Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 196 (2022) 525–532

Procedia Computer Science

www.elsevier.com/locate/procedia

CENTERIS - International Conference on ENTERprise Information Systems / ProjMAN - International Conference on Project MANagement / HCist - International Conference on Health and Social Care Information Systems and Technologies 2021

COVID-19 BR: A web portal for COVID-19 information in Brazil

Isabelle Carvalho[a,*], Filipe Andrade Bernardi[b,c], Mariane Barros Neiva[a], Vinícius Costa Lima[b,c], Lariza Laura de Oliveira[b], Newton Shydeo Brandão Miyoshi[b], Tiago Lara Michelin Sanches[b], Francisco Barbosa-Junior[b], Nathalia Yukie Crepaldi[b], Diego Bettiol Yamada[b], Domingos Alves[b]

[a]*Institute of Mathematical and Computer Sciences, University of Sao Paulo, Av Trabalhador Sao-Carlense 400, 13566590, Sao Carlos/SP, Brazil*
[b]*Ribeirão Preto Medical School, University of Sao Paulo, Av Bandeirantes 3900, 14049900, Ribeirao Preto/SP, Brazil*
[c]*São Carlos School of Engineering, Av Trabalhador Sao-Carlense 400, 13566590, Sao Carlos/SP, Brazil*

Abstract

Brazil is a large developing country that requires attention to regionalized behaviors regarding the dissemination of COVID-19. To deal with this complexity, the COVID-19 Brazil observatory was developed. The Portal aims to monitor and analyze data from different sources. Therefore, with a detailed audit, we centralized this information on the evolution of the disease, allowing for territorial and temporal monitoring. The daily publication of numbers about COVID-19 allowed anyone to follow the current scenario in several Brazilian cities. With about 1,7 million accesses, the Portal offers clarity and an easy understanding of the pandemic data in the country.

© 2021 The Authors. Published by Elsevier B.V.
This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0)
Peer-review under responsibility of the scientific committee of the CENTERIS –International Conference on ENTERprise Information Systems / ProjMAN - International Conference on Project MANagement / HCist - International Conference on Health and Social Care Information Systems and Technologies 2021

Keywords: COVID-19; Data Mining; Knowledge Discovery; Mathematical Models; Public Health

* Corresponding author.
E-mail address: isabelle.carvalho@alumni.usp.br

1877-0509 © 2021 The Authors. Published by Elsevier B.V.
This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0)
Peer-review under responsibility of the scientific committee of the CENTERIS –International Conference on ENTERprise Information Systems / ProjMAN - International Conference on Project MANagement / HCist - International Conference on Health and Social Care Information Systems and Technologies 2021
10.1016/j.procs.2021.12.045

1. Introduction

The COVID-19 pandemic is one of the most devastating infectious disease outbreaks in world history, with an estimated death toll of 3,721,000 worldwide by the time this paper was written [1]. Furthermore, the occurrence of asymptomatic patients makes it difficult to control the disease. This also challenges the process of understanding the virus's dynamics by researchers worldwide [2].

Although there is evidence of significant geographical variation in mortality rates between countries and even between cities, the disease has a spread behavior based on local contact among people. Therefore, it is essential to assess the epidemiological situation in each region. The availability of information contributes to the development of epidemiological and economic models that use trend estimates and can support the effective planning of measures to prevent the spread of the disease in a given region.

The pandemic has revealed many loss-making areas of public health, especially in low- and middle-income countries. Digital interventions offer many opportunities to strengthen health systems and be vital resources in the current public health emergency. The applications of digital technology for treatment, diagnosis, support for self-management, and surveillance are well known. Many countries have systems in place to serve a variety of health needs [3].

Governments, health agencies, and health care providers are called upon to immediately and consistently leverage the power of digital health tools. This can strengthen the capacity of their health system to respond to the COVID-19 pandemic. The use, feasibility, and importance of these and other applications vary according to the needs of the country, the existing infrastructure, and other factors [4].

In many cases, the access to relevant digitized information from research funding agencies, academics, managers, and health professionals is harmed. This is due to the non-computerization of processes, the heterogeneity and duplicity of data, and the large amount of data isolated in databases accessible only in specific contexts. Such factors often cause problems in the quality of information, making it challenging to coordinate, assess a Patient Care Network and make it impossible to support the decision-making process [5].

In addition, to accelerate progress in the fight against infectious diseases, a change from a strategy focused solely on control to a strategy focused on eliminating the disease is necessary. Successful campaigns are characterized by locally adapted responses informed by the appropriate data. Developing a locally targeted response to these diseases will require substantial investments to reconfigure existing systems, along with additional empirical data to assess the effectiveness of approaches. Without adopting an elimination strategy that uses local data to reach critical transmission points, ambitious goals to end the transmission chain will almost certainly not be met [6].

In public health surveillance and epidemiology, digital tools can be valuable to reduce the risk of exposure for public health professionals. Using a variety of remote methods, critical tasks can be performed in secure environments while collecting and analyzing high-quality data needed to mitigate the effects of the pandemic. Innovative solutions are necessary to ensure that healthcare facilities remain prepared.

1.1. Objectives

This article presents a voluntary initiative by independent researchers to organize, analyze and disseminate information related to the COVID-19 pandemic in Brazil: the COVID-19 Brazil Portal (https://ciis.fmrp.usp.br/covid19 /). Considering the decentralization of information in our country, the Portal has specific objectives for auditing, monitoring, modeling, and intuitive presentation of COVID-19 data based on data science techniques.

2. Methods

The study is considered a mixed nature and descriptive design, having adopted the participatory research methodology. This type of methodology is characterized by the involvement and identification of researchers with the investigated situation. In this sense, through actions and seeking to integrate knowledge, it is possible to benefit the study group. It is assumed that it is essential to present complete empirical research combining: the advantage of large samples and their generalizations.

THE DATA

According to Brazilian law 12527/2011 (Law of Access to Information), access to information is a fundamental right. It represents an essential condition for democracy and the monitoring of public investments [7]. In the health area, guaranteeing citizens' access to information through transparency in public management is a vital requirement for evaluating infection surveillance and control measures, as in COVID-19. However, the information provided is incomplete and does not meet the criteria of clarity, transparency, and timeliness. Even if the information is available, it is not regularly in a machine-readable format or an open and descriptive form, a requirement for knowledge production.

To meet the need for information coverage, data is collected through epidemiological bulletins from municipal and state health departments, in addition to a repository of publicly available data. The stratification process takes place automatically or manually, according to availability. The collection must meet conditions such as the source and hierarchy (municipal, state, and federal); the structure and storage of the data; the frequency of data availability; and the comparison between the sources.

Thus, the verification of the data collected manually from websites, social networks, secretaries, city halls, and municipal health secretariats are confronted with those extracted automatically from official government sources (https://covid.saude.gov.br/), third-party application programming interfaces, and the transparency portal of the Brazilian Registries (https://transparencia.registrocivil.org.br/inicio).

The high level of granularity and the heterogeneity of structures and formats (e.g., spreadsheets, images, graphs, and reports) represent a challenge in the systematic acquisition and, consequently, in data processing. The inconsistency of data regarding the COVID-19 pandemic is recurrent due to several factors, including, mainly, the diversity of sources and the level of information flow [8]. Thus, auditing was introduced in both flows, automatic and manual, with filters and automated validations to guarantee completeness, coherence, and consistency in the data.

PORTAL DEVELOPMENT

The Portal development follows the guidelines model proposed by the WHO for the development of health observatories [9], and the web development standards defined by the Informatics Department of the Brazilian Ministry of Health (DATASUS/MS, https://www.datasus.saude.gov.br). The content management system WordPress was used to develop the Portal, which helps users publish, organize, and manage content posted on the web. The Hypertext Preprocessor programming language offers native support for several databases, such as MySQL. With support for HTML and CSS markup languages and JavaScript libraries, WordPress allows incorporating graphics interfaces and reports [10]. Finally, different statistical analyses and previous mathematical models are generated daily from Python and R scripts and published.

3. Results and Discussion

3.1. Monitoring

Monitoring provides temporal visualization of COVID-19 data in six main aspects: [i] Brazil; [ii] Brazil and other countries; [iii] States of Brazil (comparison); [iv] Capitals of each state (comparison); [v] Brazilian states and their capitals (individuality); [vi] Ribeirão Preto city.

Each level of disaggregation is characterized by a set of indicators: number of daily cases, number of accumulated patients, number of daily deaths, number of accumulated deaths, contamination rate (number of patients per 100 thousand inhabitants), and lethality rate (number of deaths per number of cases). It is possible to interact with the graphics presented by clicking on the legend to choose what he sees and, thus, customize the comparisons. In addition, the Portal presents cases in georeferenced maps and interactive dynamic maps. Figure 1 shows 4 examples of monitoring views present in the Portal.

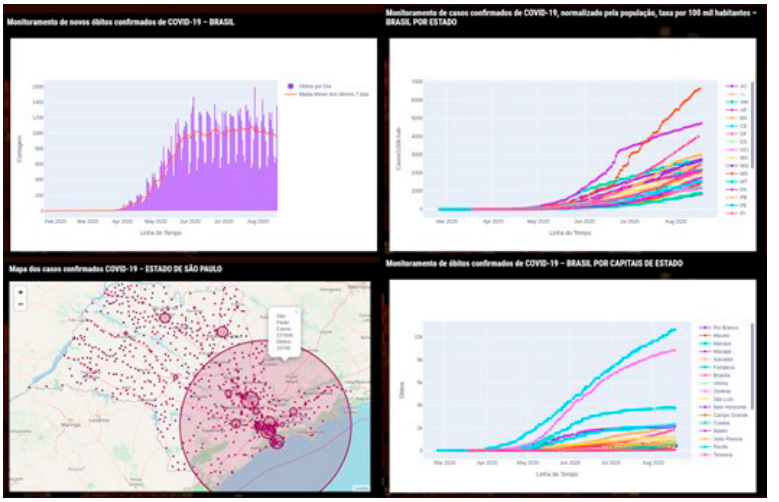

Figure 1. Examples of views. Monitoring - COVID-19 Brazil Portal.
Extracted from https://ciis.fmrp.usp.br/covid19/

3.2. Analysis

The observatory features five main types of analysis with daily updates:

- **The moving average** for daily cases and deaths, allowing to reduce the interference of delay in notifications from hospitals to the health secretaries. Some processing sites are closed on weekends, causing a sudden increase in the number of cases on the first days of the week. The moving average evaluation is made from the average of the last 7 days for each new occurrence.
- **Exponential analysis:** the current accumulated cases and deaths data are projected for the next 10 days. The calculation is done using computational methods with epidemiological knowledge of the exponential nature of infection [11]. This analysis allows us to understand the natural course of the disease considering the same circumstances of the last 14 days. This is important to develop public policies to possibly change the scenario in the evolution of COVID-19. An example is shown in Figure 2.

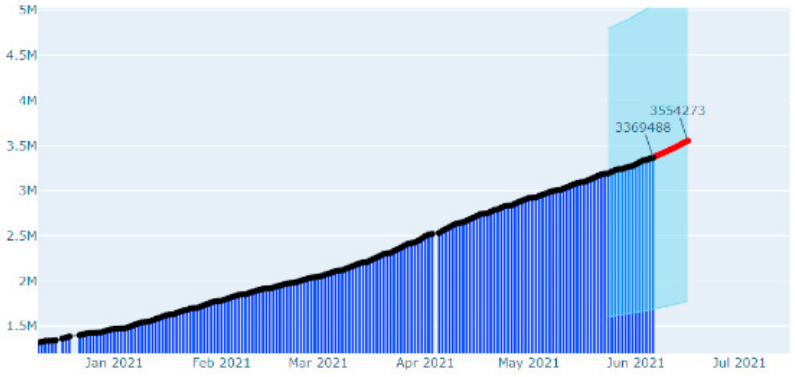

Figure 2: The exponential analysis shows the projection for the next 10 days. The technique is based on the nature of the disease' transmission and evaluates the last two weeks to project the following days.
Extracted from https://ciis.fmrp.usp.br/covid19/

- **Real data estimation:** Brazil has a low detection rate due to its low testing of suspected cases [12]. Therefore, an important analysis to be done is the estimation of real cases [13]. The study of the estimate of real cases uses the register of Brazilian deaths of a certain day, which is more reliable than the number of cases, to compute the number of patients in the last 10 days. The work is based on the alert published by Lancet [14] and also uses information on age groups to correct the case fatality rate. Figure 3 shows the estimate for Brazil and it is possible to browse by restricting the information for all states and capitals. In particular, this computational tool was developed and adapted from the methodology introduced by Lachmann et al. (2020).

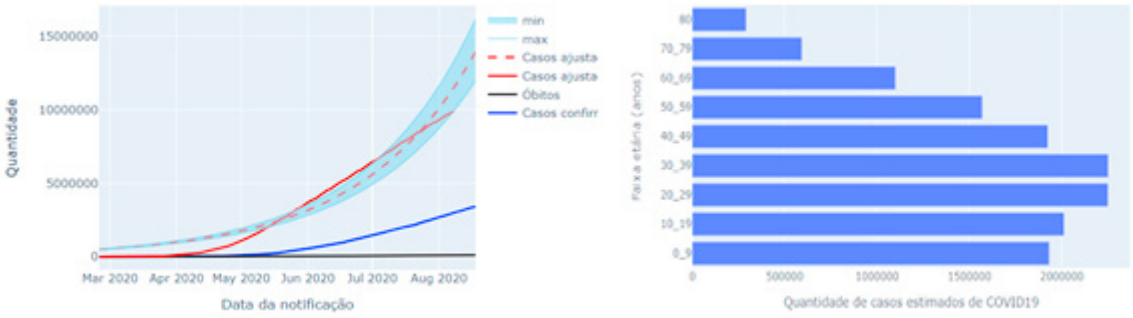

Figure 3: Estimate of cases based on Lancet lethality and alert rate [13].
Extracted from https://ciis.fmrp.usp.br/covid19/

- **Stability analysis:** in COVID-19, high exposure to the virus takes around 14 days to affect the number of accounted cases. Therefore, it is possible to compare the current values to the number of new cases from the previous two weeks. With a tolerance of 15% of cases, it is established which regions increased their patients, which decreased, and the ones that reached a plateau. The results, represented in Figure 4, are shown as colours in a map: blue for decrease, red for the increase, and yellow for stability.

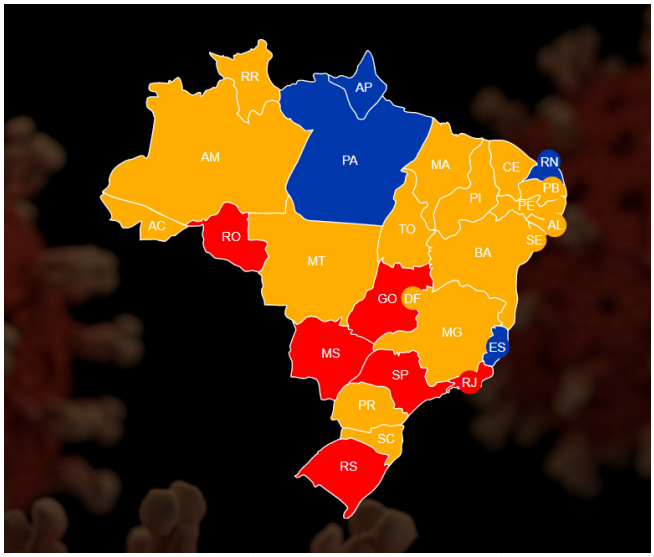

Figure 4: Colours represent the stability status of each state. Blue for decrease, red for the increase, and yellow for stability.
Extracted from https://ciis.fmrp.usp.br/covid19/

- **The vaccine run:** since the beginning of 2021, Brazil has started to vaccinate its population. Thus, we added a novel analysis in the Portal. Using a bar graph, as seen in Figure 5, we plot the percentage of people in each state immunized with the first and the second dose of the COVID-19 vaccine. We sort in a descending way

to check which state applied the most amount of the vaccine and to keep the population informed about the immunization process in Brazil.

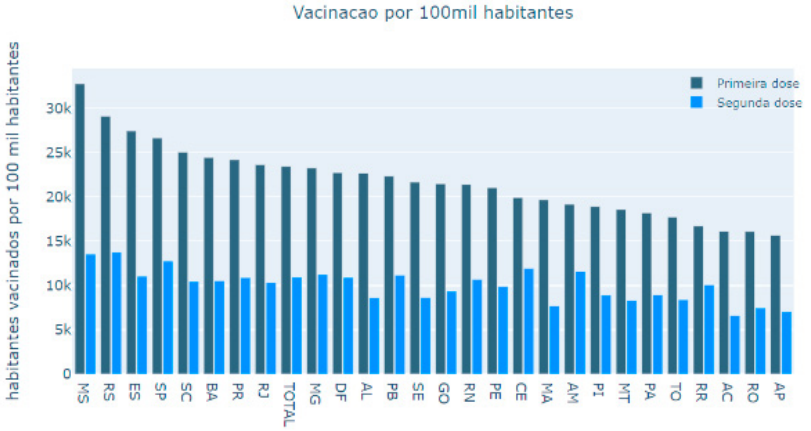

Figure 5: The vaccine graph shows the relative number of vaccinated people in each state of Brazil. We sort the values in a descending way and show both the first and second number of doses applied.
Extracted from https://ciis.fmrp.usp.br/covid19/

3.3. Follow-up tools

In addition to the direct interaction with data referring to the COVID-19 scenario in Brazil, the following tools are currently available for users to explore other issues:

(i) **Covid-calc**: Hospital pressure by Covid-19: the calculator was developed by the University of Brasília (UnB) with support from the Pan American Health Organization (PAHO/WHO), aiming to project hospital pressure according to the number of confirmed cases of Covid-19 in Brazil, states, and municipalities, in addition to hospital capacity, the epidemiological scenario, and the adopted containment measures [15].

(ii) **Epcalc - Epidemic Simulator:** Epcalc is a tool that makes it possible to analyze possible scenarios for the evolution of a hypothetical epidemic through epidemiological parameters and mathematical models. This calculator implements a classic infectious disease model - SEIR (Susceptible, Exposed, Infected, Recovered) - a model of propagation still used in the front lines of research [16]. The dynamics are characterized by four ordinary differential equations that correspond to the stages of disease progression [17]. Complementary to the transmission dynamics, this model allows supplementary time information to model the mortality rate and healthcare burden.

(iii) **Interactive transmission model**: a tool for evaluating trends from a set of pre-defined parameters. The parameters used do not have an obligation to generate values that exactly correspond to the real situation. Adapted from the "Virus in the City" model [18], the toy model was developed in the integrated development environment based on NetLogo agents.

(iv) **My Quarantine Diary**: the platform monitors the evolution of the patient's symptoms. The goal is to produce reliable data that anticipate a more accurate case since the symptoms of COVID-19 are similar to the common flu and can quickly evolve. The platform (https://www.meudiariodequarentena.com.br/) allows self-monitoring through a diary that analyses their symptoms and suggests an action in real-time. Additionally, activities that encourage mental health during quarantine are available, such as free language courses, exercise channels, and entertainment tips. A FAQ area with questions and answers is available. Finally, the platform also offers a game where it is possible to clarify divergent information and fake news.

(v) **Sentiment analysis of tweets**: the tool reports the population responses and reactions to events related to the coronavirus. To develop the tool, natural language processing techniques with sentiment analysis are applied to investigate tweets on Twitter [19].

4. Conclusion

Brazil is one of the few countries with a medium level of development with a universal public health system and electronic information systems. This study is of enormous scientific and technological relevance at the national level. It emphasizes the importance of the availability and reliability of the information, the need for interoperability of digital systems, and the empowerment of health professionals and local managers to make better decisions.

This project does not only refer to the use of existing computing techniques in the COVID-19 pandemic, but also call for developing new technologies and methods for defining paradigms in e-Science. It was possible to use and develop new performance evaluation methods for predictive models, useful for evidence-based decision-making.

The entire initiative for the construction and maintenance of the Portal has always been designed collaboratively. Over time, this collaboration network has expanded, allowing many to directly contribute to understanding the complexity of the pandemic. The Portal has been used to monitor, model, and present COVID-19 data and, therefore, contribute to controlling the coronavirus outbreak in the country. The collaborative work of the COVID-19 Brazil portal is the basis for assisting managers and making reliable information available to the population.

In particular, the Portal has been a reference for the production of various technical notes. It has supported decisions at the municipal and state levels and is taken as a reference by various official bodies and professional associations. It is important to highlight that throughout the trajectory, the Portal has been used by national and international media, both written and television, in the production of specific reports on the situation of the pandemic in Brazil (see https://ciis.fmrp.usp.br/covid19/saiu-na-midia/).

The success of this work can be measured through the visitors' statistics of the Portal (see https://ciis.fmrp.usp.br/covid19/estatisticas-do-site/). At the beginning of June 2021, the Portal had more than 1 million and 600 thousand accesses, coming from several countries worldwide. It is expected to continue this work, evaluating the scenarios of the pandemic of COVID-19 in Brazil and perhaps evolving, gradually, to face post-COVID-19 scenarios.

Finally, it is important to emphasize that the results and techniques of the Portal have the potential to generate records of computer programs with the National Institute of Industrial Property in Brazil (INPI-https://www.gov.br/inpi/pt-br).

Acknowledgements

The authors would like to thank all the contributors to COVID-19 BR Portal. The complete and updated list of people involved in this initiative can be consulted at https://ciis.fmrp.usp.br/covid19/quem-somos/.

References

[1] Worldometer (2021). "Coronavirus Cases." Available at: <https://www.worldometers.info/coronavirus/>

[2] Hao, Xingjie, Shanshan Cheng, Degang Wu, Tangchun Wu, Xihong Lin and Chaolong Wang (2020). "Reconstruction of the full transmission dynamics of COVID-19 in Wuhan." *Nature* **584(7821)**: 420-424.

[3] Mather, Thais P Salazar, Benjamin Gallo Marin, Giancarlo Medina Perez, Briana Christophers, Marcelo L Paiva, Rocío Oliva, Baraa A Hijaz, Andrea M Prado, Mateo C Jarquín, Katelyn Moretti, Catalina González Marqués, Alejandro Murillo and Elizabeth Tobin-Tyler (2020). "Love in the time of COVID-19: negligence in the Nicaraguan response." *The Lancet Global Health* **8(6)**: e773.

[4] Mahmood, Sultan, Khaled Hasan, Michelle Colder Carras and Alain Labrique (2020). "Global Preparedness Against COVID-19: We Must Leverage the Power of Digital Health" *JMIR Public Health and Surveillance* **6(2)**: e18980.

[5] Vinicius Costa Lima, Rui P C L Rijo, Felipe C Pellison, Ricardo R de Lima, Ricardo João Cruz Correia, Henrique T Giuriati and Domingos Alves. "From guidelines to decision-making: using mobile applications and semantic web in the practical case of guides to support patients". *Procedia Computer Science*, v. 121, p. 803-808, 2017.

[6] Theron,Grant, Helen E Jenkins, Frank Cobelens, Prof Ibrahim Abubakar, Aamir J Khan, Ted Cohen and Dr David W Dowdy (2015). "Data for action: collection and use of local data to end tuberculosis." *The Lancet* **386(10010)**: 2324-2333.

[7] Jardim, José Maria (2012). "A lei de acesso à informação pública." *Tendências da pesquisa brasileira em ciência da informação* **5(1)**.

[8] Spalluto, Lucy B, Virginia B Planz, LeAnn S Stokes, Richard Pierce, David M Aronoff, Melissa L McPheeters and Reed A. Omary (2020). "Transparency and trust during the coronavirus disease 2019 (COVID-19) pandemic." *Journal of the American College of Radiology* **17(7)**: 909-912.

[9] World Health Organization (2016). "Guide for the establishment of health observatories."

[10] Bo, Jiang and Qin Yanmei (2011). "Construction of Information Literacy Education Portal with WordPress [J]" *Journal of Modern Information* **10**.

[11] Park, Minah, Alex R. Cook, Jue Tao Lim, Yinxiaohe Sun and Borame L. Dickens (2020). "A systematic review of COVID-19 epidemiology based on current evidence." *Journal of clinical medicine* **9(4)**: 967.

[12] Hallal, Pedro Curi, Bernardo L Horta, Aluísio J D Barros, Odir A Dellagostin, Fernando P Hartwig, Lúcia C Pellanda, Cláudio José Struchiner, Marcelo N Burattini, Mariângela Freitas da Silveira, Ana M B Menezes, Fernando C Barros and Cesar Gomes Victora (2020). "Evolução da prevalência de infecção por COVID-19 no Rio Grande do Sul, Brasil: inquéritos sorológicos seriados." *Ciência & Saúde Coletiva* **25**: 2395-2401.

[13] Lachmann, Alexander (2020). "Correcting under-reported COVID-19 case numbers." *MedRxiv*.

[14] Baud, David, Xiaolong Qi, Karin Nielsen-Saines, Didier Musso, Léo Pomar and Guillaume Favre (2020). "Real estimates of mortality following COVID-19 infection." *The Lancet infectious diseases*, **20(7)**: 773.

[15] Zimmermann, Ivan Ricardo, Mauro Sanchez, Jonas Brant and Domingos Alves (2020). "Demand for ICU beds by COVID-19 in the Federal District, Brazil: an analysis of the impact of social distance measures with Monte Carlo simulations". *SciELO Preprints.* doi: https://doi.org/10.1590/SciELOPreprints.574.

[16] Wu, Prof Joseph T, Kathy Leung and Prof Gabriel M Leung (2020). "Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study". *The Lancet*, **395(10225)**, 689-697.

[17] Kucharski, Adam J, Timothy W Russell, Charlie Diamond, Yang Liu, John Edmunds, Sebastian Funk and Rosalind M Eggo (2020). "Early dynamics of transmission and control of COVID-19: a mathematical modelling study". *The Lancet Infectious Diseases* **20(5)**:553–8.

[18] Stonedahl, Forrest, and Uri Wilensky. (2008). "NetLogo Virus on a Network model". *Center for Connected Learning and Computer-Based Modeling*, Northwestern University, Evanston, IL.

[19] Bernardi, Filipe Andrade, Vinicius Costa Lima, Rui Pedro Charters Lopes Rijo and Domingos Alves (2021). "Mais do que palavras: uma análise das emoções brasileiras durante a COVID-19." *Revista Ibérica de Sistemas e Tecnologias de Informação*" **(E40)**: 526-541.