Gestão & Regionalidade

Recebido em: 23/04/2023
Aceito em: 01/09/2024

Factors associated with Codiv-19 deaths in the Brazilian Northeast: a multi-level approach

Fatores associados aos óbitos por Codiv-19 no Nordeste brasileiro: uma abordagem multinível

Adinésia Lima Leite[1 i], Orcid: https://orcid.org/0000-0002-0642-8353; **Thiago Henrique Carneiro Rios Lopes**[2 ii], Orcid: https://orcid.org/0000-0003-1550-1484; **Fernanda Esperidião**[3 iii], Orcid: https://orcid.org/0000-0002-8580-2090

1. Universidade Federal de Sergipe-UFS - São Cristovão - SE – Brasil. E-mail: adinesia.leite@dcomp.ufs.br
2. Universidade Federal de Sergipe-UFS - São Cristovão - SE – Brasil. E-mail: thiagohenriquerios@gmail.com
3. Universidade Federal de Sergipe-UFS - São Cristovão - SE – Brasil. E-mail: nandaesper16@gmail.com

Resumo
Em 2020 o mundo foi surpreendido pela Pandemia de COVID-19, causada pelo vírus SARS-Cov-2, que desencadeou uma crise humanitária, econômica e social a nível global. Neste sentido, este trabalho teve como objetivo analisar os fatores associados aos óbitos por COVID-19 nos municípios do nordeste brasileiro em dois períodos distintos da pandemia: julho de 2020 e abril de 2021, respectivamente. Para obtenção dos resultados, foram estimados modelos logísticos hierárquicos com variáveis a nível individual e municipal. Os resultados indicaram que a probabilidade de óbito é maior entre os indivíduos do sexo masculino, com comorbidades, idosos e com menos escolaridade. Ademais, aqueles que receberam a vacina antiviral ou contra a COVID-19 também tiveram menor probabilidade de óbito. Assim, para o caso dos municípios do Nordeste, as características individuais foram mais relevantes para a compreensão dos óbitos durante a pandemia do que as características do próprio município.
Palavras-chave: pandemia de Covid-19; modelos hierárquicos; região nordeste

Abstract
In 2020, the world was taken by surprise by the COVID-19 pandemic, caused by the SARS-CoV-2 virus, triggering a global humanitarian, economic, and social crisis. This study aimed to analyze the factors associated with COVID-19 deaths in municipalities in the Brazilian Northeast during two distinct periods of the pandemic: July 2020 and April 2021, respectively. To obtain the results, hierarchical logistic models were estimated with individual and municipal-level variables. The results indicated that the probability of death is higher among males, individuals with comorbidities, the elderly, and those with lower educational levels. Furthermore, individuals who received antiviral or COVID-19 vaccines also had a lower probability of death. Thus, for the case of Northeastern municipalities, individual characteristics were more relevant in understanding deaths during the pandemic than the characteristics of the municipalities themselves.
Keywords: Covid-19 pandemic; hierarchical models; northeast region

Citation: Leite, A. L., Lopes, T. H. C. R., & Esperidião, F. (2025). Fatores associados aos óbitos por Codiv-19 no Nordeste brasileiro: uma abordagem multinível. *Gestão & Regionalidade*, v. 41, e20259086. https//doi.org/10.13037/gr.vol41.e20259086

1 Introduction

Since the COVID-19 pandemic caused by the SARS-CoV-2 virus, several researchers have delved into understanding the factors that have contributed most to deaths caused by the novel coronavirus. Factors such as age, vaccination, and the presence of comorbidities emerged as the most relevant elements in increasing the likelihood of mortality. However, this article suggests that the characteristics of the environment in which individuals find themselves may also help to understand this issue (Santos & Teixeira, 2022).

The COVID-19 pandemic disproportionately affected economically and socially vulnerable groups. Due to the economic development process of the country, Brazil is marked by profound economic, social, and cultural differences (Cano & Guimarães Neto, 1986) which have to be considered in this study. The analysis proposed here will be restricted to municipalities in the Northeast Region of the country. This region still exhibits fragile socioeconomic indicators that could amplify the chances of death during the pandemic.

As an example of a socioeconomic indicator, we can mention nutrition. According to IBGE (2019), in the Northeastern states, 46.1% of households are below the national average (69.8%) in terms of adequate nutrition. Additionally, in terms of Per Capita Monthly Average Income, the Northeast region has the lowest income levels in the country and a high income concentration. The Gini Index of the states in the Northeast Region varies between 0.5 and 0.6, indicating a high degree of income inequality, as the closer it is to 1, the greater the inequality.

The level of human capital of individuals deserves to be highlighted as a factor capable of reducing mortality in the most critical periods. Increasing the level of education contributes to reducing social inequality between classes and regions of the country (Cruz, Teixeira, & Braga, 2010). According to Balassiano et al. (2005), the ease of finding a job has been directly associated with the issue of professional qualification, represented by a set of attributes that include aspects related to formal education, the ability to continually learn, entrepreneurship and a set of attitudes such as initiative, autonomy and versatility. These attributes would guarantee the employability of workers, that is, the ability to remain in the job market.

More specifically, this study also investigates whether a higher level of human capital among individuals in Northeast Region municipalities is associated with a lower COVID-19 mortality rate. It is expected that the higher the level of human capital among individuals, the lower the case fatality rate caused by COVID-19, as individuals with higher levels of education tend to have healthier behaviors, such as physical exercise, better nutrition, greater hygiene care, and regular preventive check-ups. Furthermore, individuals with better professional qualifications would have a greater chance of working from home, thereby avoiding greater exposure to the virus. Additionally, the beneficial effects of human capital are multidimensional: they can have positive impacts on macroeconomic, individual, and social levels, all of which may contribute to a lower mortality rate in the region (Besarria et al., 2016; Campos; Miranda, 2021).

The results indicate that the probability of death is higher among males, individuals with comorbidities, the elderly, and those with lower educational levels. Furthermore, those who received antiviral or COVID-19 vaccines also had a lower probability of death.

This paper is divided into three sections, in addition to this introduction and the concluding remarks. The second section provides a review of empirical literature on the determinants of COVID-19. The following section presents the methodology used, as well as the data source and treatment. Finally, the last section estimates the econometric models and discusses the main findings.

2 Empirical Literature Review

The COVID-19 pandemic has triggered a series of events that have highlighted how socioeconomic factors have been associated with the course of infectious diseases throughout history. In the Brazilian context, disparities in access to treatment have become a crucial factor in determining the disease's lethality rate. Some empirical studies, which will be presented throughout this section, reveal significant impacts of various characteristics on the increased risk of death, including older age groups, individuals of Black and mixed race origins, as well as those with lower levels of education. The study conducted by Batista et al. (2020) shed light on these disparities, analyzing the variation in the lethality rate in Brazil concerning socioeconomic factors. Furthermore, as will be observed, municipalities with low Human Development Index (HDI) demonstrated higher case fatality rates. These findings not only resonate in a global context, as demonstrated by studies in Detroit, Michigan (Laster Pirtle, 2020), and in countries worldwide (Chaudhry et al., 2020) but also emphasize the urgent need for effective actions to reduce socioeconomic and regional inequalities in pandemic response.

Socioeconomic factors have been important in explaining the evolution of pandemics throughout history. In the Brazilian case, in particular, the case fatality rate was influenced by disparities in access to treatment (Batista et al., 2020). Batista et al. (2020) aimed to analyze the variation in the disease's lethality rate in Brazil, taking into account socioeconomic factors such as age, gender, race, HDI, education, and the city where the case was recorded. Through a descriptive analysis, it was found that deaths are influenced by age - in age groups over 60, more than 50% of cases resulted in death. In terms of race/ethnicity, the proportion of Black and mixed-race individuals who died (54.78%) was higher than that of White individuals (37.93%). Regarding education, there was a lower proportion of deaths (22.5%) among people with higher levels of education than those without complete schooling (71.3%), reflecting that higher education levels correlate with lower disease lethality. Additionally, the research also showed that the percentage of deaths among Black and mixed-race individuals was higher than that among White individuals in all education and age groups. Finally, it was observed that in municipalities with low or medium HDI, the chance of an infected patient dying was nearly double that in municipalities with high Human Development Index.

A similar study was conducted in the city of Detroit, Michigan, in the United States. In the city, only 14% of the population is Black, however, statistics confirmed that 40% of direct COVID-19-related deaths were among Black individuals, highlighting racial and socioeconomic disparities since most of those in this group are socioeconomically vulnerable (Laster Pirtle, 2020).

Using data from the Brazilian National Health Survey (PNS) conducted by IBGE in 2013, Pires, Carvalho, and Xavier (2020) estimated the proportion of Brazilians who fell into the COVID-19 risk group. People over 60 years old with comorbidities were classified as part of the risk group. The research showed that the incidence of comorbidities is higher in Brazilians who only completed elementary education compared to other groups. Thus, the research makes it clear that the low-income population is the most vulnerable to the epidemiological and public health crisis.

Along similar lines, using the same database, Borges and Crespo (2020) aimed to characterize COVID-19 risk groups in Brazil and estimate the number of people living in the same household as individuals in the risk group. Logistic regression was used for estimation. The study took into account age, gender, region, race, education level, and labor force status of the survey respondents. The results highlighted that age is the primary risk factor for comorbidities associated with COVID-19, but a higher risk for people in more vulnerable categories, such as those with lower education levels and Black and mixed-race individuals was also observed. Moreover, the estimation revealed that 68.7% of Brazilians lived with at least one person in the risk group.

Souza et al. (2021) analyzed the risk factors for mortality among hospitalized COVID-19 patients in Brazil from February 26 to August 10, 2020. The study showed that the lethality rate was higher in elderly individuals of both sexes, with slightly higher lethality among males. Furthermore, a high lethality rate was observed in the North/Northeast region, in older age groups, in non-White populations, and among those with lower education levels. According to the authors, the higher lethality in the North/Northeast region may be due to socioeconomic conditions and the availability of ICU beds. On the other hand, it may also be due to inadequate knowledge about the characteristics of the new disease, which was more prevalent in this region at the beginning of the pandemic.

Pinheiro et al. (2020) aimed to identify the relationship between regional characteristics and epidemiological and social factors in COVID-19 mortality in Brazil. Their results revealed differences among Brazilian regions, exposing inequalities in access to healthcare services and they claimed that the epidemiological and social profile contributed to increased lethality rates in the North and Northeast regions. Therefore, the research reinforces the need for effective actions to reduce regional inequalities.

The COVID-19 pandemic has also exacerbated disparities in access to healthcare systems, increasing racial differences and worsening health outcomes in these population groups (PERES et al., 2021). In this regard, Peres et al. (2021) aimed to analyze the association between sociodemographic characteristics and hospital mortality due to COVID-19 in Brazil. For the analysis, a logistic regression model was used. The study included adult patients hospitalized with COVID-19 with a defined outcome (228,196) from February 16 to August 8, 2020. The research revealed that the overall mortality rate was 37%, with Black/mixed-race patients having a mortality rate of 42% compared to 37% for White patients. Furthermore, Black/mixed-race individuals were less frequently admitted to intensive care units (32%) compared to White individuals (36%) and more invasive mechanical ventilation (21%) was used compared to self-identified White patients (19%). Given these results, the need to implement active strategies to reduce disparities caused by broader health determinants becomes evident, leading to a sustainable change in the healthcare system (PERES et al., 2021).

Chaudhry et al. (2020) conducted an exploratory analysis of the top 50 countries in reported cases to assess the impacts of public health interventions, socioeconomic factors in each country, and the number of deaths. The results showed that a higher number of cases and overall mortality were associated with comorbidities such as obesity and an older population. In contrast, a lower income inequality within a country reduced overall mortality and critical cases. Additionally, the government policy of total lockdown was positively related to recovery rates, demonstrating that a complete lockdown and early border closures could decrease transmission peaks and consequently prevent healthcare system overload, facilitating higher recovery rates. However, countries with higher per capita GDP experienced an increase in critical cases and deaths per million inhabitants. According to the authors, this could be associated with more extensive population testing, greater data transparency, and better surveillance systems. However, the researchers also highlight that this result may be associated with international air travel and holidays taken by people in wealthier countries.

In an attempt to analyze the prevalence of COVID-19 infection in the state of Pará, Silva (2021) observed that the virus does not affect the population in a democratic manner. The research showed that the spread of COVID-19 in the state occurs among the most impoverished (social classes C, D, and E) and those with low levels of education (elementary to high school). Regarding race/ethnicity, the research identified less difference in the proportions of people with antibodies. It can therefore be concluded that in the state, social class and education level delineated a highly unequal distribution of infection risks. Furthermore, the research asserts that these are the same groups that make up the workforce most exposed to contracting the virus through their often informal jobs, which are performed outside the home, intensifying interpersonal contact.

Barbosa, Costa, and Hecksher (2020) identified which workers were most impacted in terms of job loss in Brazil during the economic crisis caused by the COVID-19 pandemic. The study revealed that the groups most likely to lose their jobs were young people and women, approximately 20%. Notably, the proportion of Black and mixed-race individuals losing their jobs consistently exceeded the average and reached 18% at the beginning of the crisis (an increase of 5 percent compared to 4 percent among White individuals). Concerning the education level, people with an incomplete high school education or less had a 15% chance of transitioning to unemployment or leaving the labor force. With regard to job positions, workers with part-time jobs, informal employment, and lower wages were highlighted among those who experienced significant losses.

In summary, the studies presented throughout this section underscore the role of socioeconomic factors in the progression of pandemics, with a focus on the COVID-19 pandemic in Brazil. Inequalities in access to treatment emerge as a critical factor that directly influences disease lethality rates. These disparities manifest in various ways, including significant differences in death rates among age groups, racial and educational groups, as well as in regions with a lower Human Development Index (HDI). The combined analysis of these studies reinforces the need for the development of strategies and public policies aimed at reducing social and regional inequalities so as to promote a more equitable and effective approach to addressing public health crises.

2.1 Interiorization of the COVID-19 Pandemic in Brazil

As previously discussed, the COVID-19 pandemic triggered a series of significant public health challenges and social impacts. However, in addition to major metropolitan areas and densely populated urban areas that initially received the most attention, the pandemic has also had a profound influence on interior regions. Less densely populated areas, often considered distant from urban epicenters, did not escape the consequences of the spread of the virus. In this section, we will explore the dynamics of the interiorization of COVID-19, examining how areas farther from major urban centers also faced significant challenges related to health, infrastructure, and access to medical care.

Silva et al. (2020) conducted an exploratory analysis of the increase in mortality attributed to the COVID-19 pandemic in Brazil, covering the period from March to May 2020. This investigation encompassed both the capital cities and municipalities of the country, using data from civil registry death records. The study compared the number of deaths actually recorded with the expected number of deaths in 2019, disaggregating the information by gender and locality. The authors observed a significant increase in standardized mortality ratios in the municipalities in the interior of the North, Northeast, and Southeast regions, which occurred primarily in May 2020. The mortality profile initially changed in the capitals of these regions, but as the epidemic progressed, there was an excess of deaths in municipalities outside the capitals.

Some studies have focused on assessing the interiorization of the pandemic in certain states in the Northeast. For example, Pedrosa and Albuquerque (2020) analyzed the distribution of COVID-19 cases and specific ICU beds for the disease in the state of Ceará using Geographic Information Systems resources. The authors conducted an exploratory analysis of the spatial distribution of COVID-19 cases recorded between March 15, 2020, and April 18, 2020. The Moran's Index was calculated to assess the spatial dependence of the distribution of the crude detection coefficient, global and local Bayesian rates. According to these authors, the interiorization of the pandemic refers to the spread of the disease to areas where cases had not been previously registered. In the case of Ceará, the interiorization of COVID-19 was observed through the identification of clusters with a High-High pattern of the disease in regions outside the capital and metropolitan area, suggesting a broader spread of the disease in the state of Ceará. In general, this spread may have been caused by various factors, such as the travel of infected individuals, gatherings at social events, increased mobility on roads, among others. These causes are common in many other regions of the world where the pandemic rapidly spread to previously disease-free areas. Understanding the interiorization of pandemics is important for understanding the propagation of the disease and for decision-making regarding the allocation of specific resources, such as ICU beds for patients, to better meet the needs of affected regions.

This is a challenge in responding to pandemic control in the more remote regions of the country, given the rapid spread in small municipalities in Brazil. Research such as that of Gomes et al. (2021) highlight that the distribution of COVID-19 occurred heterogeneously in health regions, with the initial cases identified in Brazilian capitals and subsequently new cases detected in more distant regions. By describing the epidemiological profile and spatial distribution of deaths and confirmed cases in the Western health macroregion of Bahia, the authors believe that the interiorization of COVID-19 may have impacted the healthcare system in the macroregion, as many municipalities do not even have a hospital. Therefore, the interiorization of pandemics is a concerning phenomenon, as it can impact regions where the availability and quality of healthcare services may be limited.

Quinino et al. (2021) conducted a spatial and temporal analysis of the incidence and interiorization of COVID-19 in the state of Pernambuco, Brazil, between March and June 2020, while also examining related socioeconomic factors. According to the authors, the COVID-19 pandemic had a disproportionate effect on the metropolitan and rural interior regions of Pernambuco, showing a concentration of cases in urban areas. The first cases were recorded in municipalities in the metropolitan region, with its spread to the interior occurring later. They identified significant spatial clusters of municipalities with high and low incidences, as well as transitional municipalities. Municipalities with the highest incidence coefficients with neighbors with high rates were concentrated in the Metropolitan Region of Recife (RMR), while those with low incidence coefficients and neighbors with low rates were in the Sertão and São Francisco mesoregions. The main risk areas for COVID-19 in Pernambuco were located in the Metropolitan Region of Recife, while the most affected municipalities were in the interior of the state, such as Caruaru and Palmares, which had the highest incidence coefficients.

The aforementioned study suggests some reasons for the interiorization of the pandemic, such as: i) the introduction of the disease in cities where population growth was not accompanied by urban development; ii) pendular migration of people to cities with better living conditions and where infrastructure, healthcare services, and education are concentrated; iii) socioeconomic differences between urban and rural populations and lack of access to information and quality healthcare services; and iv) social inequalities, where the impact of the pandemic falls more problematically and adherence to recommendations is compromised. Quinino et al. (2021) point out some consequences of the interiorization of COVID-19, such as the limited capacity of healthcare services to provide intensive care in the interior. Furthermore, as rural areas often house historically disadvantaged populations, there are greater difficulties in accessing adequate healthcare services. Consequently, the increase in COVID-19 cases in the interior can lead the pandemic to progress silently, without detection, and eventually evolve into severe situations and healthcare system collapse.

3 Data and Methodology

To achieve the objectives of this study, a hierarchical model, specifically a hierarchical logistic regression model, will be adopted. This model is suitable because it takes into account the data's clustering structure. In other words, in addition to the importance of individual characteristics, it considers the specificities of the environment in which the individual is located, as these specificities can affect the probability of the individual's survival in the context of infection caused by the SARS-CoV-2 virus (Soares, Mendonça, 2003, & Lopes, 2015).

In hierarchical models, behavioral and social data often have a clustered structure and each level in the structure is formally represented by its own submodel. These submodels express relationships between variables within a given level and explicitly show how they influence relationships that occur at another level (Raudenbush & Bryk, 2002).

In specifying the multilevel model, both the intercept and the slope coefficient are considered random parameters, dependent on the influence of the higher hierarchical level (Soares & Mendonça, 2003). It is assumed that the data is hierarchically structured into two levels: the first level consists of individuals who are nested within the second level, composed of municipalities in the Northeast of Brazil.

As highlighted by Lopes (2015), the variance component model implies random effects, such that the variation in the intercept is captured by the variance in the second-level residuals. Estimating this model will allow us to assess whether the hierarchical approach is indeed necessary. If the need for such an approach is confirmed, the next step is to include first-level variables with random intercepts. Specifically,

$$y_{ij} = \beta_{0j} + \beta_1 X_{ij} + e_{ij} \tag{1}$$

In which

$$\beta_{0j} = \beta_0 + u_{0j} \tag{2}$$

Note that the row of each unit at the second level is given by:

$$y_j = (\beta_0 + u_{0j}) + \beta_1 X_{ij} \tag{3}$$

In which $(\beta_0 + u_{0j})$ i s the specific intercept of the second-level units. It should be noted that the slope coefficient is, for now, common to all second-level units; β_0 and β_1 are fixed coefficients of the regression ; u_{0j} and e_{ij} are random effects or multilevel residuals. The random parameters mentioned earlier, $\sigma_u^2 + \sigma_e^2$, should be estimated together with the fixed coefficients of the regression (LOPES, 2015).

The hierarchical model to be estimated is indeed a logit, as the dependent variable is binary. Therefore, the random intercept model will look something like this:

$$log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \beta_0 + \beta_1 x_{ij} + u_j \tag{4}$$

In which $u_j \sim N(0, \sigma_u^2)$.

Considering that we have a first-level explanatory variable, x_{1ij}, and a second-level variable, x_{2j}, the random intercept logit model represented by equation 4 can be extended to include both predictors:

$$log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \beta_0 + \beta_1 x_{ij} + \beta_2 x_{2j} + u_j \quad (5)$$

Investigations that incorporate the data clustering structure into their models have several advantages, namely: they are based on more flexible and structured models that make better use of the information present in the sample; the utilization of data clustering information allows for the formulation and testing of hypotheses regarding effects between levels (SOARES; MENDONÇA, 2003)

Therefore, equation 5 can be described as follows:

$$Deaths\_Cov_{it} = A_i + B_1 Educat\_Level_{it} + B_2 race_{it} + B_3 age_{it} + B_4 sex_{it} + B_5 risk\_factor_{it} + B_6 vaccine\_COV_{it} + B_7 IDHM_{it} + B_8 Pop\_water\_piped_{it} + B_9 Pop\_garbade\_collection_{it} + B_{10} GPDpc_{it} + \varepsilon_{it} \quad (6)$$

The description of each variable is provided in Table 1 in the following section.

3.1 Data Source and Data Processing

Three databases were utilized in this study. The first one is derived from daily information obtained from the Severe Acute Respiratory Syndrome Database - including COVID-19 data (SRAG 2020), which is populated by notifications made by public and private healthcare units and made available by the Ministry of Health (OpenDataSUS). The data was updated on 11/24/2021 and 11/29/2021, referring to the years 2020 and 2021, respectively. It is worth noting that, despite the fact that this database provides information on Severe Acute Respiratory Syndrome, only the data related to COVID-19 was selected for this research.

The second source was the Human Development Atlas, from which it was possible to obtain information on the Municipal Human Development Index (MHDI), as well as data related to sanitation and waste collection, which will be used as control variables in the econometric model. The research focused on municipalities in the Northeast region, totaling 1793.

Finally, variables related to per capita GDP were collected from the Brazilian Institute of Geography and Statistics (IBGE). The description of the variables used is presented below:

OUTCOME: This refers to the outcome of the case, whether the patient died or not due to complications caused by COVID-19. The data was obtained from the OpenDataSUS website.

RISK_FACTOR: This variable indicates whether the patient had any comorbidities. Risk factors are associated with higher rates of COVID-19-related deaths, as shown in the empirical part of this study. The data was obtained from the OpenDataSUS website.

AGE: This indicates the patient's age. Like risk factors, age is directly related to a higher prevalence of COVID-19-related deaths. The data was also extracted from the OpenDataSUS website.

RACE: Represents the race/ethnicity declared by the patient. The aim was to investigate whether race/ethnicity contributed to a higher or lower rate of COVID-19 deaths.

EDUCAT_LEVEL: Represents the level of education reported by the patient. As a proxy for human capital, the goal is to verify the inverse relationship between human capital and COVID-19 deaths. It is expected that higher levels of this capital will result in lower COVID-19 death rates. The data was also obtained from the OpenDataSUS website.

SEX: Indicates the patient's gender, male or female. The purpose is to examine the relationship of this variable with the higher likelihood of COVID-19-related deaths.

VACCINE: Indicates whether the patient received the flu vaccine in the last campaign. The intention here is to verify if antiviral immunization contributes to reducing the probability of deaths caused by SARS-CoV-2.

VACCINE_COV: Indicates whether the patient was immunized with the COVID-19 vaccine. It is assumed that the vaccine has contributed to a lower death rate from the disease.

IDHM: A succinct indicator of progress in three basic dimensions of human development: income, education, and health. The inclusion of this variable is related to the hypothesis that municipalities with a better HDI correlate with a lower COVID-19 death rate. The data was obtained from the Human Development Atlas in Brazil (2010).

ln GDPpc: Natural logarithm of per capita GDP at current prices, base year 2018. The data was obtained from the IBGE website.

POP_WATER_PIPED: Percentage of the population in households with piped water. The purpose is to verify whether better sanitation conditions contribute to a lower COVID-19 death rate. The data was extracted from the Human Development Atlas in Brazil (2010).

POP_GARBADE_COLLECTION: Percentage of people in urban households with garbage collection. The purpose is also to verify whether better sanitation conditions contribute to a lower COVID-19 death rate. The data was extracted from the Human Development Atlas in Brazil (2010).

Table 1 provides a summary of the model variables, some studies that made use of the same variables in their statistical analyses, and their expected signs in relation to the probability of COVID-19-related deaths.

Table 1 – Variables and authors who used the same variables in their studies.

Variable	Source	Studies that analyzed the variables in question	Expected sign in relation to the probability of COVID-19 death
Education Level	OpenDataSUS (2021)	Batista et al. (2020); Pires, Carvalho e Xavier (2020); Borges e Crespo (2020); Pinheiro et al. (2020)	(-) Increasing educational attainment is associated with a lower probability of COVID-19-related mortality.
Race/Ethnicity	OpenDataSUS (2021)	Batista et al. (2020); Borges e Crespo (2020); Pinheiro et al. (2020); Peres et al. (2021)	(+) Blacks and Indigenous individuals are expected to have a higher probability of COVID-19 mortality compared to Whites
Age	OpenDataSUS (2021)	Batista et al. (2020); Pires, Carvalho e Xavier (2020); Borges e Crespo (2020); Pinheiro et al. (2020)	(+) The older the age, the higher the probability of COVID-19 mortality

SEX_Gender (male)	OpenDataSUS (2021)	Borges e Crespo (2020); Pinheiro et al. (2020)	(+) Men are expected to have a higher probability of COVID-19 mortality compared to women
Risk Factor (comorbidity)	OpenDataSUS (2021)	Pires, Carvalho e Xavier (2020); Borges e Crespo (2020); Pinheiro et al. (2020); Chaudhry et al. (2020).	(+) Individuals with comorbidities are expected to have a higher probability of COVID-19 mortality compared to those without comorbidities
COVID Vaccine	OpenDataSUS (2021)	Rossman et al. (2021)	(-) vaccinated individuals are expected to have a lower probability of COVID-19 mortality compared to those who are not vaccinated
IDHM: Municipal Human Development Index	Human Development Atlas in Brazil (2010)	Batista et al. (2020)	(-) Individuals residing in cities with a higher MHDI are expected to have a lower probability of COVID-19 mortality
GDP per capita	IBGE (2018)	Chaudhry et al. (2020)	(-) Individuals residing in cities with a higher MHDI are expected to have a lower probability of COVID-19 mortality
Population with piped water	Human Development Atlas in Brazil (2010);	-	(-) Individuals residing in cities with greater access to piped water are expected to have a lower probability of COVID-19 mortality
Population with garbage colletion	Human Development Atlas in Brazil (2010);	-	(-) Individuals residing in cities with greater access to garbage collection are expected to have a lower probability of COVID-19 mortality

Source: Self-developed based on the presented literature

3.2 Missing Data and Imputation

The dataset in this study contains missing values, which were imputed using the multiple imputation process. As highlighted by Rubin (1987), multiple imputation is a technique that replaces each missing value with two or more acceptable values, representing a distribution of possibilities. This method consists of three stages: i) m sets of complete datasets are created, and each missing value is filled in m times through an appropriate imputation model, given the observed values; ii) the m imputed complete datasets are analyzed using standard complete data procedures, as if the imputed data were the actual data obtained from non-respondents; and iii) the m results obtained from the imputed complete datasets are combined in a simple and appropriate manner to obtain what is called repeated imputation inference (Rubin, 1987; Zhang, 2003; Nunes, 2007).

Among the multiple imputation methods, the Multivariate Imputation by Chained Equations (MICE) method was used. According to Van Buuren (2007), this method assumes that for each variable with missing data, the user specifies a conditional distribution for the missing values given the observed values. For example, for incomplete binary variables, logistic regression can be used, for categorical data, polynomial regression, and for continuous data, linear regression, as was done in this research and specified below. Assuming that there is a multivariate distribution from which these conditional distributions can be derived, MICE constructs a Gibbs sampler from the specified conditionals. This sampler is used to generate multiple imputations (Van Buuren, 2007).

Due to the particularities of the variables analyzed in this research, the regression models were adjusted for each specific variable, for the two analyzed periods, July 2020 and April 2021, which correspond to the peaks of infection in each year. The imputed variables were: outcome, race, education level and age for both periods; vaccine and COVID vaccine for July 2020 and April 2021, respectively.

To impute missing data for the AGE variable, a linear regression model was estimated, as it is a continuous variable. This variable was regressed against the following variables: age, outcome, race, education level, vaccine, logarithm of GDP per capita, risk factor, gender, and population with piped water.

Regarding the OUTCOME variable, a logistic regression model was estimated because it is binary, referring to the case outcome (death/no death). To impute the missing values of the RACE variable, a multinomial logistic regression model was estimated since it is a categorical variable, meaning that there is no logical ordering between the category classifications.

In order to impute the missing values of the education level variable, an ordinal logistic regression was estimated, where data is presented in categories with ordering (Abreu, Siqueira, & Caiaffa, 2009). The imputation of missing data for the VACCINE and VACCINE_COV variables was done using logistic regressions since they are binary variables. Once the missing data were imputed, the next step was to estimate the econometric models, the results of which will be presented and discussed in the following section.

4 Results and Discussion

When considering the data for the peak contamination periods, July 2020 and April 2021, it can be observed that the average age of individuals was approximately 60 and 57 years, respectively. It is worth noting that during these periods, 36.8% and 35.4% of cases resulted in death, respectively. The HDI of the municipalities in the Northeast Region analyzed in this study had an average of approximately 0.68 for both periods considered. This indicator deserves attention because it may be related to the higher number of COVID-19 deaths, as observed by Alberti et al. (2021).

Furthermore, in July 2020, 64.89% of individuals had some risk factor, while in the second analysis period, April 2021, 59.17% of people had one or more comorbidities. It is important to note that these risk factors are directly related to a higher number of deaths due to COVID-19 complications, as presented earlier.

The table below shows the results of the econometric procedures, with OUTCOME as the dependent variable. This is a binary variable where 1 indicates that the individual died, and 0 indicates otherwise. It is important to highlight that there are four models in Table 1. The first two models consider the data for the entire year of 2020 - Complete Data 2020 - and the entire year of 2021 - Complete Data 2021, both without considering missing data. The last two models - Imputation 2020 and Imputation 2021 - contain data only for the peak period of the pandemic, July 2020 and April 2021, respectively, and missing data were imputed as explained in the previous section. The objective is to verify the robustness of the variables.

Table 2: Hierarchical Logistic Model.

Dependent variable = Deaths due to COVID-19								
Explanatory Variables	Complete data 2020		Complete data 2021		Imputation 2020		Imputation 2021	
Vaccine	-0.43***	-0.43***	-0.68***	-0.67***	-0.25***	-0.25***	-0.52***	-0.52***
	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.06)	(0.06)
Risk factor	0.55***	0.55***	0.68***	0.68***	0.35***	0.35***	0.47***	0.47***
	(0.05)	(0.05)	(0.09)	(0.06)	(0.04)	(0.04)	(0.03)	(0.03)
Education								
2	0.06	0.05	-0.14**	-0.15***	0.10	0.10	-0.14*	-0.14*
	(0.08)	(0.08)	(0.05)	(0.05)	(0.09)	(0.09)	(0.07)	(0.07)
3	-0.37***	-0.37***	-0.41***	-0.42***	-0.28**	-0.27**	-0.31***	-0.32***
	(0.05)	(0.05)	(0.07)	(0.07)	(0.09)	(0.09)	(0.08)	(0.08)
4	-0.76***	-0.75***	-0.61***	-0.62***	-0.61***	-0.60***	-0.47***	-0.48***
	(0.07)	(0.08)	(0.14)	(0.14)	(0.15)	(0.15)	(0.12)	(0.12)
Age	0.04***	0.04***	0.03***	0.03***	0.04***	0.04***	0.03***	0.03***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Sex	-0.33***	-0.33***	-0.10**	-0.10**	-0.15***	-0.15***	-0.08**	-0.08**
	(0.04)	(0.04)	(0.05)	(0.05)	(0.03)	(0.03)	(0.03)	(0.03)
Race								
2	-0.03	-0.04	0.08	0.07	-0.01	-0.01	-0.08	-0.08
	(0.12)	(0.12)	(0.14)	(0.14)	(0.10)	(0.10)	(0.09)	(0.09)
3	-0.46**	-0.46**	-0.07	-0.08	-0.43**	-0.43**	-0.06	-0.05
	(0.19)	(0.19)	(0.30)	(0.30)	(0.16)	(0.16)	(0.15)	(0.15)
4	-0.03	-0.03	-0.02	-0.02	-0.21***	-0.21***	-0.24***	-0.24***
	(0.07)	(0.07)	(0.13)	(0.13)	(0.06)	(0.06)	(0.04)	(0.04)
5	0.07	0.04	0.22	0.23	0.18	0.17	-0.07	-0.06
	(0.60)	(0.61)	(0.45)	(0.45)	(0.48)	(0.48)	(0.52)	(0.52)
LnGPD		0.10		-0.04		0.03		-0.14*
		(0.12)		(0.10)		(0.08)		(0.08)
IDH		-3.75***		1.14		-2.90***		2.42**
		(1.27)		(1.35)		(0.90)		(0.95)
Water		0.01**		0.00**		0.00***		-0.00***
		(0.00)		(0.00)		(0.00)		(0.00)
garbage colletion		0.00		-0.00***		0.00*		-0.00***

Dependent variable = Deaths due to COVID-19								
Explanatory Variables	Complete data 2020		Complete data 2021		Imputation 2020		Imputation 2021	
		(0.00)		(0.00)		(0.00)		(0.00)
Constant	-3.11***	2.79***	-2.38***	-3.20***	-2.64***	-2.50***	-2.17***	-1.84***
	(0.14)	(1.01)	(0.17)	(0.88)	(0.12)	(0.06)	(0.14)	(0.62)
OBS	11323	11323	11759	11759	17475	17475	21832	21832
VAR	0.77	0.75	1.08	1.05	0.37	0.34	0.42	0.42

Note: (***), (**), (*) mean that they are statistically significant at 1%, 5% and 10%, respectively.
Robust standard errors in parentheses
OBS = number of observations
VAR = variance of random intercept

The results indicate that the VACCINE variable was negative and statistically significant in all models, indicating that those who were vaccinated had a lower probability of dying from Covid-19 than those who were not vaccinated. It is important to emphasize that for the year 2020, " VACCINE " corresponds to the flu vaccine, while in 2021, it refers to the Covid-19 vaccine. Note that the parameters for 2021 are much higher, suggesting the effectiveness of the Covid-19 vaccine. Those who were vaccinated against the flu in 2020 had a lower chance of dying than the unvaccinated. This may reveal that vaccinated individuals are more likely to take preventive measures than the unvaccinated which is consistent with findings in the literature. For example, Fink et al. (2021) found that patients who received a recent influenza vaccine had, on average, a 7% lower chance of needing intensive care, a 17% lower chance of needing invasive respiratory support, and a 16% lower chance of death. The authors also observed that the protective effects were greater when the vaccine was administered after the onset of symptoms, as well as among younger patients.

The development of vaccines represents a significant advance in combating diseases, with the primary goal of stimulating the body's immune defenses (Bousada & Pereira, 2017). The results obtained in this study indicate a lower probability of deaths among those who received both the antiviral vaccine and the Covid-19 vaccine. In a preliminary analysis conducted in the state of Mato Grosso, even with the vaccination program in its initial stages, evidence pointed to the effectiveness of the Covid-19 vaccine in reducing mortality among the priority group of individuals over 70 years old (Oliveira et al., 2021). Furthermore, other preliminary studies reinforce the thesis of vaccine effectiveness in reducing deaths from the SARS-CoV-2 virus, as can be seen throughout this chapter.

It was also observed that individuals with comorbidities (RISK FACTORS) and older individuals (AGE) had a higher probability of dying from Covid-19. Both variables were positive and significant at the 1% significance level. In addition, gender (SEX) also proved to be significant, indicating that women had a lower probability of dying from Covid-19 than men. Both of these results remained consistent in all models and are therefore robust. This finding has also been confirmed by Souza et al. (2021). The explanations for this finding are partly related to the immune response that each biological sex presents – women generally have a more robust immune response to infections, vaccines, and some malignant diseases. On the other hand, men have higher rates of comorbidities, such as cardiovascular diseases, chronic lung diseases, and hypertension, which are risk factors for the exacerbation of Covid-19 complications (Lotter & Altfeld, 2019; Moreira & Oliveira, 2020).

It was found that men have the highest probability of dying from Covid-19. Gender disparities in disease outbreaks and health outcomes involve both physical mechanisms, such as sex-based biological factors that modulate the host's immune response, and socially constructed profiles that include social, behavioral, and lifestyle aspects. Covid-19 is no exception when it comes to gender differences, which manifest in the vulnerability and severity of the disease and access to healthcare services (OPAS, 2021). Considering the biological mechanisms that combat viral infections, women, in general, tend to produce a more effective and adaptive immune response to viruses, contributing to a less severe course of Covid-19 (OPAS, 2021). Another factor associated with a higher mortality rate from Covid-19 among males is the existence of previously underdiagnosed diseases, such as hypertension, diabetes, and cardiovascular disease, which are, in theory, more prevalent among men (Porto et al., 2021). Additionally, men tend to engage in higher-risk activities that increase the potential for contracting the disease and are more likely to underestimate the severity of the virus's potential to harm them (Griffith et al., 2020).

With regard to race, those who identified as yellow and brown had a lower probability of dying from Covid-19 than those who identified as white, but these results were not robust. Therefore, it is not possible to ensure their real association with the likelihood of death. Porto et al. (2021), in a study aimed at identifying Covid-19 mortality in Brazil in the first 6 weeks after the confirmation of the first death case, found higher mortality rates among white individuals. However, other studies indicate that lethality has been higher among brown and black individuals (De Negri et al., 2021; De Lima & Lima, 2021; Galvão & Roncalli, 2021). As evidenced by Porto et al. (2021), there are still no hypotheses for such findings.

Finally, the variable of interest – level of education – is in line with expectations. That is, for individuals with higher levels of education, it was negative and significant in all models. This indicates that higher levels of education are associated with a lower probability of dying from Covid-19. This may be a consequence of different socioeconomic conditions, which were not captured by the model, as well as a greater knowledge about prevention and containment of the virus, in addition to the ability to adhere to social isolation measures, such as working remotely (De Negri et al., 2021). Furthermore, this result is consistent with findings in the literature, such as those by Batista et al. (2020); Pires, Carvalho, and Xavier (2020); Borges and Crespo (2020); Pinheiro et al. (2020).

When evaluating the second-level variables, greater attention should be given to the IDHM. In 2020, those who lived in municipalities with a higher IDHM had a lower probability of dying from Covid-19. This result is consistent with studies on the subject and was consistent in both the general model and the reduced model with imputation. Aiming to verify the association between a higher number of Covid-19 deaths and the HDI (Human Development Index) of cities in Santa Catarina, Alberti et al. (2021) found that cities with a poor HDI had a higher death rate compared to cities with a good HDI. In a study conducted for the state of Ceará, Araujo et al. (2020) observed that the predominance of cases and deaths occurred among the population in places with low and very low HDI. Thus, highlighting social difficulties.

However, there was a difference in relation to the year 2021. For this year, both in the model with total data and in the reduced model with imputation, the parameter related to IDHM was positive, although significant only in the imputation model, whose sample is restricted to the peak month of cases in that year. This result requires further investigation, which can be conducted in future studies. However, it is possible that there was a change in pattern from one year to another for the Northeastern municipalities. In the first wave, individuals living in municipalities with a higher IDHM had a lower chance of dying from Covid-19, but it is possible that this logic was reversed in the second wave due to the need to open activities in more developed cities.

As for basic sanitation and per capita income variables, they did not show statistical significance in both analyzed periods. In other words, there was no evidence of their importance in the context of the pandemic, despite studies by Ferreira, Silva, and Figueiredo Filho (2021), Chaudhry et al. (2020), Vasconcelos and Muylder (2022) indicating their importance. It is understood that to some extent such indicators are incorporated into the IDHM.

As mentioned, the Northeast Region is characterized by socioeconomic inequality and poverty, which results in inequalities among the population in general. Inequality of opportunities can place individuals in different socioeconomic positions, depending on their race, gender, social group, and sex, making it difficult to access education, work, and income. Socioeconomically vulnerable people tend to be exposed to the virus in different ways because people in these situations have poorer housing, live in overcrowded homes, use public transportation with more crowding, have occupations that make it difficult to adopt social distancing measures, and face food insecurity and restrictions in access to medical care, which contribute to higher death rates in this population subgroup (Demenech et al., 2020).

Populations with a low income are more exposed to infectious diseases, including SARS-CoV-2, due to lack of access to basic sanitation and treated water, lack of access to education and healthcare services, difficulty isolating, living in precarious homes, or being more frequently unemployed or engaged in informal jobs (Kerr et al., 2020). It is precisely in this subgroup that the less educated are located. It is because of this that Souza et al. (2020) argue that this pandemic has race, color, gender, and social class dimensions. A portion of the population has the privilege of benefiting from virus containment measures without major setbacks, while the more vulnerable portion is forced to expose themselves to the virus in the fight for survival.

In summary, the results presented highlight the importance of formulating and implementing public policies in the field of Public Health, particularly in the face of the possibility of new health crises. Furthermore, considering the vulnerability of these population segments, such policies have become crucial for mitigating social and regional inequalities, protecting the most vulnerable, and promoting equity in access to healthcare.

The analysis demonstrates that vaccination plays a crucial role in reducing mortality, underscoring the need for comprehensive and accessible immunization campaigns. Additionally, variables such as comorbidities, age, gender, and level of education play distinct roles in Covid-19 mortality, emphasizing the importance of specific and targeted approaches for these risk groups. Finally, the influence of the Municipal Human Development Index (IDHM) on mortality rates suggests the need for regionally sensitive policies to address socioeconomic disparities. In summary, addressing social and regional inequalities should be at the core of strategies for preparing for and responding to public health crises, ensuring that all citizens have equal access to protection and healthcare.

5 Final Remarks

The aim of this study was to assess the factors associated with COVID-19 deaths in municipalities in the northeastern region of Brazil. To obtain the results, hierarchical logistic models were estimated with individual and environmental-level variables. As can be seen, the results indicated a statistical significance of individual-level variables in almost all estimated models in both analyzed periods. Regarding environmental-level variables, only the Municipal Human Development Index (IDHM) presented a relatively higher parameter that was significant and as expected.

In line with the results obtained, women have a lower probability of dying from Covid-19 than men. Explanations for this finding are supported by biological, social, and cultural factors. Older individuals and/or those with comorbidities are among the groups with a higher probability of dying from COVID-19. These individuals have weakened immune systems, leading to an increased severity of infectious diseases. Here vaccines play a crucial role. As observed, individuals vaccinated with both the antiviral vaccine and the specific Covid-19 vaccine had a lower risk of death due to COVID-19 complications.

Regarding race, those who identified as brown and yellow had a lower probability of dying than those who identified as white. This result, although supported by the literature, differed from the majority of studies on the subject.

The parameters of variables reflecting the level of education of individuals were significant, indicating that those with higher levels of education have a lower risk of death from COVID-19. However, this result should be analyzed cautiously because it is not that the more educated individuals die less simply because they have a higher level of education, but rather because this segment of the population has higher purchasing power, jobs with adequate sanitary conditions, housing with satisfactory infrastructure, as well as better access to healthcare, quality nutrition, and healthy habits. Therefore, these individuals are able to adhere to isolation and social distancing measures without major harm to their survival, making them less susceptible to virus exposure.

Regarding second-level variables, the Municipal Human Development Index (IDHM) was significant for both 2020 and 2021 when using the total data. In other words, in municipalities with a better IDHM, individuals have a lower chance of dying from COVID-19. The other variables were either not significant or their parameters were significant but close to zero. Therefore, for municipalities in the Northeast, it cannot be stated that variables related to basic sanitation and per capita income had a significant impact on the chance of deaths from SARS-CoV-2. It should be noted that the result for 2021 showed discrepancies when using total and imputed data. Thus, these results are more robust for the first wave in 2020.

The occurrence of missing data is a problem in scientific research, especially in the fields of health and social sciences. The amount of missing data here constituted the main limitation of this study because the database provided by the Ministry of Health (OpenDataSUS) had a large amount of missing data, necessitating the use of statistical data imputation techniques. However, despite the limitations, satisfactory and coherent results were obtained with the literature on the subject, as observed throughout the text.

As an agenda for future work, it is suggested to consider variables associated with the prior health structure of municipalities in subsequent models in order to investigate their potential effects on pandemic-related deaths. Furthermore, this study did not capture regional heterogeneities as it focused on municipalities in the northeastern region of Brazil. Thus, comparing these results with those found in other regions can provide valuable insights for policymakers.

Therefore, in light of the results obtained, this study should support strategies and public policies aimed primarily at the economically and socially vulnerable groups that tend to suffer more in adverse crisis circumstances.

References

Abreu, M. N. S., Siqueira, A. L., & Caiaffa, W. T. (2009). Regresión logística ordinal en estudios epidemiológicos. *Revista de Saúde Pública, 43*(1), 183–194.

Alberti, A., et al. (2021). Associação do maior número de mortes por COVID-19 e o Índice de Desenvolvimento Humano (IDH) de Cidades Catarinenses/Association of the highest number of deaths by COVID-19 and the Human Development Index (HDI) of cities in Santa Catarina. *ID on line Revista de Psicologia, 15*(54), 427–434.

Araujo, I. G., et al. (2020). COVID-19 no Estado do Ceará: Impacto do índice de desenvolvimento humano (IDH) na casuística. *Brazilian Journal of Health and Pharmacy, 2*(3), 21–30.

Balassiano, M., Seabra, A. A., & Lemos, A. H. (2005). Escolaridade, salários e empregabilidade: tem razão a teoria do capital humano? *Revista de Administração Contemporânea, 9*(4), 31–52.

Barbosa, A. L. N. H., Costa, J. S., & Hecksher, M. (2020). Mercado de trabalho e pandemia da covid-19: Ampliação de desigualdades já existentes? Nota técnica. *Mercado de Trabalho*(69).

Batista, A., et al. (2020). Análise socioeconômica da taxa de letalidade da COVID-19 no Brasil. *Núcleo de Operações e Inteligência em Saúde (NOIS).*

Besarria, V. S. C., et al. (2016). Análise da relação entre escolaridade e a saúde da população brasileira. *Revista Espacios, 37*(2). https://www.revistaespacios.com/a16v37n02/16370210.html

Borges, G. M., & Crespo, C. D. (2020). Aspectos demográficos e socioeconômicos dos adultos brasileiros e a COVID-19: uma análise dos grupos de risco a partir da Pesquisa Nacional de Saúde, 2013. *Cadernos de Saúde Pública, 36*.

Bousada, G. M., & Pereira, E. L. (2017). Produção de vacinas virais parte I: engenharia de bioprocessos. *Revista da Universidade Vale do Rio Verde, 15*(1), 309–332.

Cano, W., & Guimarães Neto, L. (1986). A questão regional no Brasil: Traços gerais de sua evolução histórica. In *Pensamento Iberoamericano. Revista de Economia Política.*

Campos, L. C., & Miranda, G. J. (2021). Relevância do capital humano para enfrentamento de crises: Um olhar a partir da Teoria do Capital Humano nas organizações em tempos de pandemia. In *21st USP International Conference in Accounting*. São Paulo.

Chaudhry, R., et al. (2020). A country level analysis measuring the impact of government actions, country preparedness and socioeconomic factors on COVID-19 mortality and related health outcomes. *EClinicalMedicine, 25*, 100464.

Cruz, A. C., Teixeira, E. C., & Braga, M. J. (2010). Os efeitos dos gastos públicos em infraestrutura e em capital humano no crescimento econômico e na redução da pobreza no Brasil. *Economia, Selecta, 11*(4), 163–185.

De Lima, A. L. B., & Lima, K. C. (2021). Fatores prognósticos de mortalidade por Covid-19 no Rio Grande do Norte, Brasil. *Revista Brasileira de Ciências da Saúde, 25*(2), 225–234.

Demenech, L. M., et al. (2020). Desigualdade econômica e risco de infecção e morte por COVID-19 no Brasil. *Revista Brasileira de Epidemiologia, 23*.

De Negri, F., et al. (2021). Aspectos socioeconômicos da Covid-19: Evidências dos trabalhadores formais do estado do Rio de Janeiro. *Texto para Discussão, IPEA*.

Ferreira, D., Silva, L., & Figueiredo Filho, D. B. (2021). Saneamento importa? Uma análise da relação entre condições sanitárias e COVID-19 nas capitais brasileiras. *Engenharia Sanitária e Ambiental, 26*, 1079–1084.

Fink, G., et al. (2021). Inactivated trivalent influenza vaccination is associated with lower mortality among patients with COVID-19 in Brazil. *BMJ Evidence-Based Medicine, 26*(4), 192–193.

Galvão, M. H. R., & Roncalli, A. G. (2021). Fatores associados a maior risco de ocorrência de óbito por COVID-19: análise de sobrevivência com base em casos confirmados. *Revista Brasileira de Epidemiologia, 23*.

Griffith, D. M., et al. (2020). Men and COVID-19: A biopsychosocial approach to understanding sex differences in mortality and recommendations for practice and policy interventions. *Preventing Chronic Disease, 17*, E63.

IBGE. (2021). *Pesquisa Nacional por Amostra de Domicílios Contínua*. https://www.ibge.gov.br

IBGE. (2023). *Indicadores sociais municipais: Uma análise dos resultados do universo do Censo Demográfico*. https://www.ibge.gov.br/home/estatistica/populacao/censo2010/indicadores_sociais_municipais/default_indicadores_sociais_municipais.shtm

IBGE. (2019). *Síntese de Indicadores Sociais*. https://biblioteca.ibge.gov.br/visualizacao/livros/liv101678.pdf

IPEA. (2015). *Índice de Vulnerabilidade Social*. http://ivs.ipea.gov.br/index.php/pt/

Kerr, L., et al. (2020). COVID-19 no Nordeste brasileiro: sucessos e limitações nas respostas dos governos dos estados. *Ciência & Saúde Coletiva, 25*, 4099–4120.

Laster Pirtle, W. N. (2020). Racial capitalism: A fundamental cause of novel coronavirus (COVID-19) pandemic inequities in the United States. *Health Education & Behavior, 47*(4), 504–508.

Lopes, T. H. C. R. (2015). *Democracia, confiança e crescimento econômico* (Tese de doutorado). Universidade Federal de Minas Gerais.

Lotter, H., & Altfeld, M. (2019). Sex differences in immunity. In *Seminars in Immunopathology* (pp. 133–135). Springer Berlin Heidelberg.

Ministério da Saúde (BR). (2021). *Boletim epidemiológico especial: Doença pelo coronavírus Covid-19, Semana epidemiológica, 51 (19/12 a 25/12)*. https://www.gov.br/saude/pt-br

Moreira, J. A., & Oliveira, C. S. (2020). Tempo e gênero na crise do covid-19. In C. Sales, E. Araújo, & R. Costa (Orgs.), *Tempo e Sociedade em Suspenso* (cap. 11). Lisboa: CIES_Iscte.

Nunes, L. N. (2007). *Métodos de imputação de dados aplicados na área da saúde* (Tese de doutorado). Universidade Federal do Rio Grande do Sul.

Oliveira, L. R., et al. (2021). Vacinação contra a covid-19 em Mato Grosso: primeiros resultados. *Instituto de Saúde Coletiva - UFMT. Nota Técnica 02/2021.*

Organização Pan-Americana da Saúde (OPAS). (2021). *Desfechos de saúde e COVID-19 nas Américas: diferenças de sexo. Janeiro de 2020 a janeiro de 2021.* https://iris.paho.org/handle/10665.2/53602

Peres, I. T., et al. (2021). Sociodemographic factors associated with COVID-19 in-hospital mortality in Brazil. *Public Health, 192*, 15–20.

Pinheiro, F. M. G., et al. (2020). Iniquidades regionais e sociais na mortalidade por covid-19 no Brasil. *Revista Brasileira de Gestão e Desenvolvimento Regional, 16*(4).

Pires, L. N., Carvalho, L., & Xavier, L. L. (2020). COVID-19 e desigualdade: a distribuição dos fatores de risco no Brasil. *Experiment Findings, 21*.

Programa das Nações Unidas para o Desenvolvimento (PNUD). (2019). *Relatório do Desenvolvimento Humano 2019 – Além do rendimento, além das médias, além do presente: Desigualdades no desenvolvimento humano no século XXI*. PNUD.

Porto, E. F., et al. (2021). Mortalidade por Covid-19 no Brasil: perfil sociodemográfico das primeiras semanas. *Research, Society and Development, 10*(1).

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Sage Publications.

Santos, P. R., & Teixeira, N. A. (2022). As sociologias da pandemia: contribuições sobre a Covid-19 e sociedade. *Sociologias, 24*(60), 18–30. https://doi.org/10.1590/18070337-126449

Rossman, H., et al. (2021). Patterns of COVID-19 pandemic dynamics following deployment of a broad national immunization program. *medRxiv*. https://doi.org/10.1101/2021.02.08.21251325

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. John Wiley & Sons.

Silva, V. V. A. (2021). A COVID-19 enquanto questão social: classe, escolaridade e cor da pandemia no Pará. *Holos, 1*, 1–14.

Soares, T. M., & Mendonça, M. C. M. (2003). Construção de um modelo de regressão hierárquico para os dados do SIMAVE-2000. *Pesquisa Operacional, 23*(3), 421–441.

Souza, F. S. H., et al. (2021). On the analysis of mortality risk factors for hospitalized COVID-19 patients: A data-driven study using the major Brazilian database. *PLOS ONE, 16*(3). https://doi.org/10.1371/journal.pone.0248580

Souza, L. P., et al. (2020). COVID-19 no Brasil: seguimos no mesmo mar, mas não nos mesmos barcos. *Comunicação em Ciências da Saúde, 31*(3), 41–48.

Van Buuren, S. (2007). Multiple imputation of discrete and continuous data by fully conditional specification. *Statistical Methods in Medical Research, 16*(3), 219–242.

Van Buuren, S., & Oudshoorn, K. (1999). *Flexible multivariate imputation by MICE*. Leiden: TNO.

Vasconcelos, I. M. P., & Muyllder, C. F. (2022). Os vazios institucionais, saneamento e déficit de água e os primeiros números Covid-19 no Brasil. *Revista de Administração da UFSM, 14*, 1221–1238.

World Bank. (2021). *Global economic prospects, June 2021*. https://www.worldbank.org/pt/publication/global-economic-prospects

Zhang, P. (2003). Multiple imputation: Theory and method. *International Statistical Review, 71*(3), 581–592.

[i] Mestre em Economia pela Universidade Federal de Sergipe (UFS). Técnico Administrativo do Departamento de Computação da Universidade Federal de Sergipe (UFS). (79) 3194-6678. Cidade Universitária Prof Aloísio de Campos, São Cristóvão/SE.

[ii] Doutor em Economia pela Universidade Federal de Minas Gerais. Professor Adjunto do Departamento de Economia da Universidade Federal de Sergipe (UFS) – São Cristovão – SE.

[iii] Doutora em Desenvolvimento Econômico pela Universidade Federal do Paraná (UFPR). Professora Associada do Departamento de Economia da Universidade Federal de Sergipe (UFS). Cidade Universitária Prof Aloísio de Campos , Jardim Rosa Elze, São Cristóvão/SE.