inhabitants, there are 9.73 ICU beds in the Unified Health System (Sistema Único de
27 Saúde - SUS) in the North region, 10.46 in the Northeast region, 11.87 in the
28 Central-West region, 13.24 in the Southeast region, and 15.88 in the South region,
29 according to data from 2020 (13).

1 Between December 2019 and April 2020, ICU beds were expanded from
2 46,045 to 60,265 (13) nationwide with an overall increase of 30.88%, with the
3 increase in ICU beds under SUS being almost four times smaller than the increase in
4 ICU beds in the private network (13.46% increase under SUS and 48.33% in the
5 private sector), signaling a need for public investment, especially in view of the gaps
6 identified during the pandemic.

7 A study that evaluated the first 250,000 SARS-Cov-2 hospitalizations in Brazil
8 from the national Severe Acute Respiratory Syndrome case database (SIVEP-Gripe)
9 highlighted the in-hospital mortality rate, which reached 38%, the number of people
10 with demands compatible with the ICU (39%) and regional discrepancies, revealing
11 considerable challenges in managing the pandemic in Brazil (14). However, the
12 results showed a high proportion of missing data for some variables, especially
13 comorbidities and symptoms, and non-standardized reporting protocols (14).
14 Additionally, these data did not include the so-called "second wave", caused
15 predominantly in Brazil by new highly transmissible variants such as P1 and B.1.1.7
16 (15-17).

17 Data from the two waves observed in Brazil using the same dataset (SIVEP-
18 Gripe) were published in 2021. The cross-sectional study of hospitalized patients
19 with RT-PCR positive for SARS-CoV-2 in Brazil comprised data entered into SIVEP-
20 Gripe from February 25, 2020 to April 30, 2021, separated into two waves on
21 November 5, 2020. As opposed to what was seen in the rest of the world, in-hospital
22 mortality in Brazil increased from 34.81% in the first wave to 39.30% in the second
23 wave. In the second wave, there were more ICU admissions, use of non-invasive
24 and invasive ventilation, and increased mortality in the younger age groups. The
25 Brazilian South and Southeast regions had the highest hospitalization rates per
26 100,000 inhabitants. Nevertheless, the in-hospital mortality rate was higher in the
27 northern and northeastern states. Racial differences were observed in clinical

1 outcomes, with white people being the most hospitalized, though black/brown-
2 skinned people had higher mortality rates. The younger age groups displayed more
3 considerable differences in mortality compared to the groups with and without
4 comorbidities in both waves. The results reflect the inequalities observed in Brazil
5 and the overload of health systems upon the second wave's arrival (18).

6 It is critical to maintain surveillance on the pandemic's evolution over time and
7 across the country to measure its impact on the health care system. In Brazil, SUS
8 consists of a public network of services at all levels of care, on which about 75% of
9 the population depends exclusively. Today, 66.34% of hospital beds belong to SUS
10 (19), which characterizes the Brazilian public system as responsible for most of the
11 health care during the COVID-19 pandemic in the country. Even so, both public and
12 private hospitals do not use a unified electronic record system, which increases the
13 heterogeneity of reports and represents an obstacle for data collection in research.

14 As a strategy to overcome the challenges of data collection for health
15 research and management purposes, the Pan American Health Organization
16 (PAHO)-WHO proposed a collaboration in 2020 between hospitals in all regions of
17 Brazil to join databases and collect clinical data from COVID-19 patients. Through
18 this initiative, it was possible to gather data from different locations around the
19 country, and the world, to characterize the disease's natural history, identify risk
20 factors for severe cases and describe treatment interventions and outcomes. Thus, it
21 directly contributes to national and global public health with scientific evidence on
22 COVID-19 clinical management and the strengthening of research networks and
23 technology-supported data collection.

24 As far as we know, this is the first initiative in Brazil to create a "Big Data"
25 database in health care from hospitals and, afterwards, a *data hub* with clinical
26 conditions and usage patterns from the health care system.

27 The aim of this study is to provide standardized quality clinical data on the
28 COVID-19 pandemic in Brazil. Furthermore, it aims to describe the resources and

1 methods used in this collaboration to identify and validate the cases in each
2 database, with patients confirmed as infected with SARS-CoV-2 by laboratory tests
3 (case group) and suspects (control group).

4

5 **Methods**

6 This study is being conducted in 43 hospitals located in all regions of Brazil.
7 From March 2020 to June 2022, a committee of experts worked on defining a
8 blueprint for this platform, specifying variables of interest, data extraction
9 standardization and data analysis.

10 Aiming to standardize data collection, the project uses *case report forms*
11 (CRF) to extract clinical information from adults and children hospitalized with
12 COVID-19 (accessed from the platform URL: [https://www.who.int/teams/health-care-](https://www.who.int/teams/health-care-prontidão-clínica-unidade/COVID-19/data-platform)
13 [prontidão-clínica-unidade/COVID-19/data platform](https://www.who.int/teams/health-care-prontidão-clínica-unidade/COVID-19/data-platform)). The data from all the centers
14 surveyed can be gathered and analyzed - either together or by center or region of
15 the country - to provide a better understanding of the disease and determine the
16 public health response to the pandemic. Collecting anonymized clinical data globally
17 will provide WHO the following:

- 18 1. Identification of the main clinical characteristics and prognostic factors of
19 hospitalization cases for suspected or confirmed COVID-19, expanding our
20 knowledge of the severity, spectrum and impact of the disease on the
21 hospitalized population globally and in different countries.
- 22 2. Identification of clinical interventions, facilitating global and country operational
23 planning during the COVID-19 pandemic.

24 The research questions and hypotheses of the Brazilian protocol of the Global
25 Clinical Platform on COVID-19 are presented in Table 1 below.

26
27

Table 1 - Project research questions and hypotheses	
Questions	Hypotheses
Are the descriptions presented in the international literature regarding infections by the SARS-COV-2 applicable to the Brazilian reality?	The descriptions presented in the existing academic literature on the clinical, laboratory, radiological and therapeutic profile of patients diagnosed with the SARS-COV-2 and admitted to Brazilian hospitals are different, a factor that makes it essential to consider Brazilian peculiarities during this disease's screening, diagnosis and treatment.
What are the main symptoms presented by Brazilian patients with COVID-19, especially those admitted to different hospital units in Brazil?	The most common clinical manifestations in Brazilian patients tested for SARS-COV-2 are related to fever, myalgia, respiratory symptoms, and headache.
Is in-hospital mortality from COVID-19 similar in Brazil and the rest of the world?	The in-hospital mortality profile of patients admitted to Brazilian hospitals with COVID-19 is not similar to that presented in the literature.

1 The primary and secondary objectives of the Brazilian protocol of the Global
2 Clinical Platform on COVID-19 are as follows:

- 3 ● Primary objective: to determine the clinical, laboratory and radiological profile,
4 therapy and mortality of patients confirmed having COVID-19 infection and
5 admitted to Brazilian hospitals.
- 6 ● Secondary objectives: 1) to identify outcomes such as ICU admission, length
7 of hospital stay and duration of mechanical ventilation and of respiratory
8 therapy; 2) to assess the incidence of adult respiratory dysfunction syndrome
9 and need for mechanical ventilation and intensive care, and days of intensive
10 care in patients admitted to hospitals; and 3) to identify the incidence of signs
11 and symptoms of multisystemic inflammatory syndrome and associated
12 chronic diseases.

13 **Study design, population and sample**

14 This is a multicenter cohort study with convenience sampling from the
15 population of patients hospitalized with COVID-19, because of either clinical
16 suspicion or laboratory confirmation. As inclusion criteria for cases, patients with
17 confirmed COVID-19 (RT-PCR tests or rapid tests), admitted to hospitals with a
18 minimum length of stay of 24 hours, are considered. Patients admitted to one of the
19 hospital units who tested positive for COVID-19 in the period from March 2020 to
20 June 2022 will be entered in the study retrospectively.

21 **Main outcomes and parameters**

22 The CRFs were developed by WHO for the collection and submission of
23 harmonized data to the COVID-19 Global Clinical Platform, and the procedures for
24 further data collection were performed by the committee's Technical Team.

25 Key descriptive parameters include socio-demographic data (age and sex),
26 presence of preexisting conditions, use of chronic medications, clinical
27 characteristics on admission and during hospitalization, laboratory findings on

1 admission and during hospitalization, clinical interventions on admission and during
2 hospitalization (use of oxygen, use of mechanical ventilation, use of therapeutics),
3 and patient outcomes (death, discharge and referral).

4 The patient outcomes described above will be used in secondary analysis to
5 investigate associations between baseline characteristics and disease severity and
6 outcomes.

7 **Statistical considerations**

8 **General considerations**

9 Statistical analyses (numerical summaries, point estimates, confidence
10 intervals, hypothesis tests, fitting regression models, tables, graphs and maps) will
11 be performed in R³ version 4.0.5 or higher. Specification of sets of adjustment
12 variables in certain association analyses will be conducted in DAGitty⁴ software.

13 **Missing data**

14 For each analysis, the denominator will represent the number of available
15 observations.

16 The date of hospital admission will be considered essential and critical to
17 know the hospitalized population to be analyzed. Records with missing admission
18 dates will be excluded from the analysis unless there is confirmation from
19 participating hospitals that the data refer to a cohort of hospitalized patients.

20 Data imputation will not be performed but can be considered for each
21 particular analysis.

22 **Descriptive statistics**

23 According to the type of variable, the descriptive statistics will be as follows:

- 24 • Quantitative variables: number of observations (N), mean, standard deviation
25 (SD), median and interquartile range (IQR), depending on the distribution of
26 the variable.

- 1 • Qualitative variables: number of observations (N), absolute frequency (n), and
2 relative frequency (%).
- 3 • Percentages will be calculated for the number of participants with observed
4 data (n/N).

5 **Objectives of the analysis** (See Supplementary 1)

6 1. Description of socio-demographic and clinical characteristics

7 To describe the socio-demographic and clinical characteristics of hospital
8 admissions, interventions and clinical outcomes among:

- 9 • people infected with the SARS-CoV-2 virus cared for by hospitals within the
10 Global Platform;
- 11 • specific population groups, such as children, pregnant women and
12 populations with co-infections or comorbidities, among others.

13 2. Characterization of access profiles and hospital resource utilization patterns

14 To describe access profiles and hospital resource utilization patterns such as
15 emergency care, inpatient admissions in wards and ICU, including non-invasive and
16 invasive ventilation devices and non-traditional care modalities such as intermediate
17 units, ventilatory support and others.

18 3. Association between clinical/socio-demographic characteristics and outcomes To I

19 To investigate factors associated with disease severity, ICU admission and
20 in-hospital mortality among the following:

- 21 • people infected with the SARS-CoV-2 virus cared for by hospitals within the
22 Global Platform;
- 23 • specific population groups, as described above.

24 4. Sources of inequities

25 To investigate potential sources of inequities in timely access to diagnostic and
26 therapeutic services and resources and the association with clinical outcomes.

1 **Data governance**

2 Defining data governance is always necessary in a research protocol,
3 considering that data is protected by legislation, including the rights of research
4 participants. This model dictates that the hospitals involved understand the research
5 aspects, the information security points, and complete the process of sending data to
6 the central repository in a clear, unambiguous manner. Furthermore, regarding
7 information security, the central repository's teams are required to follow procedures
8 ensuring ethical and secure actions, in compliance with the legislation in force, when
9 processing information.

10 Privacy and Consent Policies are important points for discussion on health
11 data sharing. Therefore, for the preparation of this proposal, the concepts and
12 technical principles of the General Data Protection Act (LGPD) were respected,
13 considering health data as sensitive, as described in these documents (20). The
14 LGPD regulates the processing of personal data and aims to protect citizens' right to
15 freedom, privacy and free development.

16 Considering all these aspects, three levels of governance/access were
17 defined: strategic, tactical and operational:

- 18 1. The Advisory Committee of the PAHO/WHO COVID-19 Platform Project
19 represents the first, whose main objective is to technically support PAHO/WHO in the
20 terms of the COVID-19 Global Clinical Data Platform, and the participating
21 institutions in the activities of preparation for analysis and publication of clinical and
22 epidemiological data;
- 23 2. The second consists of those responsible for translating the strategy and
24 integrating all parties, setting the objectives and managing the project, including
25 defining the work processes, schedule, team engagement and communication,
26 quality, and results follow-up;
- 27 3. The third performs activities from the tactical level always in compliance with data

1 security standards and rules. It is at this level that statements like "data must be
2 stored securely" translate into technological tools ensuring the semantics read into
3 the sentence.

4 **Risks**

5 Risks are related to the loss of confidentiality. However, they have been and
6 will be constantly mitigated/controlled by the anonymization of the participants'
7 identification data. The researchers take responsibility for maintaining their
8 commitment to preserve data security and confidentiality, and hospital institutions
9 are asked to take all necessary steps to protect their access credentials and
10 passwords to the COVID-19 Data Platform.

11 The collected and anonymized data will be stored by the aforementioned
12 WHO Platform, a secure, password-protected, restricted-access electronic platform.
13 WHO commits to the following: i) to protect the confidentiality of anonymous data on
14 COVID-19 and to prevent its unauthorized disclosure; ii) to implement and maintain
15 technical and organizational security measures to protect the security of anonymous
16 data on the Platform; and iii) to follow the International Health Regulations (21).

17 **Benefits**

18 It is expected that this protocol will produce specific and reliable knowledge of
19 the COVID-19 pandemic - clinical and epidemiological analysis, leading to
20 discussions on Brazil and other countries and their reality and plans to fight the
21 pandemic. It is expected that it will deepen our knowledge of the pandemic scenario
22 and help hospital institutions to develop preventive measures and health service
23 protocols and to strengthen staff training in existing complications. The information
24 generated by the research can serve as a basis for developing maps of the SARS-
25 CoV-2 infection evolution.

26 **Clinical data extraction**

27 As the WHO representative in the country, the PAHO/WHO office in Brazil,

1 through the Technical Unit of Health Systems and Services, coordinated the
2 project by dialoguing with the partner institutions to make this project effective. In
3 this context, Grupo Hospitalar Conceição (GHC) became the national coordinating
4 Brazilian institution for submission procedures to the Research Ethics Committee
5 (CEP) and the National Research Ethics Commission (CONEP).

6 Data collection started retrospectively in 43 Brazilian hospitals - as shown in
7 Figure 1 - starting from the first COVID-19 case diagnosed in Brazil (March 2020)
8 and ending in June 2022. We chose to carry out this first stage of data collection
9 directly in the hospital units, consulting the databases of the clinical and care
10 information systems available in each institution or the physical records of the
11 hospitalizations of patients with suspected or confirmed COVID-19 in these units.
12 Figure 2 shows the flowchart of the study protocol.

13 The selected hospitals are responsible for providing all the conditions for the
14 project's execution, such as access to data collection and feeding the platform to
15 the PAHO/WHO consultant assigned to each institution. Accordingly, each hospital
16 signed a commitment form that contains all the guarantees offered by the WHO to
17 maintain data confidentiality and anonymization. The collection and access
18 process, in agreement with each hospital unit, respects the specificities and
19 internal regulations, and also follows the ethical and data security norms.

20 Considering that the information contained in the hospital records of each
21 institution can be divided between structured and open text, and in view of the
22 heterogeneities of Brazilian hospitals in terms of management, there is a need to
23 organize data collection in several ways: one through manual consultation of
24 electronic citizen records (ECR) - or physical records for hospitals that do not have
25 ECRs - and filling out the WHO-specific forms; another through the development of
26 an automated programming interface (API) for filling out these forms.

27 The standardized API was developed by information technology (IT)
28 professionals to share databases from different hospitals. These databases, with

1 provision for secure cloud-based storage, contain de-identified electronic health
2 records (EHRs) corresponding to admission notes for the first 48 hours of care and
3 the 24 hours prior to discharge/outcome for each research participant.

4 The structured ECR data contains demographic information, prescriptions,
5 laboratory results, vital signs and general patient characterization such as height and
6 weight and admission and outcome dates. Since the lack of data for different
7 variables can be significant, there is the possibility for institutions to search and fill in
8 part of the variables from unstructured data by open text extraction.

9 For this procedure, the Smart Health Connect (SHC) software is used, which
10 operates on the basis of an algorithm using Deep Neural Networks, extracting
11 information from the clinical evolutions and transferring to the CRF data that were
12 previously missing, reinforcing the completeness of the database exported to the
13 WHO. The main variables of interest incorporated through this software were
14 comorbidities, symptoms and clinical events, especially those related to COVID-19,
15 such as oxygen support and ventilation type, and outcomes. The software is capable
16 of recognizing and categorizing data regarding history (past), current events
17 (present) and expected actions or outcomes (future) and identifying variables related
18 to drug treatment and clinical condition, including vital signs, symptoms, clinical
19 devices, affections/pathologies, and other relevant data that could not be captured in
20 the structured ECR data fields.

21 SHC uses the newest natural language processing techniques available to the
22 scientific community and a software architecture model that enables mass data
23 extraction.

24 Regarding the technological standard used by the software, there are two
25 approaches supported by artificial intelligence. The first is Named Entity Recognition
26 (NER), which has the function of locating previously defined entities in open text. For
27 this project, the categories are disease names, signs and symptoms, procedures,

1 medications, ventilatory support and test results. SHC identifies and marks in the text
2 each of the identified categories in a process known as inference.

3 Once the relevant terms are identified, the other step is the analysis of the
4 findings in a given text, called Assertion Detection. Through this approach it is
5 possible to differentiate the status of what has or has not actually occurred with that
6 patient. This is necessary because the medical records can indicate something
7 related to the past, or to a third party, such as family history, or even state that the
8 patient does not have a certain condition or does not use a certain medication.

9 **Expected results**

10 At the end of the project, in addition to the reports produced by WHO with
11 the data provided globally and the ensured access to the databases, PAHO/WHO
12 in Brazil will organize publications in partnership with the institutions participating in
13 the project in the country, highlighting the results and Brazil's contribution in
14 fighting the COVID-19 pandemic.

15 Based on this protocol design, it will be possible to continue data collection
16 among the participating hospitals and continuously add valuable information for
17 public health in the context of facing the pandemic in Brazil. Once collaboration
18 between institutions has taken place and the form of data collection, variables of
19 interest and technologies to support the search for information of interest have
20 been standardized, there is significant potential to advance the analysis of
21 evidence and response measures in the national health system.

22 Until now, the identification of this strategy's partner hospitals and their
23 characterization, with number and type of beds, provide an analysis perspective of
24 the health care network scenario in the fight against the pandemic (Table 2).

25 **Ethics**

26 This protocol was approved by the Research Ethics Committee (CEP) of the
27 Research Coordinating Center of Brazil (CEP of the Hospital Nossa Senhora da

1 Conceição), on January 29, 2021, under approval No. 4.515.519 and by the National
2 Research Ethics Commission (CONEP), on February 5, 2021, under approval No.
3 4.526.456. Afterwards, it received the approval of the local CEPs of each institution
4 involved. In addition, this study follows the resolutions 466/12 and 510/16 and the
5 resolutions CNS No. 466/2012 and No. 510/2016 of the National Health Council,
6 which deal with the ethical aspects involving the use of data directly obtained from
7 participants or identifiable information, or that may entail risks greater than those
8 existing in everyday life. The results of this study will be published in international
9 peer-reviewed journals and summaries will be provided to the funders of the study.

10 **Strengths and Limitations**

11 Through this initiative it is possible to gather data from different locations
12 around the country to characterize the disease's natural history, identify risk factors
13 for severe cases and describe treatment interventions and outcomes. Thus, it
14 directly contributes to national and global public health with scientific evidence on
15 COVID-19 clinical management and to the strengthening of research networks and
16 technology-supported data collection.

17 As far as we know, this is the first initiative in Brazil to create a "Big Data"
18 database in health care from hospitals and, afterwards, a *data hub* with clinical
19 conditions and usage patterns from the health care system.

20 Because this is a convenience, non-probabilistic sample of patients admitted
21 to health care facilities, it may not represent the population of COVID-19 patients
22 hospitalized in all the country. However, the information generated by this research
23 can serve as a basis for the development of maps of the evolution of SARS-CoV-2
24 infection and public policies to face pandemics.

25 **Patient and public involvement**

26 This project is already involving both Brazilian communities and public agencies
27 that are facing the fight against the COVID-19 pandemic, as well as the Brazilian

1 Ministry of Health and PAHO/WHO. We intend to expand partnerships and involve
2 more stakeholders from Brazil.

3 **Discussion**

4 One of the main concerns regarding the COVID-19 pandemic is its burden on
5 the organization of the health care system (22). The need for a general and effective
6 organization of health services to treat the population safely and fight the disease's
7 complications has required a great deal of effort. This includes reallocating health
8 care workers, using appropriate medicines, producing, distributing and delivering
9 vaccines to the population, and increasing health care capacity (23).

10 The outbreak of the highly transmissible delta variant (B.1.617.2) raises new
11 questions among policy makers and health system managers regarding the impact
12 on hospitalization rates, a point reinforced by a recent published study that indicates
13 a possible worst-case scenario compared to the alpha variant (B.1.1.7) (24).
14 Additionally, according to Twohig KA et al. (2021), this recent research adds to
15 previous evidence and reinforces the notion that the delta variant, in unvaccinated
16 populations, may increase the burden on the overall health care system when
17 compared to the SARS-CoV-2 alpha variant (24).

18 The Brazilian ICU occupancy condition became critical in most of the country
19 in March 2021, when 20 states reached critical rates (equal to or greater than 80%
20 occupancy) and became intermediate in most states by June 2021 (between 60 and
21 80%), in concurrence with high SARS-CoV-2 transmission rates (25,26).

22 It is necessary to maintain biosafety education for the general population
23 (social distancing and the use of masks), encourage them to seek the vaccine when
24 called upon by age/group and complete the recommended number of doses. At the
25 organizational level of hospitals, it is critical to measure resources and indicators,
26 planning strategies according to the literature and lessons learned during the
27 pandemic to meet the next demand. For the public policy level, it is crucial to

1 intensify vaccination throughout the country and support the health system by
2 consolidating its structure and strategies for serving the population.

3 The initial ignorance regarding the virus and the associated disease led
4 hospitals and health services to seek different internal management strategies to
5 meet the demand coming with COVID-19 (27). Berwick DM used the term "choices
6 for the new normal" to point to six properties of care for lasting change: time,
7 standards, working conditions, proximity, readiness and equity. For the author, the
8 changes in care required by the pandemic may allow health systems to become
9 more patient-centered and efficient and consequently provide better care to the
10 public (28).

11 Since the spectrum of SARS-CoV-2 infection can range from asymptomatic
12 infection to potentially fatal complications of COVID-19, we chose to list patients with
13 suspected and confirmed infection who were admitted to hospital units (29). Thus,
14 we believe that the strategy of the "COVID-19 Global Clinical Data Platform for
15 clinical characterization and management of hospitalized patients with suspected or
16 confirmed COVID-19" contributes to the understanding of the Brazilian context of
17 pandemic crisis in the health care system by providing data analysis of clinical
18 aspects and epidemiology.

19 It also reinforces a portrait of the reality of the institutions based on the current
20 demand, discussion on aspects of the main organizational frameworks for fighting
21 the pandemic and how to keep the institutions functioning and meeting the health
22 needs of the population amidst the challenges.

1 **Conclusions**

2 This project describes the collaboration of several Brazilian hospitals to build a
3 joint database, with special attention to protocols aimed at reducing heterogeneity
4 between different platforms and ensuring the quality of research data.

5 As part of the process, the project management carried out an organized
6 multicenter research framework involving PAHO/WHO, as the center of the strategy,
7 and Brazilian public hospitals, resulting in substantial data collection. In addition, the
8 data extraction used innovative techniques in the technology area, promoting quality
9 and mitigating the lack of relevant information, moving towards data collection in
10 unstructured fields, such as clinical evolution which are the texts with the most
11 detailed health data in the medical records.

12 This proposal accomplishes an important step towards fighting COVID-19 in a
13 developing country of continental dimensions such as Brazil, by building a foundation
14 for new collaborations between referral hospitals and demonstrating the feasibility of
15 joint initiatives for the emergence of "Big Data" analysis strategies.

16 **Acknowledgements**

17 This project received the institutional support from the hospital institutions
18 mentioned here through access to data and research participants, which were critical
19 for this research, the reason for our gratitude.

20 The authors would like to thank the research participants who made the
21 development of the research project possible.

22 **Conflict of interests**

23 The authors declare that they have no conflict of interest with the subject of the
24 referred protocol.

References

1. Wu F, Zhao S, Yu B, Chen YM, Wang W, Song ZG, et al. A new coronavirus associated with human respiratory disease in China. *Nature*. 2020;579(7798):265–9.
2. World Health Organization. Coronavirus Disease 2019. Situat Rep [Internet]. 2020;54:1–9. Available from: <https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200314-sitrep-54-covid-19.pdf>
3. Roussel Y, Giraud-gatineau A, Jimeno M, Rolain J. SARS-CoV-2: fear versus data. *International J Antimicrob Agents*. 2020;55(January):1–3.
4. Shen Y, Cui Y, Li N, Tian C, Chen M, Zhang YW, et al. Emergency responses to Covid-19 Outbreak: experiences and lessons from a general hospital in Nanjing, China. *Cardiovasc Intervent Radiol* [Internet]. 2020;43(6):810–9. Available from: <https://doi.org/10.1007/s00270-020-02474-w>
5. Li T. Diagnosis and clinical management of severe acute respiratory syndrome Coronavirus 2 (SARS-CoV-2) infection: an operational recommendation of Peking Union Medical College Hospital (V2.0): Working Group of 2019 Novel Coronavirus, Peking Union Medical Colle. *Emerg Microbes Infect*. 2020;9(1):582–5.
6. Cimerman S, Chebabo A, Cunha CA da, Rodríguez-Morales AJ. Deep impact of COVID-19 in the healthcare of Latin America: the case of Brazil. *Brazilian J Infect Dis*. 2020;24(2):93–5.
7. Barberia LG, Gómez EJ. Political and institutional perils of Brazil's COVID-19 crisis. *Lancet*. 2020;396(10248):367–8.
8. Gonçalves MR, Padilha dos Reis RC, Tólio RP, Pellanda LC, Schmidt MI, Katz N, et al. Social distancing, mask use, and transmission of severe acute respiratory syndrome Coronavirus 2, Brazil, April–June 2020. *Emerg Infect Dis*. 2021;27(8):2135–43.
9. Flaxman S, Mishra S, Gandy A, Unwin HJT, Mellan TA, Coupland H, et al. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature*.

2020;584(7820):257–61.

10. Walker PGT, Whittaker C, Watson OJ, Baguelin M, Winskill P, Hamlet A, et al. The impact of COVID-19 and strategies for mitigation and suppression in low- and middle-income countries. *Science* (80-). 2020;369(6502):413–22.
11. Federação Brasileira de Hospitais, Confederação Nacional de Saúde. Cenário dos hospitais no Brasil 2018 [Internet]. 2018. 108 p. Available from: http://fbh.com.br/wp-content/uploads/2020/11/1-Relatorio-FBH-CNS_web.pdf
12. Ministério da Saúde B. Cadastro Nacional de Estabelecimentos de Saúde [Internet]. 2021. Available from: http://cnes.datasus.gov.br/Cabecalho_Reduzido_Competencia.asp?VCod_Unidade=3205304044916
13. Cotrim Junior DF, Cabral LM da S, Machado FR de S. Fila única de leitos: urgência, dificuldades e enfrentamentos. *P2P E Inovação*. 2020;7:212–29.
14. Ranzani OT, Bastos L, Gelli J, Marchesi J, Baião F, Hamacher S, et al. Characterisation of the first 250,000 hospital admissions for COVID-19 in Brazil: a retrospective analysis of nationwide data. *Lancet Respir Med* 2021; 2020;9(January):407–18.
15. Claro IM, Cristina F, Ramundo MS, Candido DS, Silva CAM, Jesus JG De, et al. Local transmission of SARS-CoV-2 Lineage B.1.1.7, Brazil, December 2020. *Emerg Infect Dis* [Internet]. 2021;27(3):2020–2. Available from: www.cdc.gov/eid
16. da Silva MS, Demoliner M, Hansen AW, Gularte JS, Silveira F, Heldt FH, et al. Early detection of SARS-CoV-2 p.1 variant in Southern Brazil and reinfection of the same patient by p.2. *Rev Inst Med Trop Sao Paulo*. 2021;63:1–8.
17. Franceschi VB, Caldana GD, de Menezes Mayer A, Cybis GB, Neves CAM, Ferrareze PAG, et al. Genomic epidemiology of SARS-CoV-2 in Esteio, Rio Grande do Sul, Brazil. *BMC Genomics*. 2021;22(1):1–17.
18. Zeiser FA, Donida B, da Costa CA, Ramos G de O, Scherer JN, Barcellos NT, et al. First

and second COVID-19 waves in Brazil: A cross-sectional study of patients' characteristics related to hospitalization and in-hospital mortality. *Lancet Reg Heal - Am.* 2022;6(November 2021):100107.

19. Brasil M da S do. DATASUS - Ministério da Saúde [Internet]. Datasus. 2021 [cited 2021 Aug 30]. Available from: www.datasus.gov.br
20. Brasil P da R. Lei Geral de Proteção de Dados Pessoais (LGPD). Diário oficial da união [Internet]. 2018; Available from: <file:///C:/Users/usuario/Desktop/L13709compilado.html>
21. Merianos A, Peiris M. International Health Regulations (2005). *Lancet.* 2005;366(January):1249–51.
22. Bartsch SM, Ferguson MC, McKinnell JA, O'shea KJ, Wedlock PT, Siegmund SS, et al. The potential health care costs and resource use associated with COVID-19 in the United States. *JAMA J Am Med Assoc.* 2020;39(6):927–35.
23. Glied S, Levy H. The potential effects of coronavirus on national health expenditures. *JAMA - J Am Med Assoc.* 2020;323(20):2001–2.
24. Twohig KA, Nyberg T, Zaidi A, Thelwall S, Sinnathamby MA, Aliabadi S, et al. Hospital admission and emergency care attendance risk for SARS-CoV-2 delta (B.1.617.2) compared with alpha (B.1.1.7) variants of concern: a cohort study. *Lancet Infect Dis.* 2021;3099(21):1–9.
25. Fiocruz. Observatório Covid-19 [Internet]. <https://portal.fiocruz.br/>. 2021 [cited 2021 Aug 30]. p. 5. Available from: https://portal.fiocruz.br/sites/portal.fiocruz.br/files/documentos/boletim_covid_2021_extraordinario_30dejunho.pdf
26. Fiocruz. Leitos Uti Covid-19 Para Adultos [Internet]. <https://portal.fiocruz.br/>. 2021. p. 1. Available from: https://portal.fiocruz.br/sites/portal.fiocruz.br/files/documentos/serie_historica_leitos_uti_covid-19_adultos.pdf

27. Kim MK, Rabinowitz LG, Satish Nagula M, Dunn A, Chalil J, Xu T, et al. A primer for clinician deployment to the medicine floors from an epicenter of Covid-19. *NEJM Catal.* 2020;
28. Berwick DM. Choices for the “new Normal.” *JAMA - J Am Med Assoc.* 2020;323(21):2125–6.
29. Ortiz-prado E, Simbaña-rivera K, Barreno LG-, Rubio-neira M, Cevallos-robalino D, Sanches-sanmiguel H, et al. Clinical, molecular, and epidemiological characterization of the SARS-CoV-2 virus and the Coronavirus Disease 2019 (COVID-19), a comprehensive literature review. *Diagn Microbiol Infect Dis.* 2020;98(115094):1–31.

Table 2 - WHO Global Clinical Platform for COVID-19 characterization in Brazil: collaborating centers

STATE	Hospital Institution	Hospital beds					Medical Specialties	Hospital Day	Total
		Surgical	Clinical	Complementary	Obstetrical	Pediatric			
AM	HU Getúlio Vargas (HUGV-UFAM)	58	59	32	0	0	0	0	149
PA	Hospital Universitário João de Barros Barreto (HUJBB-UFPA)	40	178	38	0	28	12	10	306
CE	HU Walter Cantídio (HUWC-UFC)	68	72	21	0	20	22	4	207
CE	Maternidade Escola Assis Chateaubriand (MEAC-UFC)	21	58	65	104	6	0	0	254
CE	Hospital Regional Norte (HRN-ISGH)	58	183	177	33	70	16	0	537
CE	Hospital Geral Waldemar Alcântara (HGWA-ISGH)	18	188	58	0	33	33	0	330
CE	Hospital Regional do Cariri (HRC-ISGH)	88	112	113	0	0	29	0	342
CE	Hospital Regional do Sertão Central (HRSC-ISGH)	58	104	106	30	1	14	0	313
CE	Hospital Estadual Leonardo Da Vinci (Helv-ISGH)	1	112	179	0	0	0	0	292
CE	Hospital Regional Vale do Jaguaribe (HRVJ-ISGH)	0	30	10	0	0	0	0	40
DF	HU da Universidade de Brasília (HUB-UnB)	49	76	41	30	20	3	12	231
ES	HU Cassiano Antônio Moraes (HUCAM-UFES)	82	90	55	20	24	0	2	273
MA	HU da Universidade Federal do Maranhão	134	75	81	71	92	0	0	453
MT	HU Julio Muller (HUJM-UFMT)	12	20	31	24	11	0	0	98
MS	HU Maria Aparecida Pedrossian (Humap-UFMS)	60	45	53	32	18	0	12	220
MS	HU da Univ. Federal da Grande Dourados (HU-UFGD)	29	41	59	25	35	1	0	190
MG	HC da Univ. Federal do Triangulo Mineiro (HC-UFTM)	93	86	57	22	42	2	8	310
MG	HU da Univ. Federal de Juiz de Fora (HU-UFJF)	19	61	29	0	17	0	17	143
MG	HC da Univ. Federal de Minas Gerais (HC-UFMG)	119	114	107	35	63	1	30	469
PB	HU Lauro Wanderley (HULW-UFPB)	31	80	43	36	22	0	33	245
PE	HU da Univ. Fed. Vale do S. Francisco (HU-UNIVASF)	86	35	21	0	0	8	0	150
PE	HC da Univ. Federal de Pernambuco (HC-UFPE)	118	119	27	30	38	13	76	421
RJ	HU Antônio Pedro (HUAP-UFF)	83	54	59	21	17	0	10	244
RN	HU Ana Bezerra (HUAB-UFRN)	8	6	10	33	10	0	0	67
RS	HE da Univers. Feder. de Pelotas (HE-UFPe)	8	109	31	21	20	0	10	199
RS	HU da Univ. Federal de Santa Maria (HUSM-UFSM)	75	163	60	35	41	1	10	385
SP	HU da Univ. Federal de São Carlos (HU-UFSCAR)	0	56	12	0	17	0	0	85

PB	HU Alcides Carneiro (HUAC-UFCG)	29	84	31	0	33	0	0	177
PB	HU Julio Bandeira (HUJB-UFCG)	12	2	1	0	15	0	2	32
RJ	HU Gaffree e Guinle (HUGG-UNIRIO)	95	89	38	19	6	2	3	252
SE	HU da Universidade Federal de Sergipe (HU-UFS)	49	48	25	0	14	10	10	156
SE	Hospital Universitário de Lagarto (HUL-UFS)	25	46	33	0	16	0	10	130
TO	Hospital de Doenças Tropicais (HDT-UFT)	0	37	2	0	9	0	6	54
BA	Hospital Especializado Octávio Mangabeira	23	105	28	0	17	42	1	216
PR	Complexo Hospitalar do Trabalhador	126	37	132	35	30	24	0	384
PR	Hospital do Trabalhador	4	22	0	2	4	0	0	32
PR	Hospital Osvaldo Cruz	0	29	0	0	1	4	8	42
DF	Hospital da Criança de Brasília José Alencar	0	0	58	0	150	2	8	218
PI	Hospital Getúlio Vargas	170	87	94	0	0	2	11	364
RS	Hospital Nossa Senhora da Conceição - Grupo Hospitalar Conceição	207	357	332	62	105	30	13	1106
RS	Hospital de Clínicas de Porto Alegre - HCPA/UFRGS	228	224	259	50	95	30	17	903
SP	Instituto de Infectologia Emilio Ribas, São Paulo	4	82	232	0	16	10	31	375
SP	Hospital Nestor Goulart Reis Americo Brasiliense	0	20	0	0	0	73	0	93
Total		2388	3595	2840	770	1156	384	354	11487

STATE; AM: Amazonas; PA: Pará; CE: Ceará; DF: Distrito Federal; PB: Paraíba; PI: Piauí; RS: Rio Grande do Sul; RN: Rio Grande do Norte; SP: São Paulo; RJ: Rio de Janeiro; BA: Bahia; TO: Tocantins; SE: Sergipe; PA: Pará; PE: Pernambuco; ES: Espírito Santo; MA: Maranhão; MT: Mato Grosso; MS: Mato Grosso do Sul; MG: Minas Gerais.

HU: University Hospital; HE: Teaching Hospital

Source: The Brazilian National Register of Health Service Providers (CNES) (13).

Data Statment Section

Data from all surveyed centers can be aggregated and analyzed – together or by country center or region – to provide a better understanding of the disease and present the public health response to the pandemic via the platform URL: <https://www.who.int/teams/health-care-clinic-readiness-unit/COVID-19/dataplatform>

Author Statement

- Fernando Anschau¹,

The author contributed to the organization of the group of researchers, data collection, analysis and writing of the protocol.

- Natália Del' Angelo Aredes²

The author made substantial contributions to the design of work, data collection based on the protocol (in progress), literature review and writing. Revised the manuscript and have approved the submitted version.

- Ludovic Reveiz³;

The aforementioned author worked in the organization of the group of researchers and logistical support.

- Monica Padilla³;

The aforementioned author worked in the organization of the group of researchers and logistical support.

- Rosane de Mendonça Gomes⁴;

Logistical support and data analysis.

- Wellington Mendes Carvalho⁴;

Logistical support and data analysis.

- Fernando Antonio Gomes Leles⁴;

Logistical support and data analysis.

- Fernanda Baeumle Reese⁵;

Database organization and data collection.

- André Hostílio Hubert⁵;

Database organization and data collection.

- Elisandrea Sguario Kemper⁶;

Database organization and data collection.

- Renilson Rehem de Souza⁶;

Database organization and data collection.

- Cristiane Feitosa Salviano⁶;

Database organization and data collection.

- Hevelin Silveira e Silva⁶;

Database organization and data collection.

- Eduardo Barbosa Coelho⁷;

Paper writing and data analysis.

- Giuseppe Cesare Gatto⁷;

Database organization and data collection.

- Rafael Freitas de Moraes⁷;

Database organization.

- Leonardo Nunes Alegre⁷,

Database organization.

- Rodrigo Citton Padilha dos Reis⁸,

This researcher contributed with data collection.

- Joaquim Francisco dos Santos Neto⁸,

Database organization and data collection.

- Andresa Fontoura Garbini¹;

This researcher contributed with data collection and analysis.

- Cesar Perdomo Purper¹,

Database organization.

- Veridiana Baldon dos Santos¹;

This researcher contributed with data collection and analysis.

- Rafaela dos Santos Charão de Almeida¹;

This researcher contributed with data collection and analysis.

- Bruna Donida¹;

Database organization and data collection.

- Rogério Farias Bitencourt¹;
Database organization.
- Luciane Kopittke¹;
This researcher contributed with data collection, paper writing and analysis.
- Fernanda Costa dos Santos¹;
This researcher contributed with data collection and analysis.
- Raquel Lutkmeier¹;
Paper writing and data analysis.
- Daniela dos Reis Carazai¹;
This researcher contributed with data collection and analysis.
- Virgínia Angélica Silveira Reis⁹;
Database organization and data collection.
- Flávio Clemente Deulefeu⁹;
Database review, partial data analysis and data collection.
- Fernanda Gadelha Severino⁹;
This researcher contributed with data collection and analysis.
- José Gustavo da Costa Neto⁹;
This researcher contributed with data collection and analysis.
- Nirvania do Vale Carvalho¹⁰,
Database review, partial data analysis and data collection.
- André Jamson Rocha de Andrade¹⁰;
This researcher contributed with data collection and analysis.
- Adriana Melo Teixeira¹¹;
The author contributed to the organization of the group of researchers.
- Olavo Braga Neto¹¹,
The author contributed to the organization of the group of researchers.
- Gabriel Cardozo Muller⁸,

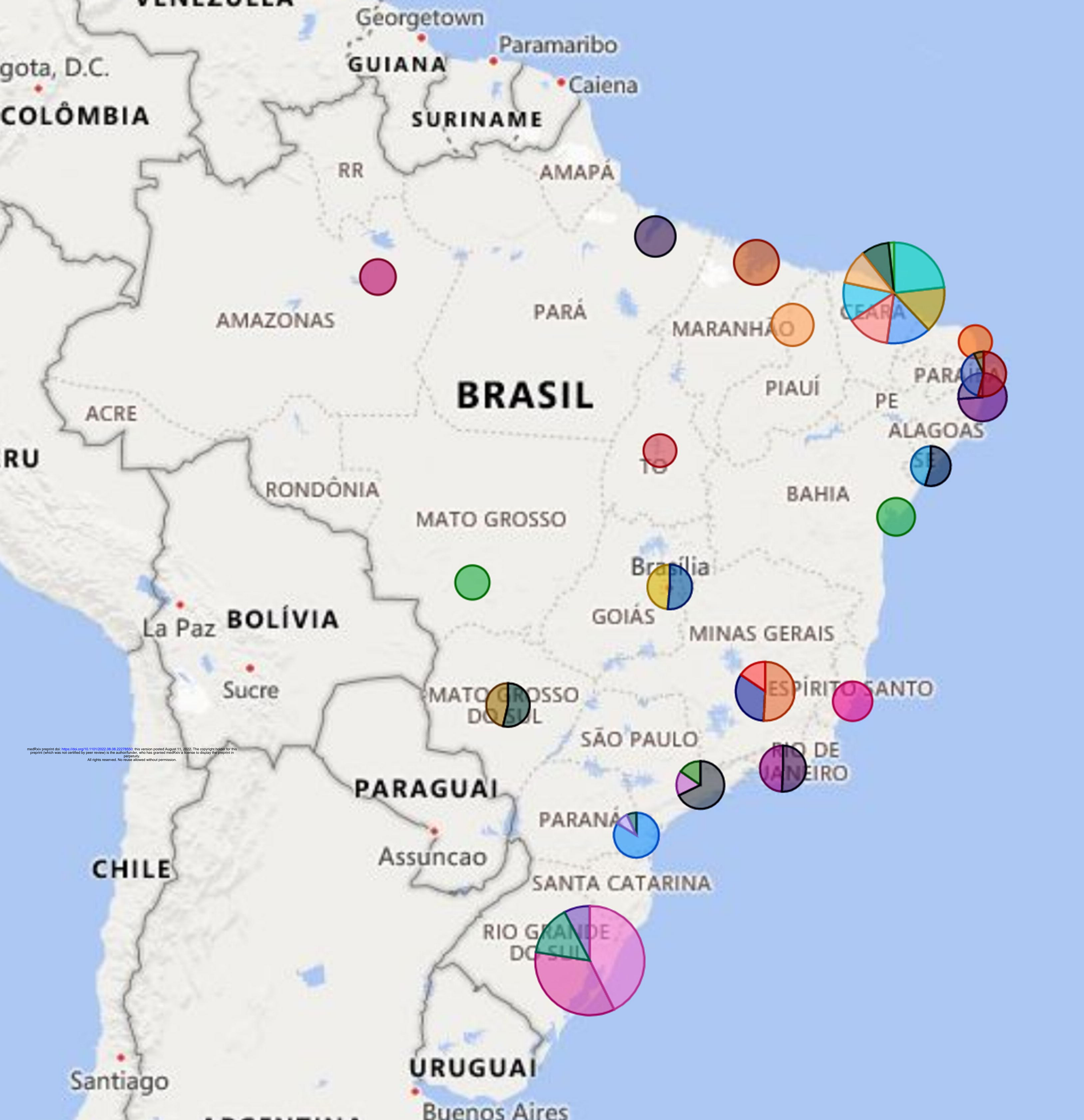
This researcher contributed with data collection and analysis.

- Ricardo de Souza Kuchenbecker⁸

The author contributed to the organization of the group of researchers, substantial organization of the protocol, improvements in the course of the project, data collection and analysis and writing of the protocol.

Figure 1. Hospitals participating in the survey by size classification

Figure 2. Illustration of flow diagram of study protocol



medRxiv preprint doi: <https://doi.org/10.1101/2022.08.08.22278560>; this version posted August 11, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

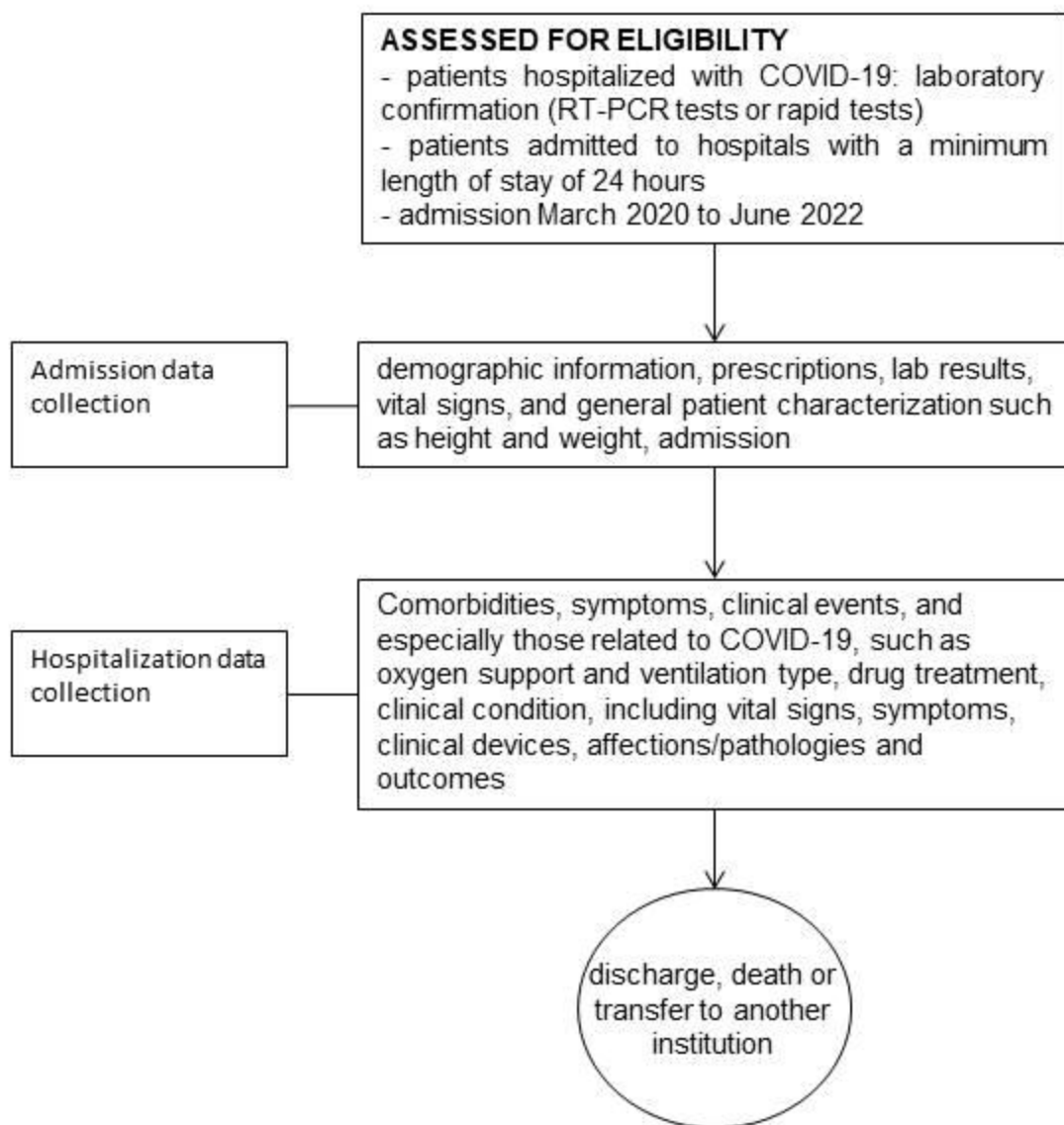


Figure 1. Illustration of flow diagram of study protocol