2023

Association of Income, Education, Employment and COVID-19 Mortality in Brazil

Marcelina Machado
Walden University

Walden University

College of Health Sciences and Public Policy

This is to certify that the doctoral dissertation by

Marcelina Machado

has been found to be complete and satisfactory in all respects,
and that any and all revisions required by
the review committee have been made.

Review Committee
Dr. Peter Anderson, Committee Chairperson, Public Health Faculty
Dr. Patrick Dunn, Committee Member, Public Health Faculty
Dr. Tina Cunningham, University Reviewer, Public Health Faculty

Chief Academic Officer and Provost
Sue Subocz, Ph.D.

Walden University
2023

Abstract

Association of Income, Education, Employment and COVID-19 Mortality in Brazil

by

Marcelina Machado

MS, Southern New Hampshire University, 2017

BS, Federal University of Lavras, 2009

AS, Bunker Hill Community College, 2012

Dissertation Submitted in Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

Public Health

Walden University

February 2023

Abstract

The increased mortality rates from COVID-19 in Brazil have adversely affected vulnerable populations and placed a severe burden on its public health system as compared to other nations. Focusing on socioeconomic status as a risk factor in this retrospective cohort study, I investigated the association between Brazilian COVID-19 mortality rate, income, school enrollment rate, and employment rate after controlling for the demographics and immunization rate. This is an important issue as an increase in incidence and mortality rates affects how governmental and public health leaders can effectively control and prevent disease spread. Using the fundamental causes of disease theory as the theoretical framework, I evaluated the underlying influence of socioeconomic factors on COVID-19 mortality at different geographic levels (region, state, and capital). Therefore, the foci of these three manuscripts is the need to assess this influence to encourage governmental response through better planning, strategies, and actions toward socioeconomic equality in Brazil, which could promote a reduction in the COVID-19 mortality rate. Most findings do not agree with the theory tested, which can be explained by a series of other factors, such as population density and heterogeneity. Thus, there are still inconsistencies regarding how minority groups are at greater risk of disease contamination and death. The social change implications of this study include a potential increase in opportunities for socioeconomic equality. A decrease in the COVID-19 mortality rate could also allow Brazilians to improve individual lives with healthier families and communities.

Association of Income, Education, Employment and COVID-19 Mortality in Brazil

by

Marcelina Machado

MS, Southern New Hampshire University, 2017

BS, Federal University of Lavras, 2009

AS Bunker Hill Community College, 2012

Dissertation Submitted in Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

Public Health

Walden University

February 2023

Dedication

To my Lord and Savior, Jesus Christ, who never abandoned me.

Acknowledgments

I want to thank my dearest parents, José Ronaldo Carvalho and Rita Elaine Carvalho, who taught me that there is nothing better than to dream. I want to thank my little sister and brother-in-law, Mariana Protes and Clair Protes, who taught me the importance of social responsibility and justice. I want to thank my loving husband, Arnaldo Machado, for being by my side during these fourteen years, always giving love, support, and care. I want to thank my beloved grandpa, Nicodemos Cunha, who taught me that wisdom does not come from books. I want to remember the memories of my treasured grandparents 'Nicinha' Cunha, João da Cruz, and 'Dica' Leite, who will always be in my heart. My sincere thanks to my committee chair, Dr. Anderson, and the second committee member, Dr. Dunn, for their mentorship and guidance. Dr. Anderson was the most patient and kind mentor I could have ever asked for, and it was a pleasure to be his mentee during this past year.

Reconhecimentos

Quero agradecer aos meus queridos pais, José Ronaldo Carvalho e Rita Elaine Carvalho, que me ensinaram que não há nada melhor do que sonhar. Quero agradecer a minha irmã caçula e cunhado, Mariana Protes e Clair Protes, que me ensinaram a importância da responsabilidade e justiça social. Quero agradecer ao meu amado marido, Arnaldo Machado, por estar ao meu lado durante estes quatorze anos dando amor, apoio e carinho. Quero agradecer ao meu querido avô, Nicodemos Cunha, que me ensinou que a sabedoria não vem dos livros. Quero lembrar das memórias dos meus queridos avós ‘Nicinha’ Cunha, João da Cruz e ‘Dica’ Leite, que estarão sempre no meu coração. Meus sinceros agradecimentos ao presidente do meu comitê, Dr. Anderson, e ao Segundo membro do meu comitê, Dr. Dunn por sua orientação. Dr. Anderson foi o mentor mais paciente e doce que eu poderia ter, e foi um prazer ser sua pupila durante todo o ano que se passou.

Table of Contents

Part 1: Overview

Introduction

The COVID-19 pandemic has become a prominent event in human history, crossing borders and altering lives by bringing illness and death and forcing people to adapt to a new way of living. Since the World Health Organization (WHO) declared a Public Health Emergency of International Concern on 30th January 2020, and a Global Pandemic on 11th March 2020 (WHO, 2020-a), changes such as travel restrictions, mask mandates, lockdowns, basic hygiene practices, social distancing, and remote work (Centers for Disease Control and Prevention [CDC], n.d.-a; WHO, n.d.-a) have given our daily routines a unique shape and pushed the world towards a virtual existence. However, even with all the control and preventive guidelines provided by the great public health agencies, the impact of COVID-19 on people's livelihood and health was unmeasurable as we witnessed countries facing a significant loss of human life.

Considering the diversity of each population and socioeconomic differences between and within countries, one could argue that nations with significant socioeconomic and healthcare disparities present high vulnerability to diseases. This vulnerability could result from a lack of financial and infrastructural resources, as well as the absence of implementable policies for disease monitoring and preventive measures, leading to an increase in disease spread. Although it is still unclear how socioeconomic status (SES) affects disease occurrence and growth, studies have shown that political, demographic, and social structural factors negatively impact public health, thus facilitating the spread of the virus and increasing the likelihood of death (Goutte et al., 2020; Mena et al., 2021). Based on these uncertainties, most questions can only be answered after a better understanding of countries' politics and public health systems.

The increased mortality rates from COVID-19 in Brazil can adversely affect vulnerable populations and become a severe burden to its public health system as compared to other nations. There is strong evidence that the high mortality rate in Brazil is associated with the population's socioeconomic conditions, particularly for minority groups (Miranda et al., 2020; Rocha et al., 2021). Consequently, an increase in disease mortality rates presents additional challenges to the public health population, which is socioeconomically disadvantaged, as well as the country's growth. From March 2020 to December 2021, 634,695 deaths were registered in Brazil (Secretaria de Vigilância em Saúde [Health Surveillance Secretariat], 2022). Baqui et al. (2021) and Rocha et al. (2021) explained that COVID-19 incidence and mortality increases vary depending on health-system preparedness, socioeconomic characteristics of the population, and people's adherence to best-practice protocols. These studies provide evidence that increased mortality rates from COVID-19 across Brazil can be associated with a lack of strict control measures, including policies and programs to reduce infections and individual careless behavior.

Other studies show that, despite neglectful political behavior in Brazil that influences denialism, vaccine hesitancy is low compared to other countries (Moore et al., 2021; Stojanovic et al., 2021). Historically, Brazil has a strong culture of mass vaccination as initiated in the '70s (Domingues et al., 2020), which allowed the population to become more receptive to COVID-19 vaccines. However, socioeconomic disparities also play a role in vaccine acceptance (Stojanovic et al., 2021). Therefore, in this retrospective cohort study, I focused on socioeconomic factors as a risk factor to investigate the association between COVID-19 mortality rate in Brazil, income, school enrollment rate, and employment rate, after controlling for race, age, sex, and immunization rate. Investigating the association and potential impact of socioeconomic factors on COVID-19 mortality is important as an increase in incidence and mortality rates affects how governmental and public health leaders can effectively control and prevent disease spread.

Lastly, the findings of this study could empower public health and political leaders to identify geographic areas in the country that need special attention regarding the development of policies and implementation of public health programs to encourage populations' best behavior and practices concerning preventive actions and vaccination willingness. The focus of this study is on the need to assess the influence of socioeconomic factors on disease mortality to encourage governmental response through better planning, strategies, and actions toward socioeconomic equality in Brazil that could promote a reduction in the COVID-19 mortality rate. The social change implications of this study may include a potential increase in opportunities for socioeconomic equality, considering governmental actions towards better income distribution, education, and employment are taken. Reducing the COVID-19 mortality rate will allow Brazilians, particularly socioeconomically disadvantaged individuals and families, a gradual transition toward normalcy.

Background

Since the beginning of the pandemic, studies on the association between socioeconomic disparities and COVID-19 incidence and mortality have been published (Bermudi et al., 2021; Martins-Filho et al., 2020; Miranda et al., 2020). The analysis of such disparities incorporated various factors, including income distribution, education, employment, and the population's demographic heterogeneity (Figueiredo et al., 2020; Oliveira et al., 2020). Although the influence of these factors on disease outcomes could be evaluated individually, it is more advantageous to assess these factors in parallel to know more about the potential impacts these factors may cause on infectious disease incidence and mortality from a public health standpoint. Most studies in the literature I reviewed agree with the impact of SES on COVID-19 mortality. Many go beyond the three critical factors of SES (income, education, and employment), assessing indicators such as individuals' comorbidities (Baqui et al., 2021).

However, considering that SES varies across and within countries, its impact on disease outcomes cannot be generalized. A gap found during the literature review were the minimal studies focused on the association between SES and COVID mortality in Latin American countries. For this reason, this study targeted the Brazilian population, one of the world's most ethnically and culturally diverse countries, which also presents significant wealth inequality. Another gap is that only a few studies consider vaccination or the population's immunity status when testing the association between SES and COVID-19 mortality. Therefore, besides including population demographics, race, age, and gender, I considered this an essential factor. The following is a synthesis of the existing literature that explores the association between SES and COVID-19 outcomes.

To investigate the association between hospitalized individuals' demographics, comorbidities, socioeconomic, geographic, and structural hospital-level factors, and mortality risk to COVID-19, Baqui et al. (2021) conducted a retrospective cohort study across 1,801 cities and 3,991 hospitals in Brazil. Using the Brazilian SIVEP-Gripe database and the XCOVID-BR model created by the authors, Baqui et al. (2021) found that SES, race, geographical, and structural factors are more critical than individual patients' comorbidities in determining COVID-19 mortality in Brazil. Past studies found evidence to support the association between strain on hospital capacity and increase in disease mortality during the pre-pandemic and COVID-19 period (Bravata et al., 2021; French et al., 2021). Baqui et al. (2021) stressed the importance of funding public hospitals and better managing the healthcare network, which can be achieved by improving Brazil's public health and care policies. Additionally, the authors recommend that the Brazilian Ministry of Health implement programs to increase the population's vaccination status as the primary preventive measure to cover sections of the country that are less economically developed. Although this study accounted for many of the socioeconomic factors I am interested in, such as education and income as potential contributors to COVID-19 incidence and death, the authors did not consider population immunization status as a control variable.

Based on the principles of the fundamental causes theory, Clouston et al. (2021) tested the hypothesis that the SES association with incidence and disease mortality shifts from a favorable (higher SES higher rates) to a negative (higher SES lower rates) stage over time. The authors proposed that an epidemic occurs as a two-stage process. First, higher SES communities that are not ready catch the disease, which leads to initial cases. Second, when public health preventive measures begin to be implemented, control risks are unequally distributed, and low SES communities face an increased disease burden.

In this study, the authors investigated how SES variation can predict the dynamics of COVID-19 incidence and mortality. Using daily COVID-19 incidence and mortality rates data on 3141 counties within the U.S., the authors conducted a retrospective cohort study using a multivariable analysis that included survival analyses and Poisson regression (Clouston et al., 2021). They also used the principles of the fundamental causes theory to examine the role of SES in predicting the distribution and death from COVID-19 in the U.S. The authors found that growth in incidence and case-fatality rates was lower in high SES areas concluding that vulnerability is an essential component of risk exposure and mortality. This study showed an urgent need to strengthen social protection systems to support vulnerable populations during the pandemic. It also used county-level data, which policymakers can apply to determine how the disease spreads, where hot spots emerge, and what sorts of interventions should be implemented.

However, Clouston et al.'s study (2021) disregarded individual-level factors. The authors suggested that in the absence of adequate social protection, populations in areas of low SES need to be educated about the potential for infections. Policies and legislation should be adopted to protect public health in vulnerable areas to ensure better assistance for communities with high socioeconomic inequalities. This study is essential for information purposes as I tested the same theory but in a different population. Besides the test of the theory, Clouston et al. (2021) also provided historical background on the formation of the fundamental cause theory and its key concepts, referring to the original work of Link and Phelan, who developed the theory in 1995.

Khanijahani et al. (2021) reviewed various studies that provided evidence concerning the association between racial/ethnic, socioeconomic disparities, and COVID-19. The authors reviewed 52 studies, the majority from the United States and published between late December 2019 through March 1, 2021. The main variables investigated in the studies included in this review were the population’s ethnicity and SES, COVID-19 mortality and infection, hospitalization admission, and access to testing. Past studies have shown evidence that lower SES is associated with a higher risk of infectious diseases occurrence and severity, and that demographic characteristics including ethnicity and SES inequalities are directly or indirectly associated with the main indicators of health, mortality, and life expectancy (Khalatbari-Soltani et al., 2020; Ogedegbe et al., 2020; Ren et al., 1999). Khanijahani et al. (2021) found that most studies showed that racial or ethnic minority groups had increased risks of COVID-19 infection, hospitalization, and mortality.

It was also found that some studies investigated the impact of education, income, poverty, and housing conditions on the COVID-19 incidence and mortality. The authors concluded that there is an inconsistency among studies but most still showed that minority groups are at greater risk of disease infection, hospitalization, and death (Khanijahani et al., 2021). The authors suggested further research to address the association between intensive care admission, lack of insurance and employment, and COVID-19 outcomes (Khanijahani et al., 2021). This study is important for information purposes, serving as a source and link to other studies investigating racial/ethnic and socioeconomic disparities in COVID-19. In my research, the intent was to fill one of the gaps mentioned by these authors, which is to explore the association between employment and COVID-19 outcomes (Khanijahani et al., 2021); additionally, I focused on a different population, a Latin American country.

This literature review also includes Link and Phelan's (1995) study, a systematic review of 240 articles published between November 1992 and 1993 in the *American Journal of Epidemiology* that provided evidence concerning the association between social conditions and illness. Based on their research, Link and Phelan developed the theory of fundamental causes, which stemmed from Lieberson's concept of basic causes set in 1985. Link and Phelan's (1995) formulated a theory that postulates that a fundamental social cause of a disease has an essential feature that involves access to resources that could minimize the chances of the disease occurring; these resources include money, knowledge, power, prestige, and social connections. The authors found that 13.3% of the articles reviewed focused on risk factors that could be understood as social in nature. Therefore, Link and Phelan's (1995) reminded the reader that in their theory, they consider individual-based risk factors and that demographic factors, including race and gender, are often analyzed as potential fundamental causes as they are closely tied to these resources. Link and Phelan (1995) also suggested future studies considering population-based factors, as policies are made for the population. This study is critical because it gives the basis of the theory of fundamental causes, which was used in my research.

Using Link and Phelan's theory (1995), Mackenbach et al. (2017) extended their investigation on the association between socioeconomic inequalities and mortality in Europe (Mackenbach et al. 1997; Mackenbach et al., 2003; Mackenbach et al. 2008; Mackenbach et al. 2015) by conducting an extensive retrospective cohort study in populations of 17 European countries using mortality data collected from 1980 - 2010 census data. In this study, the authors hypothesized that (1) when mortality in the whole population reduces, it generally reduces faster among the higher educated groups; (2) when mortality rate rises, the higher educated groups will be better prepared to protect themselves; therefore, mortality increase will be lower among them (Mackenbach et al., 2017). Further, (3) variation in mortality reduction is more preventable than for non-preventable causes of death; and (4) variation in mortality reduction between groups at different educational levels is higher in countries where inequality in resources is significant (Mackenbach et al., 2017). From the collected data, Mackenbach et al. (2017) classified 22 causes of death into four groups according to their "preventability." The authors found that faster mortality declined among the highly educated for most causes of death and in most populations; however, they found no support that highly educated individuals or individuals with higher income can protect themselves during increased mortality rates. Mackenbach et al. (2017) concluded that the difference in mortality decline between groups of different education levels was not more considerable when income inequalities were more significant. Thus, Mackenbach et al. (2017) suggested that mortality decline depends not on the individual but on collective action. They also recommended further testing their hypotheses with other indicators of SES. This study is critical as I tested the same theory and analyze similar variables in a Latin American population. Besides the test of the theory, Mackenbach et al. (2017) also provided historical background on the formation of the fundamental causes theory and its key concepts.

To investigate the association between social inequality and the COVID-19 fatality rate in a particular area in Northeast Brazil, Martins-Filho et al. (2020) conducted an ecological study targeting the population of Aracaju City. Martins-Filho et al. (2020) looked at confirmed cases and deaths due to COVID-19, using publicly available data from the Brazilian Institute of Geography and Statistics (IBGE) and surveillance information systems to determine COVID-19 rates of incidence and case fatality. The authors found that the poorest neighborhoods had lower reported cases of COVID-19 but higher fatality rates. The authors concluded that differences in COVID-19 incidence and mortality rates were due to differences in literacy, income, and housing between neighborhoods (Martins-Filho et al., 2020). As the study focused on a single town, I would suggest replicating the study on a larger population. Although these authors and I shared similar interests in the research topic, the authors only considered a small fraction of the population. Conversely, I investigated the association between socioeconomic factors and COVID-19 mortality nationwide in my research. Besides, Martins-Filho et al. (2020) did not consider the population's vaccination status as a covariate, and I did. Historically, disadvantaged groups are more likely to experience a higher incidence of infection and mortality in a pandemic scenario (Bambra et al., 2020). Previous studies also have shown that mortality rates are higher in developing countries among unemployed individuals and groups living in urban areas (Chowell et al., 2008; Grantz et al., 2016; Murray et al., 2006).

In another study, Martins-Filho et al. (2021) estimated the incidence and mortality rates of COVID-19 in Brazilian children and analyzed its relationship with socioeconomic inequalities across the nation. The authors used the social vulnerability index and Gini coefficient to assess socioeconomic disparities. The target population included 3,998,055 individuals with COVID-19; in the sample, 335,279 individuals were aged 0-19 years. Martins-Filho et al. (2021) conducted a secondary analysis of data obtained from the microdata catalog and official bulletins of the 27 Brazilian health department websites. Data were analyzed using Spearman's rank correlation. Martins-Filho et al. (2021) found that 8.4% of the individuals with COVID-19 were 0-19 years old, and that 0.7% of the deaths registered in the country were related to children with COVID-19. The authors found significant differences between the incidence and mortality rates among Brazilian regions, and a significant correlation between mortality rates and social and economic inequalities. This is one of the few studies investigating the association between SES and child mortality due to COVID-19 in Brazil. The authors had no suggestions for future research; however, it would be interesting to replicate the study in different populations for comparison. This study is essential for information purposes, as the authors used publicly available databases to estimate the population's Gini coefficient to evaluate income distribution; this would be helpful for my research, considering my intent to explore socioeconomic factors associated with COVID-19 mortality rates.

Mateo-Urdiales et al. (2021) evaluated the association between deprivation, COVID19 incidence, risk of hospitalization, and mortality from COVID-19 while adjusting for age, sex, and region of residence in Italy during three periods: the pre-lockdown, lockdown, and post-lockdown. Data were collected from the Italian integrated epidemiological surveillance system of COVID-19 and the Italian National Statistics Institute. The authors conducted a retrospective cohort study based on secondary data using a multilevel negative binomial regression model, finding a higher incidence of COVID-19 cases in the most deprived Italian municipalities during lockdown and post-lockdown compared with the least deprived. Mateo-Urdiales et al. (2021) found no significant differences in case hospitalization and fatality regarding deprivation; however, they only considered population-based factors, disregarding individual ones. Additionally, Mateo-Urdiales et al. (2021) evaluated the variables' level of education, employment, and living conditions, concluding that the impact of deprivation on COVID-19 outcomes was heterogeneous across the target population. Mateo-Urdiales et al. (2021) suggested future studies considering individual-based factors. This study is important as it explored three distinct periods of the pandemic, allowing for a better evaluation of public health implementations by public health leaders and politicians. In my research, I extended Mateo-Urdiales et al.'s (2021) study by including immunization rate, another population-based factor, to enunciate the importance of vaccination and its potential effect on the association between SES and COVID-19 mortality.

To evaluate the association between income inequality in Brazil and the risk of infection and death by COVID-19, Miranda et al. (2020) conducted an ecological study targeting the population of each state to look at the incidence and mortality rates of all states for comparison. Data were extracted from the Brazilian Institute for Geography and Statistics (IBGE), and Miranda et al. (2020) used the Gini coefficient as the primary independent variable to evaluate income inequality. The authors found high incidence and mortality rates of COVID-19 in areas with greater economic inequality, concluding that income inequality could play an important role in the impact of COVID-19 on the Brazilian population.

Although Brazil established a unified health system known as Sistema Único de Saúde in 1988 (Ministry of Health, n.d.; Rocha et al., 2021), past studies have indicated that Brazil's advances in healthcare are unequal within the country, being worse in groups socioeconomically disadvantaged (Barreto et al., 2011; Paim et al., 2011). Therefore, I would recommend future investigations on policies seeking improvements in the delivery of care aiming at unprivileged groups. Additional suggestions from Miranda et al. (2020) for further study were to identify the causal paths of results obtained. This study is essential as the authors provide information on Brazilian governmental databases that I can use. Whereas Miranda et al.'s (2020) focus was to examine, through space-time analyses, the association between Brazil's economic disparities using only demographic density and income distribution and risk of COVID-19 infection and death, my research accounted for other variables, including education and employment, while considering for population's demographics and immunization rate.

On a similar note, Rocha et al. (2021) examined the association between socioeconomic inequalities and vulnerabilities associated with health system preparedness and COVID-19 mortality rates across Brazilian regions in a secondary retrospective cohort study. Rocha et al. (2021) specifically looked at variables including COVID-19 mortality, the burden of chronic disease, housing, employment status, availability of hospital beds and staff, healthcare coverage, and social assistance, while adjusting for the population's age. Additionally, Rocha et al. (2021) developed a socioeconomic vulnerability index based on data from housing and employment, as well as subcomponents, including income and education. Rocha et al. (2021) found that COVID-19 death rates and socioeconomic vulnerability were higher in North and Northeast regions while the South presented the lowest. The study is considerably extensive, with a vast number of independent variables in their analysis. Rocha et al. (2021) mentioned that limitations included measurement errors regarding confirmed deaths and administrative data, which could contain errors causing biased results. Rocha et al. (2021) concluded that the spread of the disease was affected mainly by patterns of socioeconomic vulnerability rather than population age and prevalence of health risk factors. They also found that death rates increased in areas with high socioeconomic vulnerability. Rocha et al. (2021) suggested that efforts to strengthen local-level public health responses should be made instead of only counting on central governmental actions. This study is important for information purposes as I also plan to explore how population socioeconomic characteristics relate to disease mortality. Rocha et al.'s (2021) study include a comprehensive list of variables, some of which are also present in my research. However, I expanded our knowledge on the topic by accounting for other population demographic factors as well as immunization rate.

There are significant differences in how the governments responded to the pandemic. According to Shadmi et al. (2020), in Brazil, the president encouraged the population not to follow public health guidelines, while in countries like China and Israel, the pandemic was successfully under control due to the population's willingness to follow preventive practices and government preparedness healthcare responses. Other countries adopted relief measures, such as the United States, India, and Guatemala, to support small businesses and the most vulnerable population. In their study, Shadmi et al. (2020) reviewed 13 case studies from various countries to investigate the association between health equity and COVID-19 worldwide. Their intent was not to compare studies but rather to describe and discuss how health inequities have worsened the pandemic in some areas of the globe. This study brings an excellent overview of the event globally; however, it could be extended to a systematic review or meta-analysis. Although the authors do not compare the situation between countries, their study allows us to reflect on how each government prioritizes public health decisions and takes actions differently. This study is a warning of how governmental leaders must take action to fight socioeconomic, ethnic, and health disparities; therefore, this study is essential for information purposes as it provides an overview of the Brazilian government's denialist approach, encouraging the population to neither follow preventive actions nor seek immunization. This information is essential to understanding Brazil's vaccination rate explored herein. Additionally, the authors report information on racial and socioeconomic inequality, and the provided references can serve as background information for my research.

Based on this literature review, it can be observed that the authors agreed on the impact of SES on COVID-19 incidence and mortality, which changes based on populations' demographic and geographic heterogeneity. There are, however, very few studies on the topic with their focus on Latin-American countries, and most have not accounted for the impact of populations' immunization rate on disease mortality. Therefore, I intended to extend the research on the topic but with a focus on the larger and most diverse Latin-American population, exploring the potential association of SES on COVID-19 mortality across the country.

Independent Variables

Income

As employment and education, the income variable relates to this study as it was a determinant used to assess socioeconomic inequality and the vulnerability of a population. Latin American countries, including Peru, Colombia, and Brazil, present a pronounced disparity in income distribution which can be associated with a lower degree of public health policy compliance, leading to an increase in the death rate of COVID-19 (Martinez-Valle, 2021). Studies during the ongoing pandemic have shown, however, that this increase in mortality was not the result of income inequality alone but of a sum of other socioeconomic factors, including population density, education, employment, literacy, and housing, which have minimized the effectiveness of public health response (Bermudi et al., 2021; Mackenbach et al., 2017; Miranda et al., 2020; Martins-Filho et al., 2020; Rocha et al., 2021). Therefore, investigating the association between income and COVID-19 mortality while considering a population's demographic and socioeconomic heterogeneity is fundamental to creating a broader perspective of the influence of each of these factors on the mortality rate of COVID-19. Strategies to reduce income inequality include expanding access to education that would increase employment opportunities to create a source of regular income (Miranda et al., 2020).

Education

The education variable relates to this study as it was a factor used to assess a population's socioeconomic inequality and vulnerability. Studies have shown that areas with a higher number of school enrollments have a greater socioeconomic prestige as it can be associated with easier access to education by local individuals (Ribeiro et al., 2017). In contrast, vulnerable areas have a lower offer of school enrollment, often insufficient to meet the demand, meaning that vulnerable individuals have less access to education (Bartholo et al., 2020; Ribeiro et al., 2017). Population education is an essential factor in the distribution of infectious diseases (Rattay et al., 2021). Studies have also shown that education level differs depending on other demographics. For instance, unlike most countries, in Brazil, women are more likely to attain tertiary education than men; however, they are less likely to find employment (OECD, 2021). Khanijahani et al. (2021), who also investigated the impact of education on COVID-19 incidence and mortality, found that minority groups are at greater risk of disease infection, hospitalization, and death. Clouston et al. (2021) suggested that in the absence of adequate social protection, people in areas of low SES need to be educated about the potential for infections. The rationale for the study of the association between education and disease mortality is that lack of knowledge or access to information might indicate populations’ vulnerability and low-risk perception, increasing disease incidence and mortality (Barreto et al., 2021; Rattay et al., 2021).

Employment

The employment variable relates to this study as it was another factor used to assess socioeconomic inequality and the vulnerability of a population. Many authors have explored this variable as unemployment leads to loss of income and can lead to social and psychological suffering (Achdut et al., 2020; De Souza et al., 2020). According to the Brazilian Institute of Geography and Statistics (IBGE, 2022), during the second year of the pandemic, the average unemployment rate in Brazil was 13.2%, the second-highest rate since 2012 and three times higher compared to the unemployment rate in the United States (U.S. Bureau of Labor Statistics, n.d.). The rationale for studying the association between employment and disease mortality is that the Brazilian government has implemented many programs to fight infectious diseases; still these programs do not include actions to reduce socioeconomic inequality.

Fundamental Causes Theory

The theory that grounds this study is the fundamental causes theory (Link et al., 1995). Through this theory, the authors attempted to explain the relationships between SES and health, as well as why social inequalities in health continue even with medical innovation and disease elimination (Chang et al., 2010; Riley, 2020). This theory includes two concepts showing the importance of social factors in disease causation. The first concept is contextualizing individually based risk factors to evaluate why people become exposed to the disease. The second concept is the fundamental cause or determination of social conditions, such as SES, under which risk factors are related to the disease. This theory stems from Lieberson's concept of basic causes developed in 1985, who proposed that changes in basic causes (factors responsible for generating a particular outcome) create a change in the outcome (Link et al., 1995).

The connections between the framework presented and the nature of my study include Link and Phelan's theoretical work to explain the association between population SES and COVID-19 mortality. According to this theory, SES works as a fundamental cause of the disease because (1) these factors allow individuals access to basic resources, including wealth, knowledge, skills, power, and social connections to avoid disease, and (2) SES affects multiple disease outcomes and risk factors through time (Chang et al., 2010; Link et al., 1995). In this study, I focused on using the concept of fundamental cause or determination of socioeconomic conditions related to the COVID-19 mortality rates and the impact this had on the public's ability to access essential resources that could produce health benefits to prevent disease.

Overview of the Manuscripts

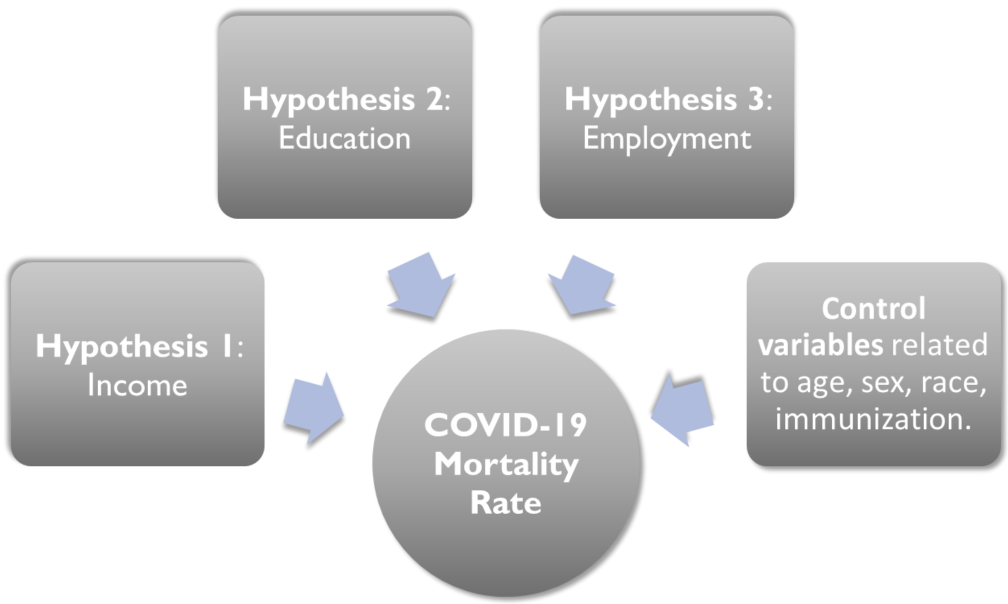
After an exhaustive search of the literature, little information was found regarding the association between SES and COVID-19 mortality at different geographic levels in Brazil, while controlling for population demographic characteristics and immunization. To address the research questions in this quantitative study, the specific research design included a retrospective cohort design (Aschengrau et al., 2020) with COVID-19 mortality examined between the period of March 2020 to December 2021. This quantitative analysis should help determine if SES and other demographic characteristics correlate to COVID-19 mortality in Brazil. The three manuscripts, part of this study, contain analyses of the association between (1) average income and COVID-19 mortality rate; (2) education, measured by school enrollment, and COVID-19 mortality rate; and (3) employment rate and COVID-19 mortality rate. All associations were examined after controlling for variables, including race, age, sex, and immunization in Brazil.

Although each of these factors (income, education, and employment) and their association with disease mortality were independently explored in each manuscript, it is more advantageous from a public health, political, and even administrative standpoint to study these factors in parallel so that more would be known about all the potential impacts these factors may cause on infectious disease spread nationwide. These indicators are the key factors in measuring SES (APA, 2022).

In addition to the socioeconomic factors investigated in this study, the population's demographic characteristics, including sex, age group, race, and immunization, were also considered to address the overall gap. Identifying and evaluating each of these social determinants of health and their influence, separately or concurrently, on the dynamics of COVID-19 in Brazil is of essential importance for determining the appropriate actions to deal with the pandemic and its repercussions (De Souza et al., 2020). Figure 1 represents the possible parallel association between the three main independent variables defining the populations' key socioeconomic factors (income, education, and employment) and the dependent variable COVID-19 mortality, while controlling for the variables representing the population's demographic characteristics (age, sex, and race), and immunization. Each hypothesis was tested in a separate manuscript to evaluate the effect of each SES factor on disease mortality during the pandemic; however, the conjunction of these manuscripts could allow us to better address the overall problem, COVID-19 mortality.

Figure 1

Scheme of Study Hypothesis, Variables, and Potential Associations for each Manuscript



Manuscript 1

The specific research problem that will be addressed through this manuscript is whether there is an association between the independent variable of average income and the dependent variable COVID-19 mortality rate, while controlling for variables related to race, age, sex, and immunization. The research question and hypotheses used to address the research problem are these:

Research Question for Manuscript 1 (RQM1): What is the association between average income and COVID-19 mortality rate at different geographic levels across Brazil after controlling for race, age, sex, and immunization rate?

Null Hypothesis (H0): There is no statistically significant association between average income and COVID-19 mortality rate at different geographic levels across Brazil after controlling for race, age, sex, and immunization rate.

Alternative Hypothesis (H1): There is a statistically significant association between average income and COVID-19 mortality rate at different geographic levels across Brazil after controlling for race, age, sex, and immunization rate.

To address the research questions in this manuscript, the specific research design includes a retrospective cohort design (Aschengrau et al., 2020) with mortality rates examined between the period of March 2020 to December 2021. This quantitative analysis should help to determine if average income and population demographic characteristics are correlated to COVID-19 mortality in Brazil.

Sources of Data

The data sources used to find public dataset repositories were the Instituto Brasileiro de Geografia e Estatistica (IBGE) [Brazilian Institute of Geography and Statistics], and the Ministério da Saúde [Ministry of Health]. Therefore, access to public dataset repositories were needed to gather specific data points, including the number of deaths by COVID-19 and the population's socioeconomic and demographic information. These data were obtained through access to governmental websites; however, because the

COVID-19 pandemic is an ongoing issue, not all data were available. Additionally, some datasets were not ready for analysis and present missing data.

The monthly average income across the Brazilian population was used, and data were retrieved from the IBGE [Brazilian Institute of Geography and Statistics] website. Data on COVID-19 mortality were retrieved from the Secretaria de Vigilância em Saúde [Health Surveillance Secretariat], created by the Ministério da Saúde [Ministry of Health]. Data on population demographics, including race, age, and sex, were retrieved from the Tabnet – DATASUS - Ministério da Saúde [Ministry of Health]. Lastly, data on immunization were retrieved from the Fundação Oswaldo Cruz (Fiocruz) [Oswaldo Cruz Foundation].

Since this is a population-based study, aggregated data from different geographic levels, including region, state, and capitals, were used for all dependent, independent, and control variables. For this study, convenience sampling was used (Warner, 2013). Based on the data retrieved for each geographic level during the first two years of the pandemic, the convenience sample size for regions is 10, for states is 54, and for capitals is 54. Additionally, information on individual-based data was used to show death distribution throughout the first years of the pandemic according to race, age group, sex, and geographic region, but were used as information only and not used to answer the study research question.

Manuscript 2

The specific research problem that will be addressed through this manuscript is whether there is an association between the independent variable of education measured by school enrollment rate and the dependent variable COVID-19 mortality rate, while controlling for variables related to race, age, sex, and immunization. The research question and hypotheses used to address the research problem are these:

Research Question for Manuscript 2 (RQM2): What is the association between education and COVID-19 mortality rate at different geographic levels across Brazil after controlling for race, age, sex, and immunization rate?

Null Hypothesis (H0): There is no statistically significant association between education and COVID-19 mortality rate at different geographic levels across Brazil after controlling for race, age, sex, and immunization rate.

Alternative Hypothesis (H1): There is a statistically significant association between education and COVID-19 mortality rate at different geographic levels across Brazil after controlling for race, age, sex, and immunization rate.

To address the research questions in this manuscript, the specific research design includes a retrospective cohort design (Aschengrau et al., 2020) with mortality rates examined between the period of March 2020 to December 2021. This quantitative analysis should help to determine if education and a population's demographic characteristics are correlated to COVID-19 mortality in Brazil.

Sources of Data

The data source used to find public dataset repositories were the Ministério da Saúde [Ministry of Health]. Therefore, access to public dataset repositories were needed to gather specific data points, including the number of deaths by COVID-19 and the population's socioeconomic and demographic information. These data were obtained through access to governmental websites; however, because the COVID-19 pandemic is an ongoing issue, not all data were available. Additionally, some datasets were not ready for analysis and presented missing data.

Education data (school enrollment) were retrieved from the Ministério da Educação [Ministry of Education] website. Data on COVID-19 mortality were retrieved from the Secretaria de Vigilância em Saúde [Health Surveillance Secretariat] created by the Ministério da Saúde [Ministry of Health]. Data on population demographics, including race, age, and sex, were retrieved from the Tabnet – DATASUS - Ministério da Saúde [Ministry of Health]. Lastly, data on immunization were retrieved from the Fundação Oswaldo Cruz (Fiocruz) [Oswaldo Cruz Foundation].

Since this is a population-based study, aggregated data from different geographic levels including region, state, and capitals were used for all dependent, independent, and control variables. For this study, convenience sampling was also used (Warner, 2013). Based on the data retrieved for each geographic level during the first two years of the pandemic, the convenience sample size for regions is 10, for states is 54, and for capitals is 54. Additionally, some information on individual-based data was still used to show death distribution throughout the first years of the pandemic according to race, age group, sex, and geographic region, but were used as information only and not to answer the study research question.

Manuscript 3

The specific research problem that will be addressed through this manuscript is whether there is an association between the independent variable employment rate and the dependent variable COVID-19 mortality rate while controlling for variables related to race, age, sex, and immunization. The research question and hypotheses used to address the research problem are as follows.

Research Question for Manuscript 3 (RQM3): What is the association between employment rate and COVID-19 mortality rate at different geographic levels across Brazil after controlling for race, age, sex, and immunization rate?

Null Hypothesis (H0): There is no statistically significant association between employment rate and COVID-19 mortality rate at different geographic levels across Brazil after controlling for race, age, sex, and immunization rate.

Alternative Hypothesis (H1): There is a statistically significant association between employment rate and COVID-19 mortality rate at different geographic levels across Brazil after controlling for race, age, sex, and immunization rate.

To address the research questions in this manuscript, the specific research design included a retrospective cohort design (Aschengrau et al., 2020) with mortality rates examined between the period of March 2020 to December 2021. This quantitative analysis should help to determine if employment and population's demographic characteristics are correlated to COVID-19 mortality in Brazil.

Sources of Data

The data sources used to find public dataset repositories were the Instituto Brasileiro de Geografia e Estatistica (IBGE) [Brazilian Institute of Geography and Statistics] and the Ministério da Saúde [Ministry of Health]. Access to public dataset repositories was needed to gather specific data points, including the number of deaths by

COVID-19 and the population's socioeconomic and demographic information. These data were obtained through access to governmental websites; however, because the COVID-19 pandemic is an ongoing issue, not all data were available. Additionally, some datasets were not ready for analysis and presented missing data.

Employment data were retrieved from the Instituto Brasileiro de Geografia e Estatistica (IBGE) [Brazilian Institute of Geography and Statistics]. Data on COVID-19 mortality were retrieved from the Secretaria de Vigilância em Saúde [Health Surveillance Secretariat] created by the Ministério da Saúde [Ministry of Health]. Data on population demographics, including race, age, and sex, were retrieved from the Tabnet – DATASUS - Ministério da Saúde [Ministry of Health]. Lastly, data on immunization were retrieved from the Fundação Oswaldo Cruz (Fiocruz) [Oswaldo Cruz Foundation].

Since this is a population-based study, aggregated data from different geographic levels, including region, state, and capitals were used for all dependent, independent, and control variables. For this study, convenience sampling was used (Warner, 2013). Based on the data retrieved for each geographic level during the first two years of the pandemic, the convenience sample size for regions is 10, for states is 54, and for capitals is 54. Additionally, information on individual-based data was still used to show death distribution throughout the first years of the pandemic according to race, age group, sex, and geographic region, but were used as information only and not used to answer the study research question.

Significance

The findings from this study could provide information on the effects of the population's socioeconomic factors on COVID-19 mortality across Brazil, empowering policymakers, political and public health leaders, and public health practitioners to work on ways to monitor, prevent, and manage disease spread. By knowing about the socioeconomic and demographic impact on disease mortality, preventive measures and public health guidelines can be developed and implemented accordingly. By filling a gap in the impact of SES on COVID-19 mortality in Brazil, leaders and public health professionals could be better equipped to control and prevent rapid disease spread and the likelihood of future pandemics.

The results of this study can demonstrate critical information used to develop protocols for monitoring, prevention, and control. With these results, public health practitioners could justify the needed attention to the distribution of resources required during the pandemic and preventive measures, including social distancing, the use of masks, and lockdowns (Talic et al., 2021). This is critical in reducing the spread of the virus within the Brazilian population, decreasing the mortality rate, and reducing the number of COVID-19 incidences in Brazil. Considering that Brazil is a country of high economic and demographic disparities, the social change implications of this study may include a potential increase in people's opportunities for socioeconomic equality. A decrease in the COVID-19 mortality rate will also allow Brazilians a gradual transition toward normalcy.

Summary

Brazil has a considerably high COVID-19 mortality rate, which may be associated with the population's SES and demographic characteristics. This quantitative study aimed to investigate the association between average income, education, employment rate, and of COVID-19 mortality rate after controlling for race, age, sex, and immunization rate. The theory that grounds this study is the fundamental causes theory. For my research design, I accessed public dataset repositories to obtain data on COVID-19 mortality and the population's demographic and socioeconomic characteristics.

These data were accessed from governmental websites. A power analysis was conducted to determine the sample size suitable to detect the effect of the statistical test used at the established level of significance. Limitations, challenges, and barriers relevant to this research study, particularly during secondary data retrieval from public dataset repositories, were (a) working with relatively large datasets; (b) finding suitable datasets that align with the study's research questions; (c) because the COVID-19 pandemic is an ongoing issue, not all data were available; (d) some datasets were not ready for analysis, and (e) some datasets present missing data.

Part 2: Manuscripts

Manuscript 1: Association of Income and COVID-19 Mortality in Brazil.

Marcelina Machado

Walden University

Outlet for Manuscript

The manuscript on income and COVID-19 mortality in Brazil will be submitted to the *American Journal of Public Health*, https://ajph.aphapublications.org/. The journal requires the original research results reported in up to 3500 words in the text, a structured abstract, up to four tables and figures combined, and no more than 35 free references. The journal also requires AMA formatting with 1.5 or double spacing and font size of 12. The American Journal of Public Health is aligned with the mission to advance public health research, policy, practice, and education. The present manuscript can provide valuable insights for public health practitioners and policymakers, allowing them to evaluate the association between public demographics, income inequalities, and COVID-19 mortality across Brazil.

Abstract

Objective. To investigate the association between the Brazilian population's income and COVID-19 mortality while controlling for variables related to race, age, sex, and immunization.

Methodology. In this retrospective study, average income data were retrieved from the Brazilian Institute of Geography and Statistics; COVID-19 mortality data were retrieved from the Health Surveillance Secretariat; population demographics data including race, age, and sex, were retrieved from the Ministry of Health; and immunization data were retrieved from the Oswaldo Cruz Foundation. All data were analyzed at region, state, and capital-levels using multiple linear regression on IBM SPSS Statistics version 27.

Results. The findings suggest a statistically significant association between income and COVID-19 mortality at the state and capital levels (p=0.026 and p=0.003, respectively). At a state level, for each unit increase in average income, an increase of 0.035 in COVID-19 death per 100,000 people can be predicted, while at a capital level, a decrease of 0.077 in COVID-19 death per 100,000 people can be predicted. Variables related to race, age, sex, and immunization were used as control variables. Results show that the model is invalid to predict COVID-19 mortality from income at the regional level.

Conclusion. To better explore the influences of income on COVID-19 mortality, further research is recommended with both individual and population-based approaches. Since COVID-19 is an ongoing pandemic, it is likely that more accurate results will only be obtained with post-pandemic data.

***Keywords*:** COVID-19 mortality, income, Brazil.

Introduction

Brazil's income inequality is considerably high compared to international standards, which could affect the public's ability to access healthcare resources to prevent and control disease spread (Miranda et al., 2020; OECD, 2020; Rocha et al., 2021). The issue that prompted me to search the literature is how the Brazilian population experienced a significantly high COVID-19 mortality rate, which could be associated with the population's socioeconomic disparities (Miranda et al., 2020). Based on this issue, I evaluated the association income could have on COVID-19 mortality in Brazil, while accounting for the population's race, age, sex, and immunization.

The significance of this study lies in the fact that its findings could help public health practitioners, politicians, and policymakers to target the areas in the country where this association is more significant and focus on developing and implementing strategies for better income distribution among groups socioeconomically disadvantaged. The concepts of the fundamental causes theory were used herein as it relates to the COVID-19 mortality rate and the effect it has on the public's ability to access essential resources that could produce health benefits to prevent disease (Link et al., 1995). This topic can make an original contribution to the research community, considering that the social change implications of this study may include a potential increase in opportunities for income distribution equality for Brazilians and, consequently, a possible decrease in the COVID-19 mortality rate; this would allow each region in the country a gradual transition toward normalcy. Additionally, income distribution equality would potentially decrease poverty and improve housing conditions to individuals and their families, enabling better infection control within households (Martins-Filho et al., 2020). Thus, my purpose was to investigate whether this association exists in Brazil, while controlling for the population's demographics and immunization.

Researchers have indicated that Brazil's advances in healthcare are unequal across the country, being worse in groups socioeconomically disadvantaged (Barreto et al., 2011; Paim et al., 2011). Rocha et al. (2021) examined the association between population vulnerability, health system preparedness, and COVID-19 mortality rates and found that COVID-19 death was higher where the socioeconomic vulnerability was more elevated. Although there are studies on the association between socioeconomic disparities and disease mortality, there is still limited evidence regarding income inequality and its impact on COVID-19 mortality in Latin-American countries, including Brazil. Thus, this gap needs to be studied.

Research Question & Hypotheses

This quantitative study addresses the following research question:

Research Question for Manuscript 1 (RQM1): What is the association between average income and COVID-19 mortality rate at different geographic levels across Brazil after controlling for race, age, sex, and immunization rate?

Based on the research question, the following hypotheses were formulated:

Null Hypothesis (H0): There is no statistically significant association between average income and COVID-19 mortality rate at different geographic levels across Brazil after controlling for race, age, sex, and immunization rate.

Alternative Hypothesis (H1): There is a statistically significant association between average income and COVID-19 mortality rate at different geographic levels across Brazil after controlling for race, age, sex, and immunization rate.

The rationale for this research question is to evaluate if Brazil's high-income inequality can adversely affect COVID-19 mortality. Moreover, there were few studies exploring this association in adjunction with immunization and other demographic factors. Therefore, to address the research question, this research included a retrospective cohort design (Aschengrau et al., 2020) with mortality rates examined between the period of March 2020 to December 2021. This quantitative analysis should help to determine if income, population demographic characteristics, and population immunization are correlated to COVID-19 mortality in Brazil.

Methodology

Study Population

The target population for this study include Brazilian individuals with different demographic and geographic characteristics, including race (White, African origin, and Mixed), age (0-14, 15-49, and 50-80 years old), sex (male and female), and immunization. Public dataset repositories were used for this research design to obtain data on the COVID-19 mortality population's socioeconomic, demographic, and geographic characteristics. Since this is a population-based study, aggregated data from different geographic levels, including region, state, and capitals were used for all dependent, independent, and control variables. For this study, convenience sampling was used (Warner, 2013). Based on the data retrieved for each geographic level during the first two years of the pandemic, the convenience sample size for regions is 10, for states is 54, and for capitals is 54. Additionally, information on individual-based data was still used to show death distribution throughout the first years of the pandemic according to race, age group, sex, and geographic region, but were used as information only and not to answer the study research question.

Variables

The variables of this study include the following.

Independent variable (IV):

- Average Income: Monthly average income values in Brazilian currency from different geographic levels (regional, state, and capital levels) were used as found in the data source.

Dependent variable (DV):

- COVID-19 mortality rate: The mortality rate was calculated for different geographic areas by dividing the number of deaths from COVID-19 in each area by the number of people in the population, then multiplied by 100,000.

Control variables (CV1 – CV8):

- White population: Numbers in percentage. The percentage of the White population was calculated for different geographic areas by dividing the number of White people in each area by the number of people in the population, then multiplied by 100.
- African origin population: Numbers in percentage. The percentage of the African-origin population was calculated for different geographic areas by dividing the

number of African-origin people in each area by the number of people in the population, then multiplied by 100.

- Mixed-race population: Numbers in percentage. The percentage of the Mixed-race population was calculated for different geographic areas by dividing the number of Mixed-race people in each area by the number of people in the population, then multiplied by 100.
- Population between 0-14 years old: Numbers in percentage. The percentage of the 0-14 years old population was calculated for different geographic areas by dividing the number of 0-14 years old people in each area by the number of people in the population, then multiplied by 100.
- Population between 15-49 years old: Numbers in percentage. The percentage of the 15-49 years old population was calculated for different geographic areas by dividing the number of 15-49 years old people in each area by the number of people in the population, then multiplied by 100.
- Population between 50-80 years old: Numbers in percentage. The percentage of the 50-80 years old population was calculated for different geographic areas by dividing the number of 50-80 years old people in each area by the number of people in the population, then multiplied by 100.
- Male population: Numbers in percentage. The percentage of the male population was calculated for different geographic areas by dividing the number of male people in each area by the number of people in the population, then multiplied by 100.

- Immunization rate: Number in percentage. A higher percentage indicates a higher immunization. Actual percentage values from different geographic levels (regional and state levels) were used as found in the data source There were no codes from missing data.

Sources of Data

Data on average income per geographic location from 2020 and 2022 were accessed on September 18, 2022, from the publicly available dataset repository below:

- Instituto Brasileiro de Geografia e Estatistica (IBGE) [Brazilian Institute of Geography and Statistics].

Data on COVID-19 mortality per geographic location from March 2020 through December 2021 were accessed on September 18, 2022, from the publicly available dataset repository below:

- Secretaria de Vigilância em Saúde [Health Surveillance Secretariat], created by the Ministério da Saúde [Ministry of Health].

Data on the population's demographics, including race, age, and sex per geographic location from 2020 and 2021, were accessed on September 18, 2022, from the publicly available dataset repositories below:

- Instituto Brasileiro de Geografia e Estatistica (IBGE) [Brazilian Institute of Geography and Statistics].
- Tabnet – DATASUS - Ministério da Saúde [Ministry of Health].

Data on the population's immunization (or percentage of population vaccinated with the second dose) per geographic location from March 2021 through September 2022 were accessed on September 18, 2022, from the publicly available dataset repository below:

- Fundação Oswaldo Cruz (Fiocruz) [Oswaldo Cruz Foundation].

Access to these dataset repositories were needed to gather specific data points, including the number of deaths by COVID-19 at regional, state and capital-levels, as well as the population's socioeconomic and demographic information. To minimize potential issues regarding the use of secondary data, datasets were screened for errors, outliers, inconsistencies, and missing values prior to analysis. Identified errors in the datasets were reported in the discussion section.

Validity and Reliability of Data

Internal validity, which is "the degree to which the results of the study can be used to make causal inferences," is often weak in non-experimental studies (Warner, 2013). To increase internal validity by limiting the influence of confounding, a series of control variables related to race, age, sex, and immunization were included in the analysis. Additionally, regression analyses were performed separately at each geographic level (region, state, and capital) to avoid violation of internal validity. On the other hand, external validity, the "degree to which the study's results can be generalized" (Warner, 2013), could be considered high in this study as the data used represent a large and heterogeneous population (Oliveira, 2020).

The reliability of secondary data depends on the sources from which the data were obtained. Datasets from government sources are often reputable; however, data screening is still recommended to identify errors and inconsistencies (Warner, 2013). The datasets used in this study were retrieved from Brazilian governmental websites listed in the previous section. Also, data were reviewed before analysis to ensure the accuracy of data transcription.

Design and Analysis

A multiple linear regression (MLR) analysis was performed using IBM SPSS Statistics version 27 to determine the association between average income and COVID-19 mortality rate at different geographic levels across Brazil after controlling for variables related to race, age, sex, and immunization rate. Multiple linear regression (MLR) is a statistical technique often used to evaluate the relationship between one continuous dependent variable and multiple continuous or categorical independent variables (Warner, 2013).

Although aggregated data (population-based data) were accessed and used in the primary statistical analysis (MLR) to answer the research question previously elaborated, individual-based data on COVID-19 mortality and population demographics, including race, age, sex, and region, were used to evaluate death distribution in the country during the first two years of the pandemic. Additionally, descriptive statistical analyses were performed on the COVID-19 mortality rate and average income variables. Lastly, prior to the performance of the primary analysis, assumptions for MLR were tested. Table 1 and Figure 2 summarize of the proposed analytical test plan and justification.

Table 1

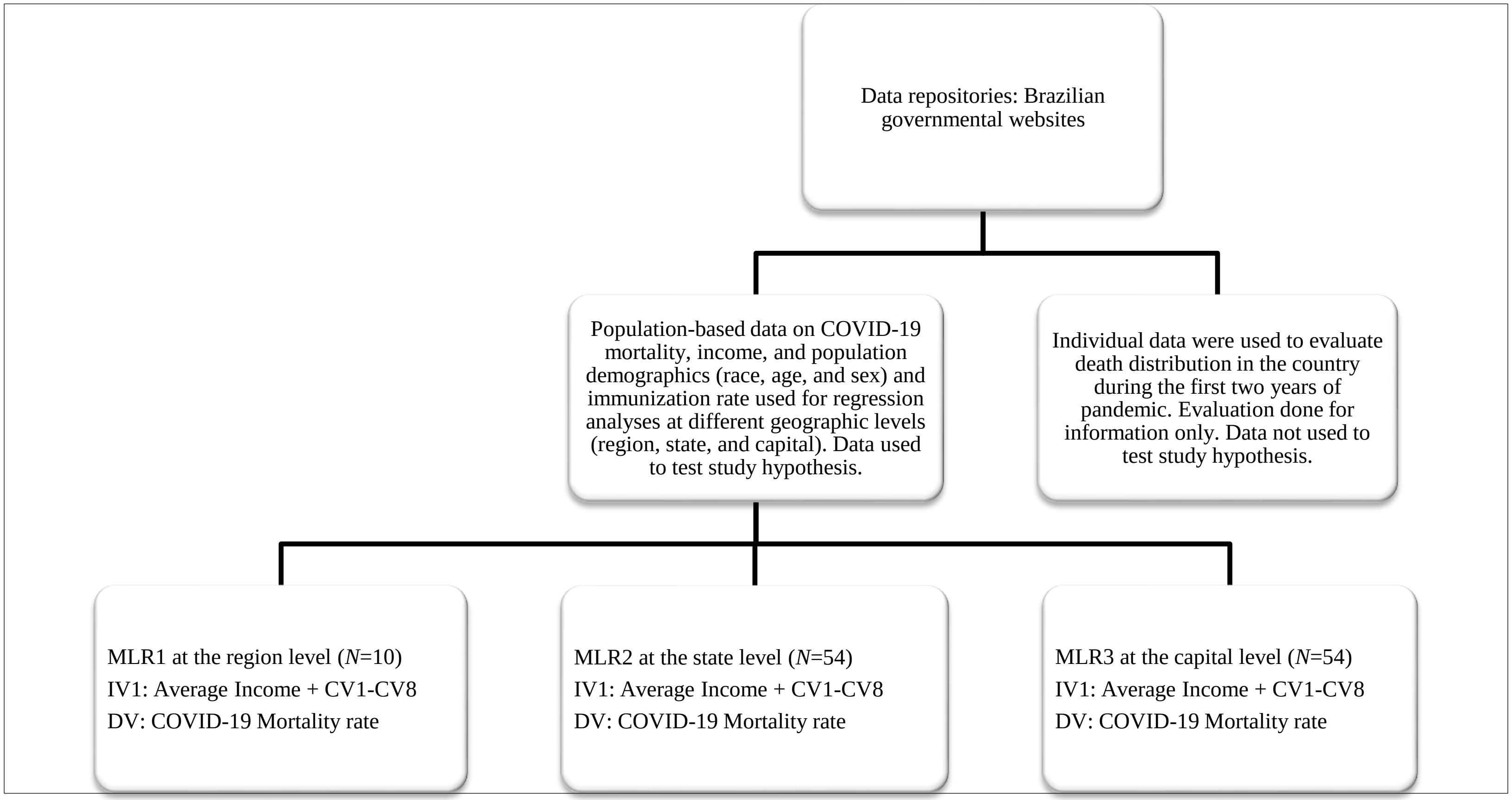
Proposed Analytical Test Plan and Justification

Variable	Type of variable	Multiple Linear Regression (MLR)	Death Distribution Evaluation
COVID-19 Mortality rate [1]	Dependent Variable **DV** (Scale)	Using a population-based dataset and including all variables listed, multiple linear regression analyses were performed to determine the association between income and COVID-19 mortality while controlling for the population's demographic characteristics and immunization. MLR is an appropriate test considering that the dependent variable (disease mortality) and the independent variable (average income) are continuous variables, and the control variables are scale (Warner, 2013). Assumptions tested: Normality, linearity, homoscedasticity, and multicollinearity. In these analyses, all variables listed were used.	An individual-based dataset was created using only COVID-19 mortality and population demographics, including race, age, sex, and region, to evaluate death distribution in the country during the first two years of the pandemic. Data not used to test study hypothesis. Analyses were done for information only.
Average Income [1]	Independent Variables **IV1** (Scale)		
Percentage of White population	Control Variable **CV1** (Scale)		
Percentage of African origin population	Control Variable **CV2** (Scale)		
Percentage of Mixed-race population	Control Variable **CV3** (Scale)		
Percentage of Population between 0-14 Years Old	Control Variable **CV4** (Scale)		
Percentage of Population between 15-49 Years Old	Control Variable **CV5** (Scale)		
Percentage of Population between 50-80 Years Old	Control Variable **CV6** (Scale)		
Percentage of Male Population	Control Variable **CV7** (Scale)		
Percentage of Immunized Population	Control Variable **CV8** (Scale)		

Note. [1] Descriptive statistical analysis were performed on the COVID-19 mortality rate and average income variables.

Figure 2

Scheme of Data Flow and Regression Analyses for Manuscript 1



Note. Regression analyses were performed at each geographic level (region, state, and capital) to avoid internal validity.

Results

The study was conducted by retrieving population-based data on COVID-19 mortality, average income, and population demographics, including race, age, and sex, as well as immunization rate from Brazilian governmental websites, and further analyzing it at different geographic levels (see Figure 2) using a multiple linear regression model on IBM SPSS software. The population-based data was used to create a dataset containing all information for the regression analyses, to test the hypothesis, and answer the research question, that is, if an association between income and COVID-19 mortality exists at different geographic levels in Brazil. The same data was used to further create an individual-based dataset only on COVID-19 mortality and population demographics to investigate the death distribution during the first two years of the pandemic.

Most data needed for the analyses were retrieved without issues, except for the immunization data, which were not available at the capital level; therefore, this variable was not included in the regression analysis for this level. Additionally, as vaccination campaigns started on January 2021 in Brazil, the data retrieved for this variable are from 2021 and 2022. It is also important to note that the regression at the national level was not performed due to the small sample size.

Descriptive Statistics for COVID-19 Mortality

After running a descriptive statistics analysis on the variable COVID-19 mortality, it was observed that among the 27 states during 2 years of the pandemic (N= 54), the COVID-19 mortality rate ranged from 0.063% to 0.284% at the state level. The three measures of central tendency for this variable were mean equals 0.14% (M = 0.14), median equals 0.13% (Median = 0.13), and mode equals 0.063% (Mode = 0.063) with a standard deviation equal 0.058 (*SD* = 0.058). The skewness and kurtosis of this distribution (0.816 and -0.284, respectively) suggest that the distribution deviates from normality and is positively skewed (Table 2 and Figure 3). In Figure 3, the variability or measure of spread distribution of the COVID-19 mortality rate was 0.18% to 0.10%; outliers were not identified.

In a second descriptive statistics analysis, it was observed that among the 27 capitals during the 2 years of the pandemic (*N*= 54), the COVID-19 mortality rate ranged from 0.107% to 0.519% at the capital level. The three measures of central tendency for this variable were mean equals 0.25% (*M* = 0.25), median equals 0.23% (Median = 0.23), and mode equals 0.107% (Mode = 0.107) with standard deviation equals 0.094 (*SD* = 0.094). The skewness and kurtosis of this distribution (0.325 and -0.886, respectively) suggest that the distribution deviates from normality and is slightly positively skewed (Table 2 and Figure 4). In Figure 4, the variability or measure of spread distribution of the COVID-19 mortality rate was 0.28% to 0.18%; outliers were identified.

Table 2

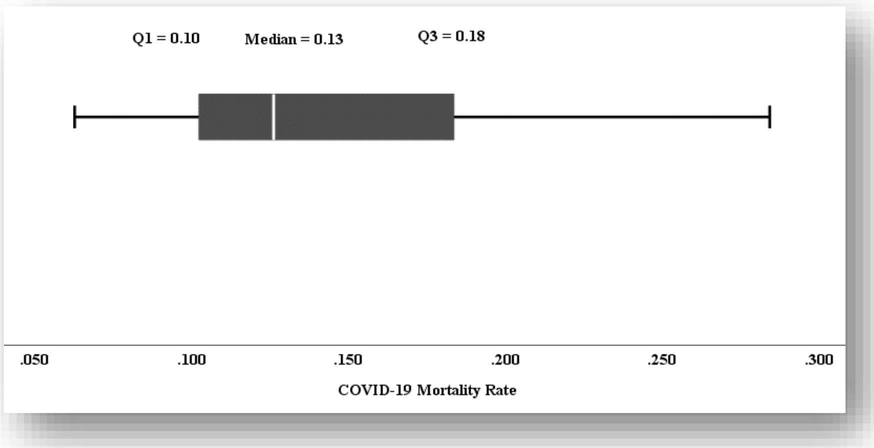
Statistics of COVID-19 Mortality at State and Capital-levels

		State-level	Capital-level
N	Valid	54	54
	Missing	0	0
Mean		0.14169	0.24528
Median		0.12622	0.23056
Mode		0.063[1]	0.107 [1]
Std. Deviation		0.058010	0.093895
Variance		0.003	0.009
Skewness		0.816	0.989
Std. Error of Skewness		0.325	0.325
Kurtosis		-0.284	0.886
Std. Error of Kurtosis		0.639	0.639
Range		0.221	0.411
Minimum		0.063	0.107
Maximum		0.284	0.519
Percentiles	25	0.10052	0.18157
	50	0.12622	0.23056
	75	0.18464	0.28142

Note. Statistics were performed at state and capital levels for the first two years of the pandemic. For that, the *Frequencies* function in SPSS was used to observe Percentile values, Central Tendency, Dispersion, and sample Distribution. [1] Multiple mode exist. The smallest value is shown.

Figure 3

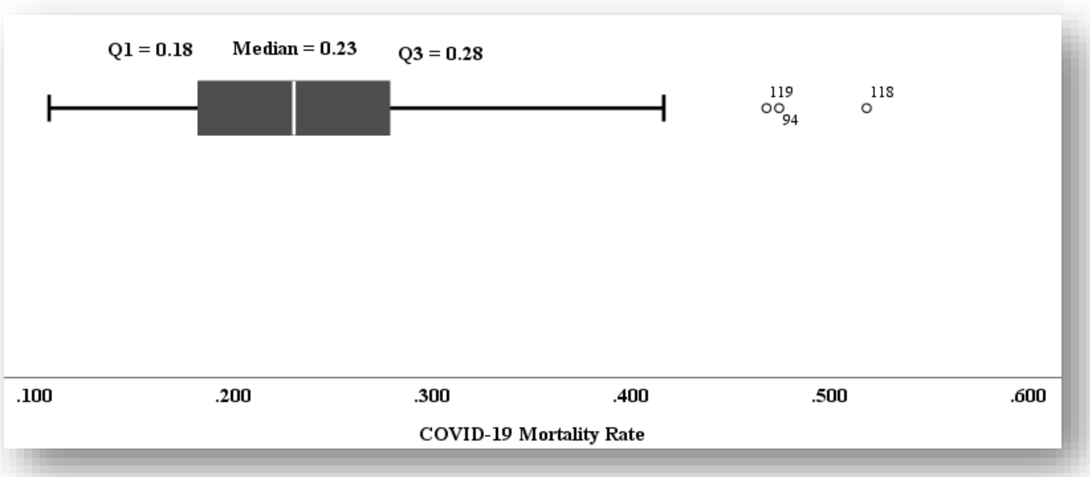
Boxplot Graph Representing a Positively Skewed Distribution of COVID-19 Mortality at State-level



Note. No outliers were identified.

Figure 4

Boxplot Graph Representing a Positively Skewed Distribution of COVID-19 Mortality at Capital-level



Note. Outliers were identified.

Descriptive Statistics for Income

After running a descriptive statistics analysis on the variable income, it was observed that among the 27 states during two years of the pandemic (*N*= 54), the income average ranged from 1595,00 to 4624,00 in Brazilian currency at the state level. The three measures of central tendency for this variable were mean equals 2448,33 (*M* = 2448.33), median equals 2321,50 (Median = 2321.50), and mode equals 1595,00 (Mode = 1595.00) in Brazilian currency, with a standard deviation equals 654.5 (*SD* = 654.5). The skewness and kurtosis of this distribution (1.373 and 2.319, respectively) suggest that the distribution deviates from normality and is positively skewed (Table 3 and Figure 5). In Figure 5, outliers were identified using a boxplot, and the variability or measure of spread distribution of the average income is 2780,00 to 1935,75 in Brazilian currency.

In a second descriptive statistics analysis, it was observed that among the 27 capitals during the two years of the pandemic (*N*= 54), the income average ranged from 2130,00 to 5129,00 in Brazilian currency at the capital level. The three measures of central tendency for this variable were mean equals 3230,02 (*M* = 3230.02), median equals 2997,50 (Median = 2997.50), and mode equals 2130 (Mode = 2130) in Brazilian currency, with standard deviation equals 821.2 (*SD* = 821.2). The skewness and kurtosis of this distribution (0.685 and -0.759, respectively) suggest that the distribution deviates from normality and is positively skewed (Table 3 and Figure 6). In Figure 6, the variability or measure of spread distribution of the average income was 3920,25 to 2497,75 in Brazilian currency; outliers were not identified. Note that for both descriptive analyses, the high value of variance (the square root of the standard deviation) may indicate highly scattered data around the mean.

Table 3

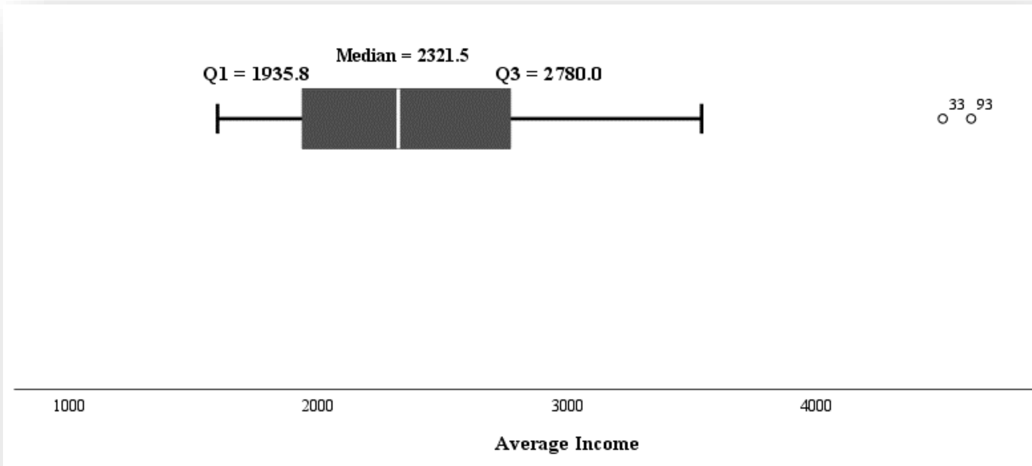
Statistics of Average Income at State and Capital-levels

		State-level	Capital-level
N	Valid	54	54
	Missing	0	0
Mean		2448.33	3230.02
Median		2321.50	2997.50
Mode		1595[a]	2130[a]
Std. Deviation		654.530	821.168
Variance [b]		428410.038	674317.113
Skewness		1.373	0.685
Std. Error of Skewness		0.325	0.325
Kurtosis		2.319	-0.759
Std. Error of Kurtosis		0.639	0.639
Range		3029	2999
Minimum		1595	2130
Maximum		4624	5129
Percentiles	25	1935.75	2497.75
	50	2321.50	2997.50
	75	2780.00	3920.25

Note. Statistics were performed at state and capital levels for the first two years of the pandemic. For that, the *Frequencies* function in SPSS was used to observe further Percentile values, Central Tendency, Dispersion, and sample Distribution. a. Multiple modes exist. The smallest value is shown. b. Note that for both descriptive analyses, the high value of variance (the square root of the standard deviation) may indicate highly scattered data around the mean.

Figure 5

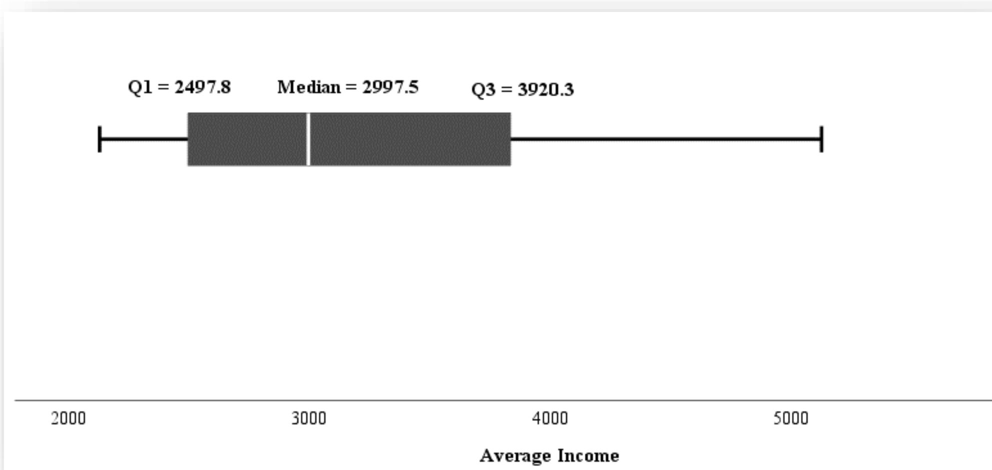
Boxplot Graph Representing a Positively Skewed Distribution of Average Income at State-level



Note. Outliers were identified.

Figure 6

Boxplot Graph Representing a Positively Skewed Distribution of Average Income at Capital-level



Note. No outliers were identified.

Assumptions of Multiple Linear Regression

The assumptions tested for the MLR analyses were normality, linearity, homoscedasticity, and multicollinearity. Normality was tested by observing the Normal P-P Plot, linearity and homoscedasticity were tested by observing the scatterplot of residuals, and multicollinearity was tested by observing the variance inflation factor (VIF) obtained from the coefficient table in the SPSS output. Additionally, the independence of residuals was tested by observing the Durbin-Watson value. Outliners were tested by looking for data points that were far from the regression line.

Table 4 summarizes the results for the assumptions of the multiple linear regression analyses. Note that the passing criterion set for normality was that residuals should be normally distributed. The passing criterion for linearity was that residuals should have a straight-line relationship with predicted DV scores. The passing criterion for homoscedasticity was that VIF values should be less than 10 (VIF<10). The passing criterion for the independence of residuals is that plotted standardized residuals are randomly scattered. The passing criterion for multicollinearity was that Durbin-Watson values should be between 1.5 – 2.5 (approximate values were acceptable). Overall, the results show that except for linear regression at the region-level, none of the assumptions were violated. Although not relevant, significant results of the analysis at the regional level were still reported in this study.

Table 4

Results of Assumptions of Multiple Linear Regression for Manuscript 1

Level	Regression/IV and DV	Normality (passing criterion is that residuals should be normal distributed) **(Pass or fail criterion)**	Linearity (passing criterion is that residuals should have a straight-line relationship with predicted DV scores) **(Pass or fail criterion)**	Homoscedasticity (passing criterion is that plotted standardized residuals are randomly scattered) **(Pass or fail criterion)**	Multicollinearity (passing criterion is VIF <10) **(Pass or fail criterion)**	Independence of Residuals (passing criterion is Durbin-Watson values between 1.5-2.5[1]) **(Pass or fail criterion)**	Outliers (passing criterion is absence of outliers) **(Pass or fail criterion)**
Region	Regression 1 (Average Income and COVID-19 mortality)	Fail	Fail	Fail	Fail	Fail	Pass
State	Regression 2 (Average Income and COVID-19 mortality)	Pass	Pass	Pass	Pass[2]	Pass	Pass
Capital	Regression 3 (Average Income and COVID-19 mortality)	Pass	Pass	Pass	Pass[2]	Pass	Pass

Note. [1]Approximate values are acceptable. [2] Failed for two non-significant variables.

Results of Multiple Linear Regression

To approach the association between average income and COVID-19 mortality in Brazilian regions, a multiple linear regression analysis was conducted to evaluate the prediction of COVID-19 mortality rate from income while controlling for variables related to the population's race, age, sex, and immunization rate (see list of variables included in the Design and Analysis section). Preliminary analyses were conducted to assess the assumptions of normality, linearity, homoscedasticity, multicollinearity, independence of residuals, and outliers. Violations were noted, indicating that the model is invalid in predicting COVID-19 mortality at the regional level (see Table 4). Therefore, regression results for this level were not reported as they cannot be used to interpret the data correctly.

The same approach was taken to evaluate the association between average income and COVID-19 mortality rate in Brazilian states. A multiple linear regression analysis was conducted to evaluate the prediction of COVID-19 mortality rate from income while controlling for variables related to the population's race, age, sex, and immunization rate (see control variables listed in Table 5). Preliminary analyses were conducted to assess the assumptions of normality, linearity, homoscedasticity, multicollinearity, independence of residuals, and outliers. No violations were noted (see Table 4). The multiple linear regression analysis results revealed that at the state level, a significant regression equation was found $F(8, 45)=4.565$, $p<0.001$), with an R^2 of 0.448. Table 5 summarizes the results of the regression performed, showing that, at the state level, for each unit increase in average income, measured in Brazilian currency, we can predict an increase of 0.035 in COVID-19 death per 100,000 people. Income is a significant predictor of COVID-19 mortality (p=0.026) while controlling for variables related to the population's race, age, sex, and immunization rate.

Lastly, to approach the association between income and COVID-19 mortality in Brazilian capitals, a multiple linear regression analysis was conducted to evaluate the prediction of COVID-19 mortality rate from income while controlling for variables related to the population's race, age, and sex (see control variables listed in Table 6). Preliminary analyses were conducted to assess the assumptions of normality, linearity, homoscedasticity, multicollinearity, independence of residuals, and outliers. No violations were noted (see Table 4). The multiple linear regression analysis results revealed that at the capital level, a significant regression equation was found F(7, 46) =2.555, p=0.026), with an R^2 of 0.280. Table 6 summarizes the regression results performed, showing that at the capital level, for each unit increase in average income, measured in Brazilian currency, we can predict a decrease of 0.077 in COVID-19 death per 100,000 people. Income is a significant predictor of COVID-19 mortality (p=0.003) while controlling for variables related to the population's race, age, and sex.

Table 5

SPSS Output: Regression Analysis of Income and COVID-19 Mortality at State-level

Coefficients

Model		Unstandardized Coeff. *B*	Unstandardized Coeff. *SE*	Standardized Coeff. β	t	*p*	95% CI for *B* *LL*	95% CI for *B* *UL*
1	(Constant)	-407.180	818.775		-0.497	0.621	-2056	1242
	Average income[1]	0.035	0.015	0.397	2.302	0.026	-0.004	0.066
	% Male population[2]	19.190	10.327	0.327	1.858	0.070	-1.610	39.990
	% Population between 0-14 years[2]	4.752	5.326	0.243	0.892	0.377	-5.975	15.479
	% Population between 15-49 years[2]	-9.173	6.867	-0.311	-1.336	0.188	-23.005	4.658
	% White population[2]	-389.826	415.916	-1.199	-0.937	0.354	-1228	448
	% African Origin population[2]	-273.592	493.849	-0.135	-0.554	0.582	-1268	721
	% Mixed-race population[2]	-297.887	435.277	-0.819	-0.684	0.497	-1175	579
	% Immunized population[2]	3.450	0.808	0.702	4.269	<0.001	1.822	5.078

Note. [1.] Predictor variable [2.] Control variables.

Regression model using the total number of deaths from COVID-19 per 100,000 people at the state-level in Brazil.

Table 6

SPSS Output: Regression Analysis of Income and COVID-19 Mortality at Capital-level

	Coefficients							
		Unstandardized Coeff.		Standardized			95% CI for *B*	
	Model	*B*	*SE*	Coeff. β	t	*p*	*LL*	*UL*
1	(Constant)	2088	1799		1.161	0.252	-1533	5709
	Average income[1]	-0.077	0.024	-0.677	-3.178	0.003	-0.126	-0.028
	% Male population[2]	29.494	13.455	0.451	2.192	0.033	2.410	56.577
	% Population between 0-14 years[2]	-11.498	8.625	-0.333	-1.333	0.189	-28.859	5.863
	% Population between 15-49 years[2]	-22.484	8.317	-0.649	-2.704	0.010	-39.225	-5.774
	% White population[2]	-1555	1440	-2.851	-1.080	0.286	-4453	1343
	% African Origin population[2]	-1576	1483	-0.764	-1.062	0.294	-4561	1409
	% Mixed-race population[2]	-1517	1453	-2.560	-1.044	0.302	-4442	1408

Note. [1.] Predictor variable [2.] Control variables.

Data on immunization was not available at capital-level; therefore, it was not included in this regression.

Regression model using the total number of deaths from COVID-19 per 100,000 people at the capital-level in Brazil.

Distribution of COVID-19 Death

Between March 2020 and December 2021, the Secretaria de Vigilância em Saúde [Health Surveillance Secretariat], created by the Ministério da Saúde [Ministry of Health], reported 634,695 deaths by COVID-19, being 212,706 deaths in 2020 and 421, 989 deaths in 2021. The number of deaths caused by COVID-19 during these first two years of the pandemic corresponds to 18.8% of all deaths in the country in the same period, and the COVID-19 mortality rate observed was 297.5 deaths per 100,000 people in all age groups. The distribution of the casualties according to age groups, sex, race, and Brazilian region are shown in Table 7 – Table 8 and Figure 7 – Figure 14. According to the distribution, in 2020, the number of deaths across the country was higher among individuals who were 80 years or older (1348.3 deaths/100,000), White (112.3 deaths/100,000), and male individuals (117.5 deaths/100,000). Similarly, in 2021 the number of deaths was still higher among individuals who were 80 years or older (1731 deaths/100,000), White (238.5 deaths/100,000), and male individuals (224.8 deaths/100,000). In 2020 death rate was higher in the Southeast region (109.9 deaths/100,000), conversely, in 2021, the death rate was higher in the South (244.6 deaths/100,000).

Note that although individual-based data was used to show death distribution throughout the first years of the pandemic according to race, age group, sex, and geographic region, this is a population-based study and the data in this section was not used to test the study hypothesis or answer the research question. This death distribution evaluation has been for information only.

Table 7

Distribution of COVID-19 Death in 2020 and 2021 according to Age Group

Demographic	[1]*N* (%) in 2020	[2] Death rate per 100,000	[1]*N* (%) in 2021	[2] Death rate per 100,000
Age				
Infant	278 (0.1)	-	347 (0.1)	-
0-4 years	168 (0.1)	1.1	212 (0.1)	1.4
5-9 years	101 (0.0)	0.7	121 (0.0)	0.8
10-14 years	128 (0.1)	0.9	171 (0.0)	1.2
15-19 years	347 (0.2)	2.2	545 (0.1)	3.5
20-29 years	1809 (0.9)	5.3	5517 (1.3)	16.2
30-39 years	5975 (2.8)	17.5	21155 (5.0)	61.7
40-49 years	13542 (6.4)	46.3	45392 (10.8)	152.0
50-59 years	26600 (12.5)	111.4	76859 (18.2)	317.1
60-69 years	47875 (22.5)	286.1	98362 (23.3)	568.7
79-79 years	55979 (26.3)	620.4	93347 (22.1)	991.3
80-older	59880 (28.2)	1348.3	79925 (18.9)	1731.0
Unknown	24 (0.0)	-	36 (0.0)	-
Total	212,706 (100.0)	100.3	421,989 (100.0)	197.6

Note. [1]*N* = number of deaths by COVID-19. Data on death were accessed from Secretaria de Vigilância em Saúde [Health Surveillance Secretariat] on 18Sep2022, using indicator ICD-10-CM code B34.2 for Coronavirus infection, unspecified. Data on total population for each age group were accessed from Ministério da Saúde [Ministry of Heath] on 18Sep2022. [2] Death rate per 100,000 was calculated by diving the number of deaths per the total population, times 100,000 (data for infant and unknown were excluded from the calculation).

Table 8

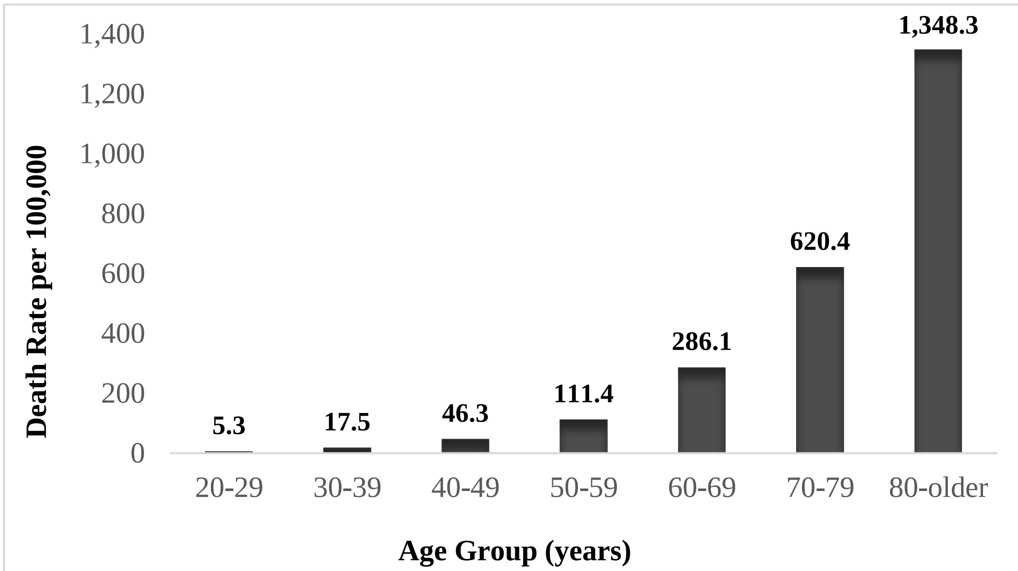
Distribution of COVID-19 Death in 2020 and 2021 according to Sex, Race, and Region

Demographic	[1]*N* (%) in 2020	[2] Death rate per 100,000	[1]*N* (%) in 2021	[2] Death rate per 100,000
Sex				
Male	121641 (57.2)	117.5	234352 (55.5)	224.8
Female	91051 (42.8)	84.1	187588 (44.5)	172.0
Unknown	14 (0.0)	-	49 (0.0)	-
Race				
White	103525 (48.7)	112.3	218852 (51.9)	238.5
African origin	18604 (8.7)	99.7	32149 (7.6)	154.1
Asian origin	1458 (0.7)	-	2591 (0.6)	-
Mixed-race	81572 (38.3)	82.7	139288 (33.0)	142.0
Brazilian Native	972 (0.5)	-	851 (0.2)	-
Unknown	6575 (3.1)	-	28258 (6.7)	-
Region				
Central-West	17381 (8.2)	105.3	39559 (9.4)	236.8
Northeast	54697 (25.7)	95.3	76020 (18.0)	131.8
North	19531 (9.2)	104.6	29861 (7.1)	157.9
Southeast	97798 (46.0)	109.9	202174 (47.9)	225.6
South	23299 (11.0)	77.2	74375 (17.6)	244.6
Total	212706 (100.0)	100.4	421989 (100.0)	197.8

Note. [1]*N* = number of deaths by COVID-19. Data on death were accessed from Secretaria de Vigilância em Saúde [Health Surveillance Secretariat] on 18Sep2022, using indicator ICD-10-CM code B34.2 for Coronavirus infection, unspecified. Data on total population for each sex and region groups were accessed from Ministério da Saúde [Ministry of Heath] on 18Sep2022. Data on total population for each race group were accessed from the Sistema IBGE de Recuperação Automática [IBGE Automatic Recovery System] on 18Sep2022. [2] Death rate per 100,000 was calculated by diving the number of deaths per the total population, times 100,000 (data for unknown, Asian, and Brazilian native population were excluded from the calculation).

Figure 7

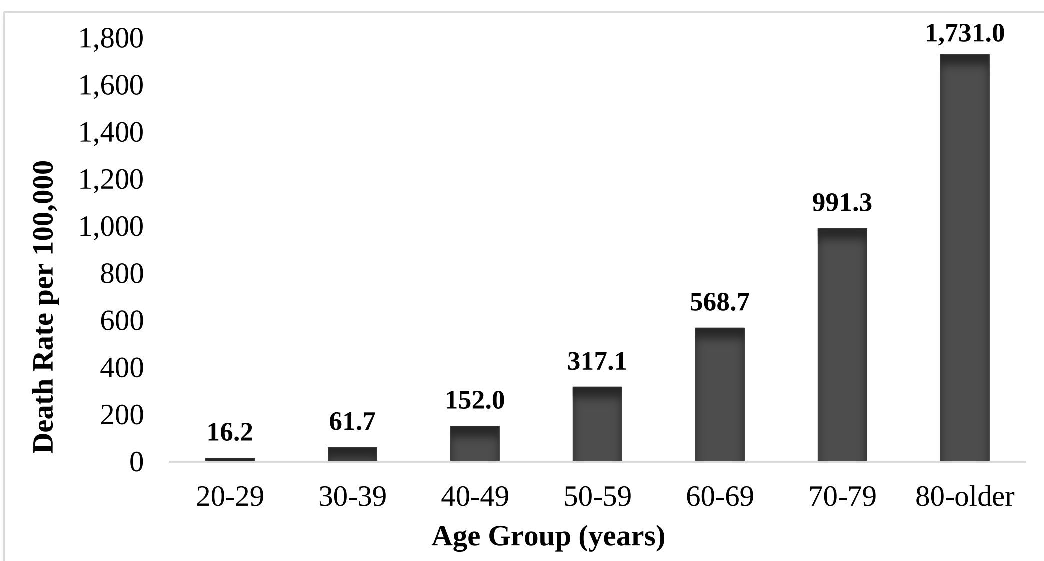
Bar Chart of COVID-19 Death Rate per 100,000 in 2020 by Age Group in Brazil



Note. Only rates > 5 are displayed.

Figure 8

Bar Chart of COVID-19 Death Rate per 100,000 in 2021 by Age Group in Brazil



Note. Only rates > 5 are displayed.

Figure 9

Pie Chart of COVID-19 Death Rate per 100,000 in 2020 by Sex in Brazil

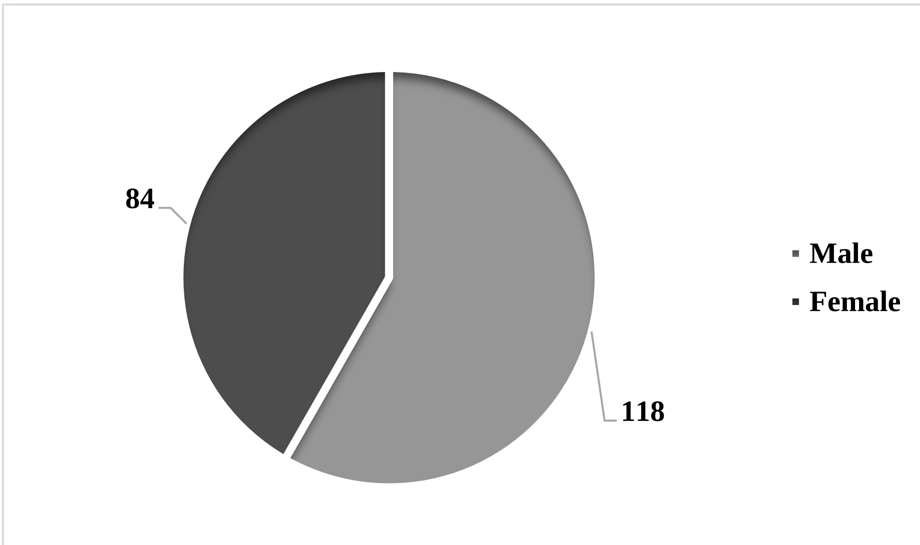


Figure 10

Pie Chart of COVID-19 Death Rate per 100,000 in 2021 by Sex in Brazil

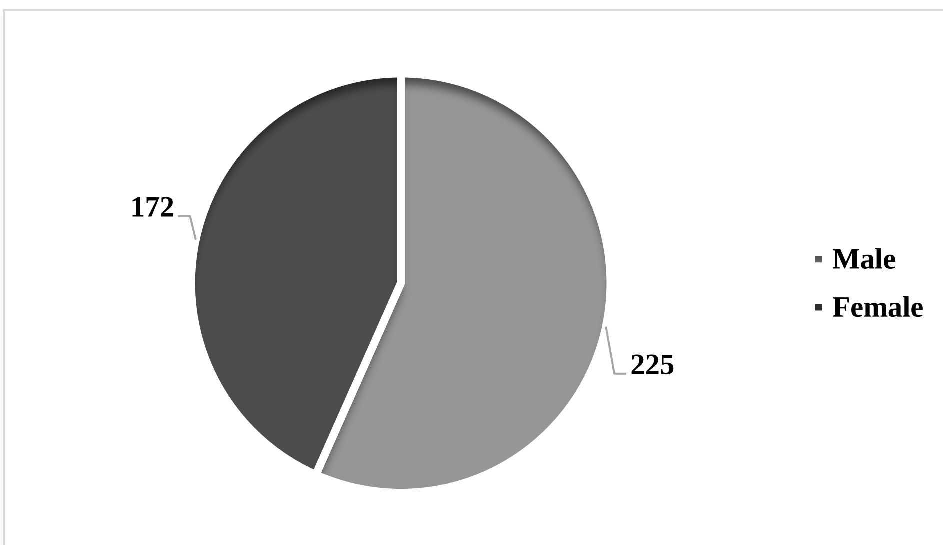


Figure 11

Bar Chart of COVID-19 Death Rate per 100,000 in 2020 by Race in Brazil

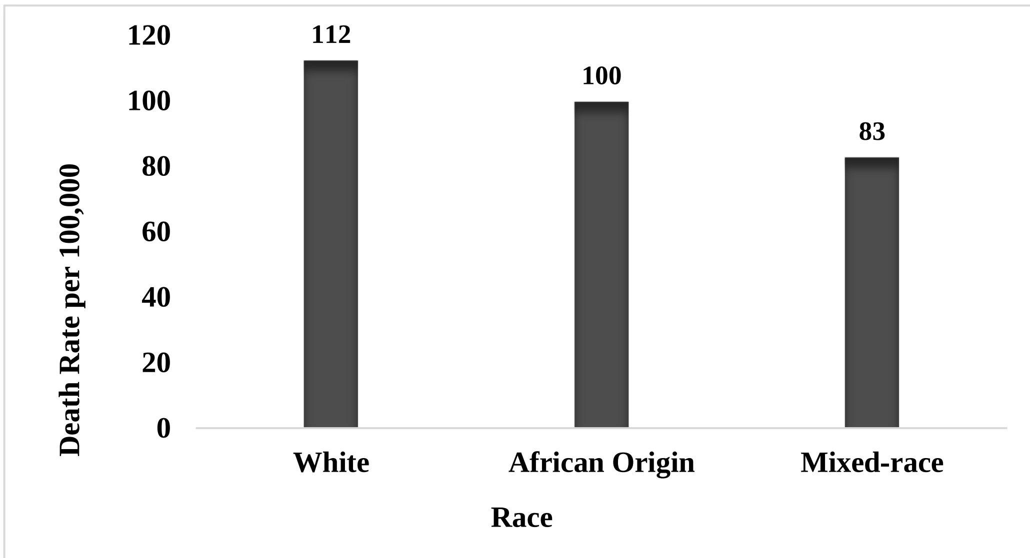


Figure 12

Bar Chart of COVID-19 Death Rate per 100, 000 in 2021 by Race in Brazil

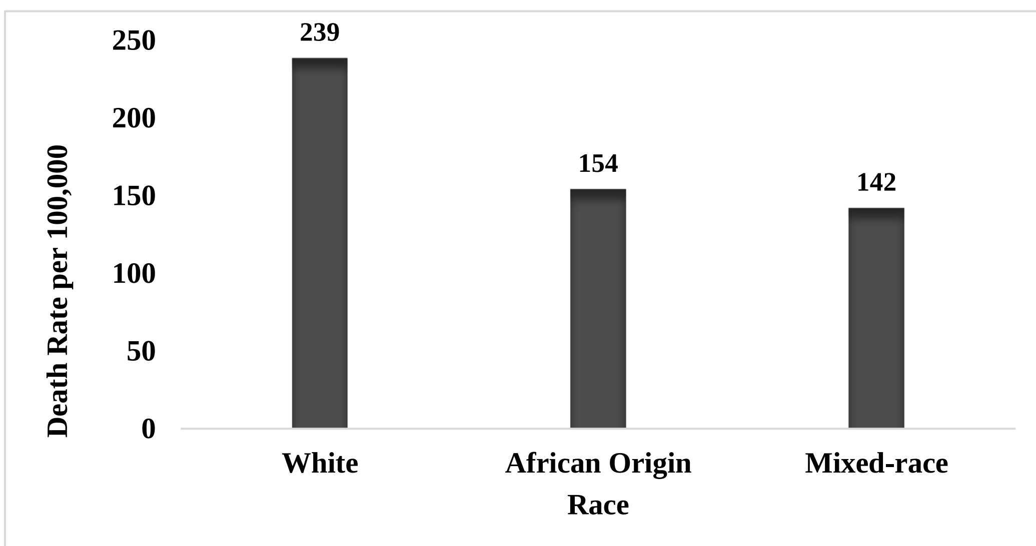


Figure 13

Bar Chart of COVID-19 in 2020 by Region

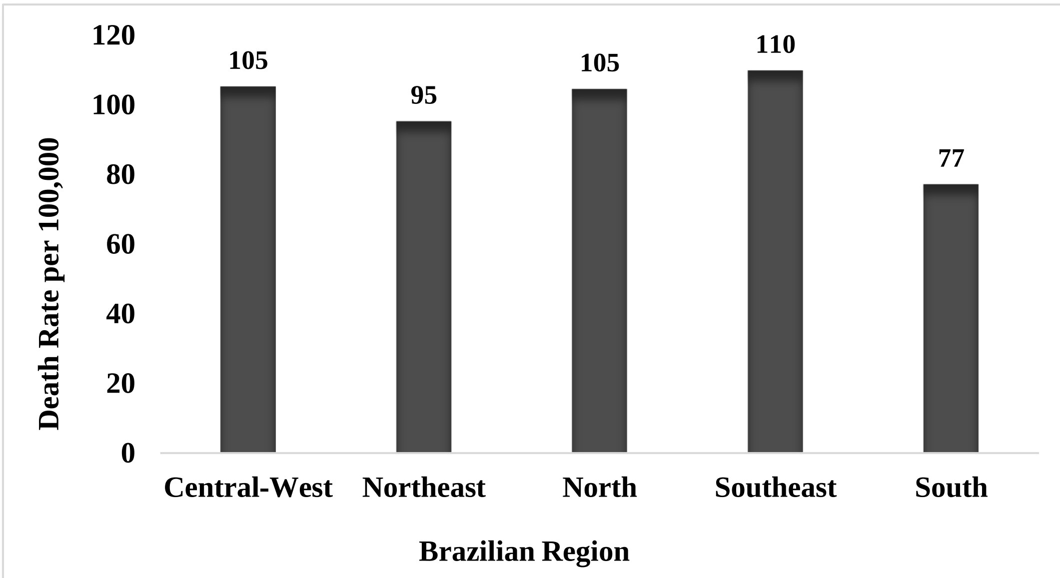
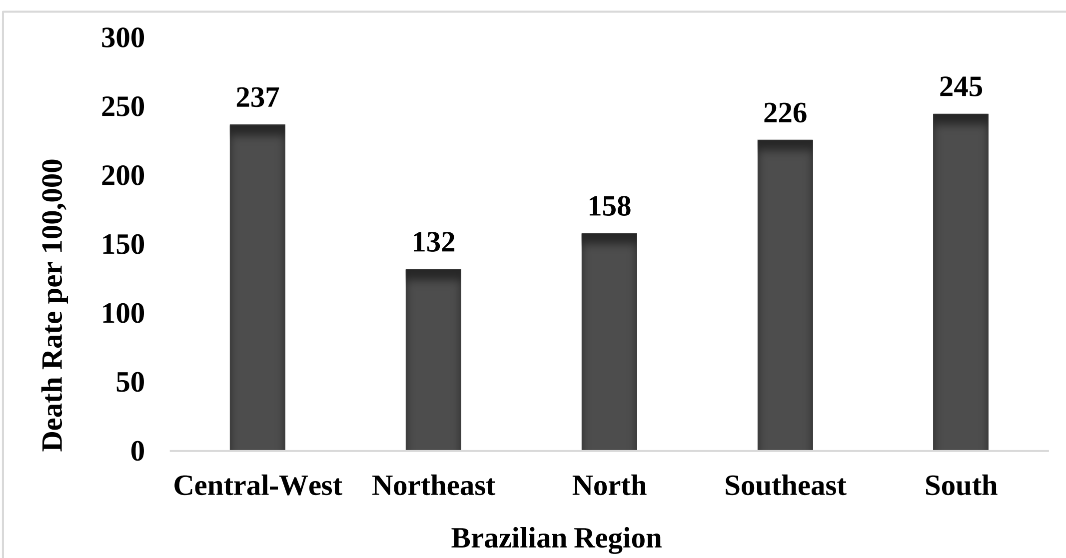


Figure 14

Bar Chart of COVID-19 in 2021 by Region



Discussion

Based on the fundamental causes theory, disease spread and mortality are greater among individuals with fewer resources (Link et al., 1995). Therefore, mortality due to COVID-19 is expected to be higher among individuals who are socioeconomic disadvantaged or part of minority groups (Mello, 2020; Ribeiro et al., 2021). During multiple linear regression analyses to investigate the association between income and COVID-19 mortality in Brazilian states and capitals while controlling for variables related to race, age, sex, and immunization rate, a statistically significant association between income and the COVID-19 mortality rate was found. At the state level, an increase in income suggested an increase in the COVID-19 death rate. The increase in mortality in areas of higher income could be explained by Bermudi et al.'s (2021) study, which shows that high-mortality risk areas tend to shift from the best to the worst socioeconomic areas. Results at the regional level were not reported, as the model was found to be invalid in predicting COVID-19 mortality from income at this level, probably due to the small sample size.

Looking at the distribution of COVID-19 mortality in Brazil during the first two years of the pandemic, it was found that the COVID-19 death rate was higher in the Southeast region in 2020 and higher in the South region in 2021, the most densely populated regions in the country. Based on this observation, we can suggest that the irregularity or heterogeneity of the population's distribution can shape disease spread, as investigated by Martins-Filho et al. (2021) and Miranda et al. (2020). The same analysis found that the COVID-19 death rate was higher among individuals 80 years old or older, male and White, in both years of the pandemic.

As mentioned in the Summary of the Manuscripts section, most of the study limitations were associated with data availability. Two-year data for disease mortality were obtained for each region, state, and capital. Although the data used for the regression analyses were population-based, the data obtained at a national level was converted to an individual-level dataset to investigate death distribution for the first two years of the pandemic. To avoid errors during the manipulation of this large dataset, the data were reviewed a second time to ensure transcription errors were not present. Another limitation was regarding data availability. Since COVID-19 is an ongoing pandemic, many data points were unavailable or ready for analyses, resulting in a need for some manipulation to guarantee dataset completion and readiness before analyses. Data manipulation was required to avoid internal validity issues, but data should be considered reliable as all information was retrieved from governmental resources.

The findings of this study empower public health and political leaders to identify geographic areas in the country that need special attention regarding the development of policies and implementation of public health programs to encourage populations' best behavior and practices concerning preventive actions and vaccination willingness. Furthermore, considering Brazil a country of high economic and demographic disparities, the social change implications of this study may include a potential increase in people's opportunities for socioeconomic equality, particularly a fair income distribution, which could allow individuals to have better access to healthcare, consequently decreasing

COVID-19 mortality rates. A decrease in the COVID-19 mortality rate could allow Brazilians to improve individual lives with healthier families and communities.

Based on this study's findings and to better explore the influences of SES as a risk factor for COVID-19 mortality, further research is recommended with both individual and population-based approaches. Since COVID-19 is an ongoing pandemic, it is more likely that more accurate results will be only obtained with post-pandemic data. Lastly, multiple linear regression models may not be the best model for analyzing population-based data considering that sample sizes may be limited compared to individual-based data.

References

Achdut, N., & Refaeli, T. (2020). Unemployment and psychological distress among young people during the COVID-19 pandemic: psychological resources and risk factors. *International Journal of Environmental Research and Public Health, 17*(19), 7163. https://doi.org/10.3390/ijerph17197163

American Psychological Association [APA] (2022). Socioeconomic Status. Retrieved from https://www.apa.org/topics/socioeconomic-status#:~:text=Socioeconomic%20status%20is%20the%20social,to%20privilege%2C%20power%20and%20control

Bambra, C., Riordan, R., Ford, J., & Matthews, F. (2020). The COVID-19 pandemic and health inequalities. *Journal of Epidemiology and Community Health*, *74*(11), 964-968. https://doi.org/10.1136/jech-2020-214401

Baqui, P., Marra, V., Alaa, A. M., Bica, I., Ercole, A., & van der Schaar, M. (2021). Comparing COVID-19 risk factors in Brazil using machine learning: the importance of socioeconomic, demographic, and structural factors. *Scientific Reports*, *11*(1), 1–10. https://doi.org/10.1038/s41598-021-95004-8.

Barberia, L. G., & Gómez, E. J. (2020). Political and institutional perils of Brazil's COVID-19 crisis. *Lancet (London, England)*, *396*(10248), 367. https://doi.org/10.1016/S0140-6736(20)31681-0

Barreto, M. L., Teixeira, M. G., Bastos, F. I., Ximenes, R. A., Barata, R. B., & Rodrigues, L. C. (2011). Successes and failures in the control of infectious diseases in Brazil: social and environmental context, policies, interventions, and research needs. *The Lancet*, *377*(9780), 1877-1889. https://doi.org/10.1016/S0140-6736(11)60202-X

Bartholo, T. L., Koslinski, M. C., de Andrade, F. M., & de Castro, D. L. (2020). Segregação escolar e desigualdades educacionais no início da escolarização no Brasil. REICE. *Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, *18*(4), 57-76. https://doi.org/10.15366/reice2020.18.4.003

Bermudi, P. M. M., Lorenz, C., Aguiar, B. S. de, Failla, M. A., Barrozo, L. V., & Chiaravalloti-Neto, F. (2021). Spatiotemporal ecological study of COVID-19 mortality in the city of São Paulo, Brazil: Shifting of the high mortality risk from areas with the best to those with the worst socio-economic conditions. *Travel Medicine & Infectious Disease*, *39*, 1-8. N.PAG. https://doi.org/10.1016/j.tmaid.2020.101945

Bravata, D. M., Perkins, A. J., Myers, L. J., Arling, G., Zhang, Y., Zillich, A. J., Reese, L., Dysangco, A., Agarwal, R., Myers, J., Austin, C., Sexson, A., Leonard, S. J., Dev, S., & Keyhani, S. (2021). Association of intensive care unit patient load and demand with mortality rates in US Department of Veterans Affairs hospitals during the COVID-19 pandemic. *Journal of the American Medical Association Network Open*, *4*(1), https://doi.org/10.1001/jamanetworkopen.2020.34266. https://doi.org/10.1001/jamanetworkopen.2020.34266

Brazilian Institute of Geography and Statistics [IBGE] (2022). Annual average unemployment rate decreases in most states in 2021. https://agenciadenoticias.ibge.gov.br/en/agencia-news/2184-news-agency/news/33054-annual-average-unemployment-rate-decreases-in-most-states-in-2021

Brazilian Institute of Geography and Statistics [IBGE] (n.d.). Densidade demográfica [Demographic density]. Retrieved from https://www.ibge.gov.br/geociencias/cartas-e-mapas/sociedade-e-economia/15955-densidade-demografica.html?=&t=acesso-ao-produto

Centers for Disease Control and Prevention (n.d.-a). Guidance for Businesses and Employers Responding to Coronavirus Disease 2019 (COVID-19). Retrieved from https://www.cdc.gov/coronavirus/2019-ncov/community/guidance-business-response.html

Cifuentes, M. P., Rodriguez-Villamizar, L. A., Rojas-Botero, M. L., Alvarez-Moreno, C. A., & Fernández-Niño, J. A. (2021). Socioeconomic inequalities associated with mortality for COVID-19 in Colombia: A cohort nationwide study. *Journal of Epidemiology and Community Health*, *75*(7), 610-615. https://doi.org/10.1136/jech-2020-216275

Chang, V. W., & Lauderdale, D. S. (2009). Fundamental Causes theory, technological innovation, and health disparities: the case of cholesterol in the era of statins. *Journal of Health and Social Behavior, 50*(3), 245-260. https://doi.org/10.1177/002214650905000301

Chowell, G., Bettencourt, L. M., Johnson, N., Alonso, W. J., & Viboud, C. (2008). The 1918–1919 influenza pandemic in England and Wales: spatial patterns in transmissibility and mortality impact. *Proceedings of the Royal Society B:*

Biological Sciences, *275*(1634), 501-509. https://doi.org/10.1098/rspb.2007.1477

Clouston, S. A. P., Natale, G., & Link, B. G. (2021). Socioeconomic inequalities in the spread of coronavirus-19 in the United States: A examination of the emergence of social inequalities. *Social Science & Medicine*, *268*, 1-6. https://doi.org/10.1016/j.socscimed.2020.113554

De Paula, A. M. N. R., Coelho, D., de Oliveira, C. L., & Moura, A. P. V. (2021). Should the convergence between the social and biological vulnerabilities of black communities and the COVID-19 pandemic be considered? *Medical Hypotheses, 146*, 1-2, 110359. https://doi.org/10.1016/j.mehy.2020.110359

De Souza, C. D. F., Machado, M. F., & do Carmo, R. F. (2020). Human development, social vulnerability, and COVID-19 in Brazil: A study of the social determinants of health. *Infectious Diseases of Poverty, 9*(04), 50-59. https://doi.org/10.1186/s40249-020-00743-x

Díaz-Castro, L., Cabello-Rangel, H., & Hoffman, K. (2021). The impact of health policies and sociodemographic factors on doubling time of the COVID-19 pandemic in Mexico. *International Journal of Environmental Research and Public Health*, *18*(5), 2354 https://doi.org/10.3390/ijerph18052354

Domingues, C., Maranhão, A., Teixeira, A. M., Fantinato, F., & Domingues, R. (2020). The Brazilian National Immunization Program: 46 years of achievements and challenges. 46 anos do Programa Nacional de Imunizações: Uma história repleta de conquistas e desafios a serem superados. *Cadernos de Saude Publica, 36*Suppl 2(Suppl 2), e00222919. https://doi.org/10.1590/0102-311X00222919

French, G., Hulse, M., Nguyen, D., Sobotka, K., Webster, K., Corman, J., Aboagye-Nyame, B., Dion, M., Johnson, M., Zalinger, B., & Ewing, M. (2021). Impact of hospital strain on excess deaths during the COVID-19 pandemic - United States, July 2020–July 2021. *Morbidity and Mortality Weekly Report*, *70*(46), 1613. https://doi.org/10.15585/mmwr.mm7046a5

Goutte, S., Péran, T., & Porcher, T. (2020). The role of economic structural factors in determining pandemic mortality rates: Evidence from the COVID-19 outbreak in France. *Research in International Business and Finance, 54,* 101281. https://doi.org/10.1016/j.ribaf.2020.101281

Grantz, K. H., Rane, M. S., Salje, H., Glass, G. E., Schachterle, S. E., & Cummings, D. A. (2016). Disparities in influenza mortality and transmission related to sociodemographic factors within Chicago in the pandemic of 1918. *Proceedings of the National Academy of Sciences*, *113*(48), 13839-13844. https://doi.org/10.1073/pnas.1612838113

Greer, S. L., King, E. J., da Fonseca, E. M., & Peralta-Santos, A. (2020). The comparative politics of COVID-19: The need to understand government responses. *Global public health*, *15*(9), 1413-1416. https://doi.org/10.1080/17441692.1783340

Khanijahani, A., Iezadi, S., Gholipour, K., Azami-Aghdash, S., & Naghibi, D. (2021). A systematic review of racial/ethnic and socioeconomic disparities in COVID-19. *International Journal for Equity in Health, 20*(1), 1-30. https://doi.org/10.1186/s12939-021-01582-4

Khalatbari-Soltani, S., Cumming, R. C., Delpierre, C., & Kelly-Irving, M. (2020). Importance of collecting data on socioeconomic determinants from the early stage of the COVID-19 outbreak onwards. *Journal of Epidemiology and Community Health*, *74*(8), 620-623. http://dx.doi.org/10.1136/jech-2020-214297

Link, B. G., & Phelan, J. (1995). Social Conditions As Fundamental Causes of Disease. *Journal of Health and Social Behavior*, 80–94. https://doi.org/10.2307/2626958

Lotta, G., Fernandez, M., Kuhlmann, E., & Wenham, C. (2022). COVID-19 vaccination challenge: what have we learned from the Brazilian process? *The Lancet Global Health*. 1-2. https://doi.org/10.1016/S2214-109X(22)00049-3

Luptáčik, M., & Nežinský, E. (2020). Measuring income inequalities beyond the Gini coefficient. *Central European Journal of Operations Research*, *28*(2), 561-578. https://doi.org/10.1007/s10100-019-00662-9

Mackenbach, J. P., Kulhánová, I., Bopp, M., Deboosere, P., Eikemo, T. A., Hoffmann, R., Kulik, M. C., Leinsalu, M., Martikainen, P., Menvielle, G., Regidor, E., Wojtyniak, B., Östergren, O., & EURO-GBD-SE Consortium. (2015). Variations in the relation between education and cause-specific mortality in 19 European populations: A test of the "Fundamental Causes" theory of social inequalities in health. *Social Science & Medicine*, *127*, 51-62. https://doi.org/10.1016/j.socscimed.2014.05.021

Mackenbach, J. P., Kunst, A. E., Cavelaars, A. E., Groenhof, F., Geurts, J. J., & EU Working Group on Socioeconomic Inequalities in Health. (1997). Socioeconomic inequalities in morbidity and mortality in western Europe. *The Lancet, 349*(9066), 1655-1659. https://doi.org/10.1016/S0140-6736(96)07226-1

Mackenbach, J. P., Looman, C. W. N., Artnik, B., Bopp, M., Deboosere, P., Dibben, C., Kalediene, R., Kovács, K., Leinsalu, M., Martikainen, P., Regidor, E., Rychtaříková, J., & de Gelder, R. (2017). "Fundamental Causes" of inequalities in mortality: an empirical test of the theory in 20 European populations. *Sociology of Health & Illness*, *39*(7), 1117–1133. https://doi.org/10.1111/1467-9566.12562

Mackenbach, J. P., Stirbu, I., Roskam, A. J. R., Schaap, M. M., Menvielle, G., Leinsalu, M., & Kunst, A. E. (2008). Socioeconomic inequalities in health in 22 European countries. *New England Journal of Medicine*, *358*(23), 2468-2481. https://doi.org/10.1056/NEJMsa0707519

Mackenbach, J. P., Bos, V., Andersen, O., Cardano, M., Costa, G., Harding, S., Reid, A., Hemström, O., Valkonen, & Kunst, A. E. (2003). Widening socioeconomic inequalities in mortality in six Western European countries. *International Journal of Epidemiology*, *32*(5), 830-837. https://doi.org/10.1093/ije/dyg209

Martinez-Valle, A. (2021). Public health matters: Why is Latin America struggling in addressing the pandemic? *Journal of Public Health Policy*, *42*(1), 27-40. https://doi.org/10.1057/s41271-020- 00269-4

Martins-Filho, P. R. (2021). Relationship between population density and COVID-19 incidence and mortality estimates: A county-level analysis. *Journal of Infection and Public Health*, *14*(8), 1087. https://doi.org/10.1016/j.jiph.2021.06.018

Martins-Filho, P. R., de Souza Araújo, A. A., Quintans-Júnior, L. J., & Santos, V. S. (2020). COVID-19 fatality rates related to social inequality in Northeast Brazil: a neighborhood-level analysis. *Journal of Travel Medicine*, *27*(7), 1-3. https://doi.org/10.1093/jtm/taaa128.

Martins-Filho, P. R., Quintans-Júnior, L. J., de Souza Araújo, A. A., Sposato, K. B., Souza Tavares, C. S., Gurgel, R. Q., Fontes Leite, D. C., de Paiva, S. M., Santos, H. P., & Santos, V. S. (2021). Socio-economic inequalities and COVID-19 incidence and mortality in Brazilian children: a nationwide register-based study. *Public Health (Elsevier)*, *190*, 4–6. https://doi.org/10.1016/j.puhe.2020.11.005

Mateo-Urdiales, A., Fabiani, M., Rosano, A., Vescio, M. F., Del Manso, M., Bella, A., Riccardo, F., Pezzotti, P., Regidor, E., & Andrianou, X. (2021). Socioeconomic patterns and COVID-19 outcomes before, during, and after the lockdown in Italy (2020). *Health & Place*, 71, 102642. https://doi.org/10.1016/j.healthplace.2021.102642

Mello, D. (05 May 2020). *Risco de morrer por coronavírus varia até 10 vezes entre bairros de SP* [Risk of dying from coronavirus varies up to 10 times between neighborhoods in SP] [Internet]. Agência Brasil [Brazil Agency]. Retrieved from Risco de morrer por coronavírus varia até 10 vezes entre bairros (ebc.com.br)

Mena, G. E., Martinez, P. P., Mahmud, A. S., Marquet, P. A., Buckee, C. O., & Santillana, M. (2021). Socioeconomic status determines COVID-19 incidence and related mortality in Santiago, Chile. *Science*, *372*(6545), eabg5298. https://doi.org/10.1126/science.abg5298

Miech, R., Pampel, F., Kim, J., & Rogers, R. G. (2011). The enduring association between education and mortality: the role of widening and narrowing disparities.

American Sociological Review, *76*(6), 913-934. https://doi.org/10.1177/0003122411411276

Ministry of Health (n.d.). Sistema Único de Saúde (SUS): Estrutura, princípios e como funciona. https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-z/s/sus-estrutura-principios-e-como-funciona

Miranda Demenech, L., de Carvalho Dumith, S., Centena Duarte Vieira, M. E., & Neiva-Silva, L. (2020). Income inequality and risk of infection and death by COVID-19 in Brazil. *Revista Brasileira de Epidemiologia*, 1–12. https://doi.org/10.1590/1980-549720200095

Moore, D. C. B. C., Nehab, M. F., Camacho, K. G., Reis, A. T., de Fátima Junqueira-Marinho, M., Abramov, D. M., & Correia, R. F. (2021). Low COVID-19 vaccine hesitancy in Brazil. *Vaccine*, *39*(42), 6262-6268. https://doi.org/10.1016/j.vaccine.2021.09.013

Murray, C. J., Lopez, A. D., Chin, B., Feehan, D., & Hill, K. H. (2006). Estimation of potential global pandemic influenza mortality on the basis of vital registry data from the 1918–20 pandemic: a quantitative analysis. *The Lancet, 368*(9554), 2211-2218. https://doi.org/10.1016/S0140-6736(06)69895-4

Nair, S. C., Gasmelseed, H. I., Khan, A. A., Khafagy, I. N., Sreedharan, J., Saleem, A. A., Abdrhman, H. I., Alhosani, A. H., Siddiqua, A. R., Ahmed, A. R., Shubbar, A. I., Aleissaee, A. R., Alanqar, A. W., Hamadeh, A. M., Safdani, F. A., Habbal, F. W., Choker, W. B., Bashir, K. M., Alblooshi, M. A., Farajallah, M. M., Alzaabi, M. N., ShiL, R. S., Alshehhi, S. S., & Douleh, W. F. (2021). Assessment of mortality from COVID-19 in a multicultural multi-ethnic patient population. *BMC Infectious Diseases, 21*(1), 1-9. https://doi.org/10.1186/s12879-021-06762-9

National Health Council (2021). https://conselho.saude.gov.br/vacinometro

Ogedegbe, G., Ravenell, J., Adhikari, S., Butler, M., Cook, T., Francois, F., & Horwitz, L. I. (2020). Assessment of racial/ethnic disparities in hospitalization and mortality in patients with COVID-19 in New York City. *Journal of the American Medical Association, Network Open, 3*(12), https://doi.org/:10.1001/jamanetworkopen.2020.26881

Organization for Economic Co-operation and Development [OECD] (2020). OECD Economic Surveys. Brazil. December 2020. Overview. https://www.oecd.org/economy/surveys/Brazil-2020-OECD-economic-survey-overview.pdf

Organization for Economic Co-operation and Development [OECD] (2021). OECDiLibrary. Education at a Glance 2021: OECD Indicators. Brazil. https://www.oecd-ilibrary.org/sites/d8d547f5-en/index.html?itemId=/content/component/d8d547f5-en#section-d12020e24001

Ortega, F., & Orsini, M. (2020). Governing COVID-19 without government in Brazil: Ignorance, neoliberal authoritarianism, and the collapse of public health leadership. *Global public health*, *15*(9), 1257-1277. https://doi.org/10.1080/17441692.2020.1795223

Paim, J., Travassos, C., Almeida, C., Bahia, L., & Macinko, J. (2011). The Brazilian health system: history, advances, and challenges. *The Lancet*, *377*(9779), 1778-

1797. https://doi.org/10.1016/S0140-6736(11)60054-8

Ren, X. S., Amick, B. C., & Williams, D. R. (1999). Racial/ethnic disparities in health: the interplay between discrimination and socioeconomic status. *Ethnicity & Disease*, *9*(2), 151-165.

Rattay, P., Michalski, N., Domanska, O. M., Kaltwasser, A., De Bock, F., Wieler, L. H., & Jordan, S. (2021). Differences in risk perception, knowledge and protective behaviour regarding COVID-19 by education level among women and men in Germany. Results from the COVID-19 Snapshot Monitoring (COSMO) study. *Plos One*, *16*(5), e0251694. https://doi.org/10.1371/journal.pone.0251694

Ribeiro, F., & Leist, A. (2020). Who is going to pay the price of COVID-19? Reflections about an unequal Brazil. *International Journal for Equity in Health*, *19*(1), 1–3. https://doi.org/10.1186/s12939-020-01207-2

Ribeiro, K. B., Ribeiro, A. F., Veras, M. A. D. S. M., & de Castro, M. C. (2021). Social inequalities and COVID-19 mortality in the city of São Paulo, Brazil. *International Journal of Epidemiology*, *50*(3), 732-742. https://doi.org/10.1093/ije/dyab022

Ribeiro, V. M., & Vóvio, C. L. (2017). Desigualdade escolar e vulnerabilidade social no território. Educar em Revista, 71-87. https://doi.org/10.1590/0104-4060.51372

Rocha, R., Atun, A., Massuda, A., Rache, B., Spinola, S., Nunes, L., Lago, M., & Castro, M. C. (2021). Effect of socioeconomic inequalities and vulnerabilities on health-system preparedness and response to COVID-19 in Brazil: A comprehensive analysis. *The Lancet Global Health*, *9*(6), e782–e792 https://doi.org/10.1016/S2214-109X(21)00081-4

Secretaria de Vigilância em Saúde [Health Surveillance Secretariat] (2022). Departamento de Análise Epidemiológica e Vigilância de Doenças Não Transmissíveis [Department of Epidemiological Analysis and Surveillance of Noncommunicable Diseases]. Portal do Governo Brasileiro. https://svs.aids.gov.br/daent/centrais-de-conteudos/paineis-de-monitoramento/mortalidade/cid10/

Shadmi, E., Chen, Y., Dourado, I., Faran-Perach, I., Furler, J., Hangoma, P., Hanvoravongchai1, P., Obando, C., Petrosyan, V., Rao, K. D., Ruano, A. L., Shi, L., Souza, L. E., Spitzer-Shohat, S., Sturgiss, E., Suphanchaimat, R., Uribe, M. V., & Willems, S. (2020). Health equity and COVID-19: global perspectives. *International Journal for Equity in Health, 19*(1), 1-16. https://doi.org/10.1186/s12939-020-01218-z

Shkolnikov, V. M., Andreev, E. M., Jdanov, D. A., Jasilionis, D., Kravdal, Ø., Vågerö, D., & Valkonen, T. (2012). Increasing absolute mortality disparities by education in Finland, Norway and Sweden, 1971-2000. *Journal of Epidemiology and Community Health, 66*(4), 372–378. https://doi.org/10.1136/jech.2009.104786

Sistema IBGE de Recuperação Automática [IBGE Automatic Recovery System] (n.d.). Pesquisa Nacional por Amostra de Domicílios Contínua Trimestral. Tabela 6403 - População, por cor ou raça (Vide Notas). Retrieved from https://sidra.ibge.gov.br/tabela/6403

Stojanovic, J., Boucher, V. G., Gagne, M., Gupta, S., Joyal-Desmarais, K., Paduano, S.,

& Lavoie, K. L. (2021). Global trends and correlates of COVID-19 vaccination hesitancy: Findings from the iCARE study. *Vaccines, 9*(6), 661. https://doi.org/10.3390/vaccines9060661

Strand, B. H., Grøholt, E. K., Steingrímsdóttir, Ó. A., Blakely, T., Graff-Iversen, S., & Næss, Ø. (2010). Educational inequalities in mortality over four decades in Norway: prospective study of middle-aged men and women followed for cause specific mortality, 1960-2000. *British Medical Journal*, *340*, 1-10. https://doi.org/10.1136/bmj.c654

Talic, S., Shah, S., Wild, H., Gasevic, D., Maharaj, A., Ademi, Z., & Ilic, D. (2021). Effectiveness of public health measures in reducing the incidence of COVID-19, SARS-CoV-2 transmission, and COVID-19 mortality: Systematic review and meta-analysis. *British Medical Association, 375*. https://doi.org/10.1136/bmj-2021-068302

U.S. Bureau of Labor Statistics (n.d.). Labor Force Statistics from the Current Population Survey. https://www.bls.gov/cps/

World Health Organization [WHO] (2020-a). Archived: WHO Timeline - COVID-19. Retrieved from https://www.who.int/news/item/27-04-2020-who-timeline---COVID-19

World Health Organization (n.d.-a). Advice for the public: Coronavirus disease (COVID-19). Retrieved from https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public

Manuscript 2: Association of Education and COVID-19 Mortality in Brazil

Marcelina Machado

Walden University

Outlet for Manuscript

The manuscript on education and COVID-19 mortality in Brazil will be submitted to the American Journal of Public Health, https://ajph.aphapublications.org/. The journal requires the original research results reported in up to 3500 words in the text, a structured abstract, up to four tables and figures combined, and no more than 35 free references. The journal also requires AMA formatting with 1.5 or double spacing and a font size of 12. The American Journal of Public Health is aligned with the mission to advance public health research, policy, practice, and education. The present manuscript can provide valuable insights for public health practitioners and policymakers, allowing them to evaluate the association between public demographics, education, and COVID-19 mortality across Brazil.

Abstract

Objective. To investigate the association between education and COVID-19 mortality in Brazil, while accounting for variables related to race, age, sex, and immunization.

Methodology. In this retrospective study, school enrollment data were retrieved from the Ministry of Education; COVID-19 mortality data retrieved from the Health Surveillance Secretariat population demographics data including race, age, and sex will be retrieved from the Ministry of Health; and immunization data retrieved from the Oswaldo Cruz Foundation. All data were analyzed at region, state, and capital-levels using multiple linear regression on IBM SPSS Statistics version 27.

Results. The findings suggest a statistically significant association between school enrollment and COVID-19 mortality rate at the state level ($p<0.001$), but not a significant association at the capital level ($p=0.102$). At the state level, for each unit increase in average income, we can predict a decrease of 9.5 in COVID-19 death per 100,000 people. Variables related to race, age, sex, and immunization were used as control variables in all the regression analyses. Results show that the model is invalid to predict COVID-19 mortality from school enrollment at the regional level, as violations of the assumptions of the model were noted.

Conclusion. To better explore the influences of education on COVID-19 mortality, further research is recommended with both individual and population-based approaches. Since COVID-19 is an ongoing pandemic, it is likely that more accurate results will only be obtained with post-pandemic data.

***Keywords*:** COVID-19 mortality, education, school enrollment, Brazil.

Introduction

In Brazil, access to education is a right based on the constitution of 1934; however, the law in which this right was under presented some gaps (Lei de Diretrizes e Bases da Educação [National Education Guidelines and Framework Law], 1961). In this law, the obligatoriness of primary education was exempted for children with severe diseases or living in poverty in areas with insufficient schools and closed enrollments (Governo Brasileiro [Brazilian Government], 1961). This exemption worsened education inequality, even opening opportunities for child labor, as individuals living in poverty were deprived of education (Curry, 2008; Governo Brasileiro [Brazilian Government], 1961). The law of 1934 was revoked in 1966, and although primary education access was extended based on the constitution of 1967, educational resources were relocated, thus compromising the quality of public education (Curry, 2008; Governo Brasileiro, 1966). Eventually, because of many public social movements throughout the decades, the constitution of 1988 placed education as the most important social right for the citizens' formation (Curry, 2008).

According to the Brazilian Institution of Geography and Statistics (IBGE, 2019), the level of education is commonly measured by the country's illiteracy rate. Among Brazilian individuals of 15 years or older, illiteracy is estimated at 6.6% (11 million illiterates), being higher among black males in the Northeast (IBGE, 2019). Besides illiteracy, school enrollments can be used to measure a specific area's education level, considering that areas with a higher number of school enrollments have a greater socioeconomic prestige associated with easier access to education. In contrast, vulnerable regions have lower rates of school enrollment, often insufficient to meet the demand, meaning that vulnerable individuals have less access to education (Bartholo et al., 2020; Ribeiro et al., 2017). The lack of access to education can lead to a reduction in employment opportunities, in income, and consequently the individual's access to healthcare, leading to an increase in the likelihood of disease incidence and mortality. The issue that prompted me to search the literature is that the Brazilian population experienced a significantly high COVID-19 mortality rate, which could be associated with the population's level of education. Based on this issue, this study aims to evaluate whether there is an association between the independent variable education and the dependent variable COVID-19 mortality while controlling for variables related to race, age, sex, and immunization.

The significance of this study lies in the fact that its findings could help political leaders to identify the areas where the association between education and COVID-19 mortality is higher, encouraging them to improve people's access to better education for more employment opportunities, increase income and consequently more access to healthcare. The concepts of the fundamental causes theory will be used herein as it relates to the COVID-19 mortality rate and the effect it has on the public's ability to reach essential resources, in this case, education, that could produce health benefits to prevent disease (Link et al., 1995). This study aims to investigate whether this association exists in Brazil while controlling for the population's demographics and immunization.

Researchers have indicated that access to education is unequal and poorly encouraged in Brazil (Curry, 2008; Governo Brasileiro [Brazilian Government], 1961).

According to Barreto et al. (2021) and Rattay et al. (2021), lack of knowledge or access to information might indicate a population's vulnerability and low-risk perception, increasing disease incidence and mortality. Although there are studies on the association between education and disease mortality, there is still limited evidence regarding education and its impact on COVID-19 mortality in Latin-American countries, including Brazil. Thus, this gap needs to be yet studied.

Research Question & Hypotheses

This quantitative study addressed the following research question:

Research Question for Manuscript 2 (RQM2): What is the association between school enrollment and COVID-19 mortality rate at different geographic levels across Brazil after controlling for race, age, sex, and immunization rate?

Based on the research question, the following hypotheses were formulated:

Null Hypothesis (H0): There is no statistically significant association between school enrollment and COVID-19 mortality rate at different geographic levels across Brazil after controlling for race, age, sex, and immunization rate.

Alternative Hypothesis (H1): There is a statistically significant association between school enrollment and COVID-19 mortality rate at different geographic levels across Brazil after controlling for race, age, sex, and immunization rate.

The rationale for this research is to evaluate if Brazil's education inequality can adversely impact COVID-19 mortality; additionally, there were few studies that explored this association. Thus, to address the research question, this research design includes a retrospective cohort design (Aschengrau et al., 2020) with mortality rates examined between the period of March 2020 to December 2021. This quantitative analysis should help determine if education and population demographic characteristics correlate to COVID-19 mortality in Brazil.

Methodology

Study Population

The target population for this study include Brazilian individuals with different demographic and geographic characteristics, including race (White, African origin, and Mixed), age (0-14, 15-49, and 50-80 years old), sex (male and female), and immunization. Public dataset repositories were used for this research design to obtain data on the COVID-19 mortality population's socioeconomic, demographic, and geographic characteristics. Since this is a population-based study, aggregated data from different geographic levels, including region, state, and capitals were used for all dependent, independent, and control variables. For this study, convenience sampling was used (Warner, 2013). Based on the data retrieved for each geographic level during the first two years of the pandemic, the convenience sample size for regions is 10, for states is 54, and for capitals is 54. Additionally, information on individual-based data was still used to show death distribution throughout the first years of the pandemic according to race, age group, sex, and geographic region, but were used as information only and not to answer the study research question.

Variables

The variables of this study include the following.

Independent variable (IV):

- School enrollment rate": Numbers in percentage, a higher percentage, indicate higher school enrollment. This rate was calculated for different geographic areas by dividing the number of school enrollment in each area by the number of people in the population, then multiplied by 100.

Dependent variable (DV):

- COVID-19 mortality rate: The mortality rate was calculated for different geographic areas by dividing the number of deaths from COVID-19 in each area by the number of people in the population, then multiplied by 100,000.

Control variables (CV1 – CV8):

- White population: Numbers in percentage. The percentage of the White population was calculated for different geographic areas by dividing the number of White people in each area by the number of people in the population, then multiplied by 100.
- African origin population: Numbers in percentage. The percentage of the African-origin population was calculated for different geographic areas by dividing the number of African-origin people in each area by the number of people in the population, then multiplied by 100.
- Mixed-race population: Numbers in percentage. The percentage of the Mixed-race population was calculated for different geographic areas by dividing the number of Mixed-race people in each area by the number of people in the population, then multiplied by 100.

- Population between 0-14 years old: Numbers in percentage. The percentage of the 0-14 years old population was calculated for different geographic areas by dividing the number of 0-14 years old people in each area by the number of people in the population, then multiplied by 100.
- Population between 15-49 years old: Numbers in percentage. The percentage of the 15-49 years old population was calculated for different geographic areas by dividing the number of 15-49 years old people in each area by the number of people in the population, then multiplied by 100.
- Population between 50-80 years old: Numbers in percentage. The percentage of the 50-80 years old population was calculated for different geographic areas by dividing the number of 50-80 years old people in each area by the number of people in the population, then multiplied by 100.
- Male population: Numbers in percentage. The percentage of the male population was calculated for different geographic areas by dividing the number of male people in each area by the number of people in the population, then multiplied by 100.
- Immunization rate: Number in percentage. A higher percentage indicates a higher immunization. Actual percentage values from different geographic levels (regional, state, and capital levels) were used as found in the data source There were no codes from missing data.

Sources of Data

Data on school enrolment per geographic location from 2020 and 2021 were accessed on September 18, 2022, from the publicly available dataset repository below:

- Instituto Nacional de Estudos e Pesquisas (INEP) - Ministério da Educação [National Institute of Studies and Research - Ministry of Education].

Data on COVID-19 mortality per geographic location from March 2020 through December 2021 were accessed on September 18, 2022, from the publicly dataset repository below:

- Secretaria de Vigilância em Saúde [Health Surveillance Secretariat], created by the Ministério da Saúde [Ministry of Health].

Data on the population's demographics, including race, age, and sex per geographic location from 2020 and 2021, were accessed on September 18, 2022, from the publicly available dataset repositories below:

- Instituto Brasileiro de Geografia e Estatistica (IBGE) [Brazilian Institute of Geography and Statistics].
- Tabnet – DATASUS - Ministério da Saúde [Ministry of Health].

Data on the population's immunization (or percentage of population vaccinated with the second dose) per geographic location from March 2021 through September 2022 were accessed on September 18, 2022, from the publicly available dataset repository below:

- Fundação Oswaldo Cruz (Fiocruz) [Oswaldo Cruz Foundation].

Access to these public dataset repositories were needed to gather specific data points, including the number of deaths by COVID-19 at region, state and capital-levels and the population's socioeconomic and demographic information. To minimize potential issues regarding the use of secondary data, datasets were screened for errors, outliers, inconsistencies, and missing values prior to analysis. Identified errors in the datasets were be reported in the discussion section.

Validity and Reliability of Data

Internal validity, which is "the degree to which the results of the study can be used to make causal inferences," is often weak in non-experimental studies (Warner, 2013). To increase internal validity by limiting the influence of confounding, a series of control variables related to race, age, sex, and immunization rate, were included in the analysis. Additionally, regression analyses were performed separately at each geographic level (region, state, and capital) to avoid violation of internal validity. On the other hand, external validity, the "degree to which the study's results can be generalized" (Warner, 2013), could be considered high in this study as the data to be used represent a large and heterogeneous population (Oliveira, 2020).

The reliability of secondary data depends on the sources from which the data were obtained. Datasets from government sources are often reputable; however, data screening is still recommended to identify errors and inconsistencies (Warner, 2013). The datasets used in this study were retrieved from Brazilian governmental websites listed in the previous section. Also, data were reviewed before analysis to ensure the accuracy of data transcription.

Design and Analysis

A multiple linear regression (MLR) analysis was performed using IBM SPSS Statistics version 27 to determine the association between school enrollment and COVID-19 mortality rate at different geographic levels across Brazil after controlling for variables related to race, age, sex, and immunization rate. Multiple linear regression is a statistical technique often used to evaluate the relationship of one continuous dependent variable and multiple continuous or categorical independent variables (Warner, 2013).

Although aggregated data (population-based data) were accessed and used in the primary statistical analysis (MLR) to answer the research question previously elaborated, individual-based data on COVID-19 mortality and population demographics, including race, age, sex, and region, were used to evaluate death distribution in the country during the first two years of the pandemic. Additionally, descriptive statistical analyses were performed for the COVID-19 mortality rate and school enrollment variables. Lastly, prior to the performance of the primary analysis, assumptions for MLR were tested. Table 9 and Figure 15 summarize of the proposed analytical test plan.

Table 9

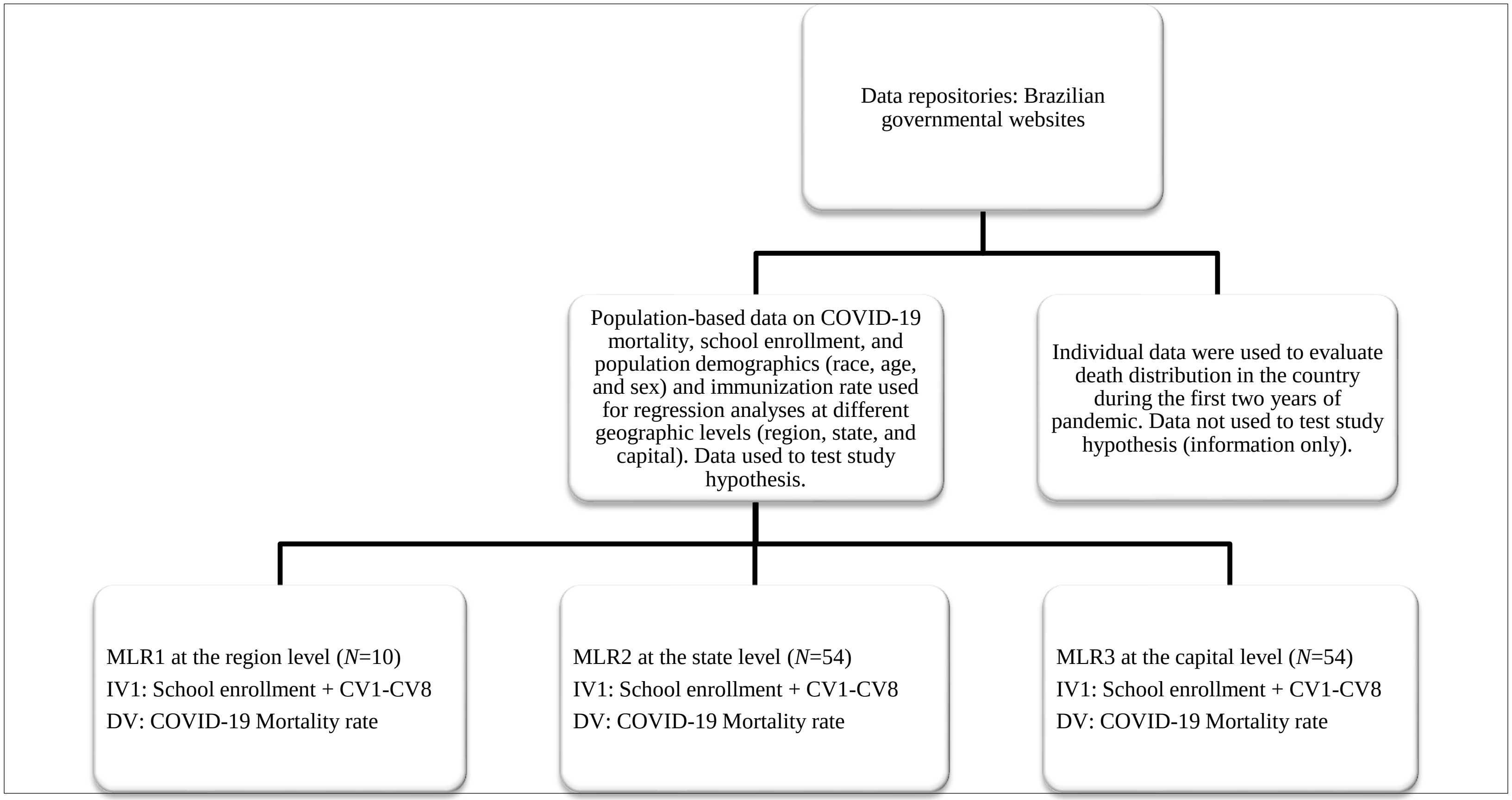
Proposed Analytical Test Plan and Justification

Variable	Type of variable	Multiple Linear Regression (MLR)	Death Distribution Evaluation
COVID-19 Mortality rate [1]	Dependent Variable **DV** (Scale)	Using a population-based dataset and including all variables listed, multiple linear regression analyses were performed to determine the association between education and COVID-19 mortality while controlling for the population's demographic characteristics and immunization. MLR is an appropriate test considering that the dependent variable (disease mortality) and the independent variable (school enrollment) are continuous variables, and the control variables are scale (Warner, 2013). Assumptions tested: Normality, linearity, homoscedasticity, and multicollinearity. In these analyses, all variables listed were used.	An individual-based dataset was created using only COVID-19 mortality and population demographics, including race, age, sex, and region, to evaluate death distribution in the country during the first two years of the pandemic. Data not used to test study hypothesis. Analyses were done for information only.
School enrollment [1]	Independent Variables **IV1** (Scale)		
Percentage of White population	Control Variable **CV1** (Scale)		
Percentage of African origin population	Control Variable **CV2** (Scale)		
Percentage of Mixed-race population	Control Variable **CV3** (Scale)		
Percentage of Population between 0-14 Years Old	Control Variable **CV4** (Scale)		
Percentage of Population between 15-49 Years Old	Control Variable **CV5** (Scale)		
Percentage of Population between 50-80 Years Old	Control Variable **CV6** (Scale)		
Percentage of Male Population	Control Variable **CV7** (Scale)		
Percentage of Immunized Population	Control Variable **CV8** (Scale)		

Note. [1] Descriptive statistical analysis were performed on the COVID-19 mortality rate and school enrollment variables.

Figure 15

Scheme of Data Flow and Regression Analyses for Manuscript 2



Note. Regression analyses were performed at each geographic level (region, state, and capital) to avoid internal validity.

Results

The study was conducted by retrieving population-based data on COVID-19 mortality, school enrollment, and population demographics, including race, age, and sex, as well as immunization rate from Brazilian governmental websites and analyzed using multiple linear regression model at different geographic levels (see Figure 15) on IBM SPSS software. The population-based data were used to create a dataset containing all information for the regression analyses. The dataset was further converted into an individual-based dataset only on COVID-19 mortality and population demographics to investigate the death distribution during the first 2 years of the pandemic.

Most data needed for the analyses were retrieved without issues, except for the immunization data, which were not available at the capital-level; therefore, this variable was not included in the regression analysis for this level. Additionally, as vaccination campaigns started on January 2021 in Brazil, the data retrieved for this study are from 2021 and 2022. It is also important to note that the regression at the nation-level was not performed due to the small sample size (N=2).

Descriptive Statistics for COVID-19 Mortality

After running a descriptive statistics analysis on the variable COVID-19 mortality, it was observed that among the 27 states during two years of the pandemic (N= 54), the COVID-19 mortality rate ranged from 0.063% to 0.284% at the state level. The three measures of central tendency for this variable were mean equals 0.14% (M = 0.14), median equals 0.13% (Median = 0.13), and mode equals 0.063% (Mode = 0.063) with a standard deviation equal 0.058 (SD = 0.058). The skewness and kurtosis of this distribution (0.816 and -0.284, respectively) suggest that the distribution deviates from normality and is positively skewed (Table 10 and Figure 16). In Figure 16, the variability or measure of spread distribution of the COVID-19 mortality rate was 0.18% to 0.10%; outliers were not identified.

In a second descriptive statistics analysis, it was observed that among the 27 capitals during the two years of the pandemic (*N*= 54), the COVID-19 mortality rate ranged from 0.107% to 0.519% at the capital level. The three measures of central tendency for this variable were mean equals 0.25% (*M* = 0.25), median equals 0.23% (Median = 0.23), and mode equals 0.107% (Mode = 0.107) with standard deviation equals 0.094 (*SD* = 0.094). The skewness and kurtosis of this distribution (0.325 and -0.886, respectively) suggest that the distribution deviates from normality and is slightly positively skewed (Table 10 and Figure 17). In Figure 17, the variability or measure of spread distribution of the COVID-19 mortality rate was 0.28% to 0.18%; outliers were identified.

Table 10

Statistics of COVID-19 Mortality at State and Capital-levels

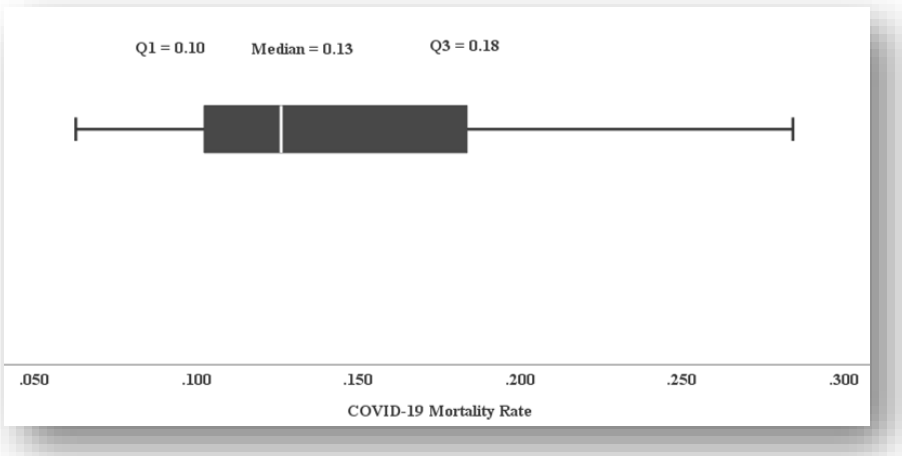
		State-level	Capital-level
N	Valid	54	54
	Missing	0	0
Mean		0.14169	0.24528
Median		0.12622	0.23056
Mode		0.063[1]	0.107 [1]
Std. Deviation		0.058010	0.093895
Variance		0.003	0.009
Skewness		0.816	0.989
Std. Error of Skewness		0.325	0.325
Kurtosis		-0.284	0.886
Std. Error of Kurtosis		0.639	0.639
Range		0.221	0.411
Minimum		0.063	0.107
Maximum		0.284	0.519
Percentiles	25	0.10052	0.18157
	50	0.12622	0.23056
	75	0.18464	0.28142

Note. Statistics were performed at state and capital levels for the first two years of the pandemic. For that, the *Frequencies* function in SPSS was used to observe further Percentile values, Central Tendency, Dispersion, and sample Distribution.

[1] Multiple mode exist. The smallest value is shown.

Figure 16

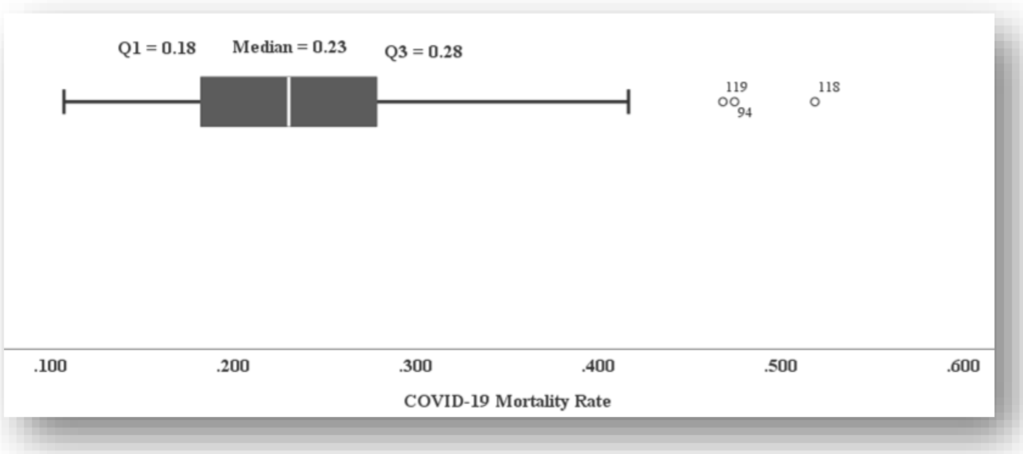
Boxplot Graph Representing a Positively Skewed Distribution of COVID-19 Mortality at State-level



Note. No outliers were identified.

Figure 17

Boxplot Graph Representing a Positively Skewed Distribution of COVID-19 Mortality at Capital-level



Note. Outliers were identified.

Descriptive Statistics for School Enrollment

After running a descriptive statistics analysis, it was observed that among the 27 states during the two years of the pandemic (*N*= 54), the school enrollment rate ranged from 38.7% to 58.4% at the state level. The three measures of central tendency for this variable are mean equals 47.1% (*M* = 47.1), median equals 46.7% (Median = 46.7), and mode equals 46.2% (Mode = 46.2) with standard deviation equals 4.9 (*SD* = 4.9). The skewness and kurtosis of this distribution (0.267 and -0.642, respectively) suggest that the distribution deviates from normality and is positively skewed (Table 11 and Figure 18). In Figure 18, the variability or measure of spread distribution of the school enrollment was 50.4% to 43.2%; outliers were not identified.

In a second descriptive statistics analysis, it was observed that among the 27 capitals during the two years of the pandemic (*N*=54), the school enrollment ranged from 34.1% to 50.5% at the capital level. The three measures of central tendency for this variable are mean equals 42.6% (*M* = 42.6), median equals 42.7% (Median = 42.7), and mode equals 35.8% (Mode = 35.8) with standard deviation equals 4.3 (*SD* = 4.3). The skewness and kurtosis of this distribution (-0.092 and -1.043, respectively) suggest a non-normal distribution and slightly negatively skewed (Table 11 and Figure 19). In Figure 19, the variability or measure of spread distribution of the school enrollment was 46.2% to 38.6%; outliers were not identified. Note that for both descriptive analyses, the high value of variance (the square root of the standard deviation) may indicate highly scattered data around the mean.

Table 11

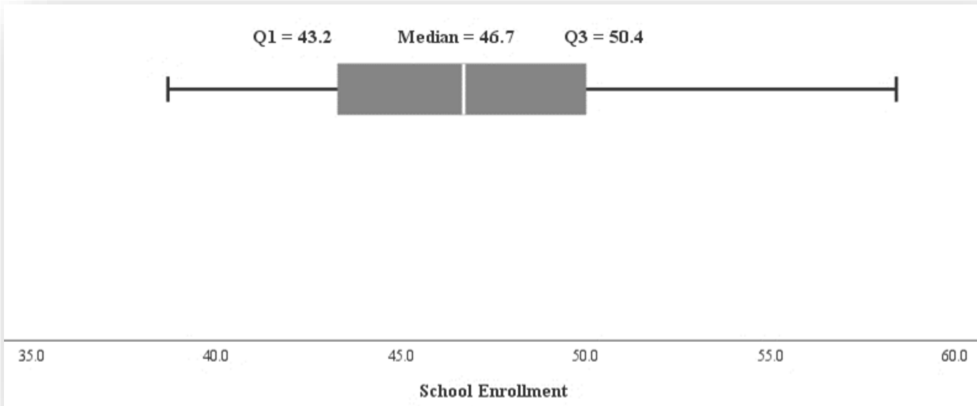
Statistics of School Enrollment at State and Capital-levels

		State-level	Capital-level
N	Valid	54	54
	Missing	0	0
Mean		47.076	42.604
Median		46.700	42.700
Mode		46.2	35.8[1]
Std. Deviation		4.9636	4.3226
Variance [2]		24.637	18.685
Skewness		0.267	-0.092
Std. Error of Skewness		0.325	0.325
Kurtosis		-0.642	-1.043
Std. Error of Kurtosis		0.639	0.639
Range		19.7	16.4
Minimum		38.7	34.1
Maximum		58.4	50.5
Percentiles	25	43.200	38.575
	50	46.700	42.700
	75	50.425	46.225

Note. Statistics were performed at state and capital levels for the first two years of the pandemic. For that, the *Frequencies* function in SPSS was used to observe further Percentile values, Central Tendency, Dispersion, and sample Distribution. [1] Multiple mode exist. The smallest value is shown. [2] Note that for both descriptive analyses, the high value of variance (the square root of the standard deviation) may indicate highly scattered data around the mean.

Figure 18

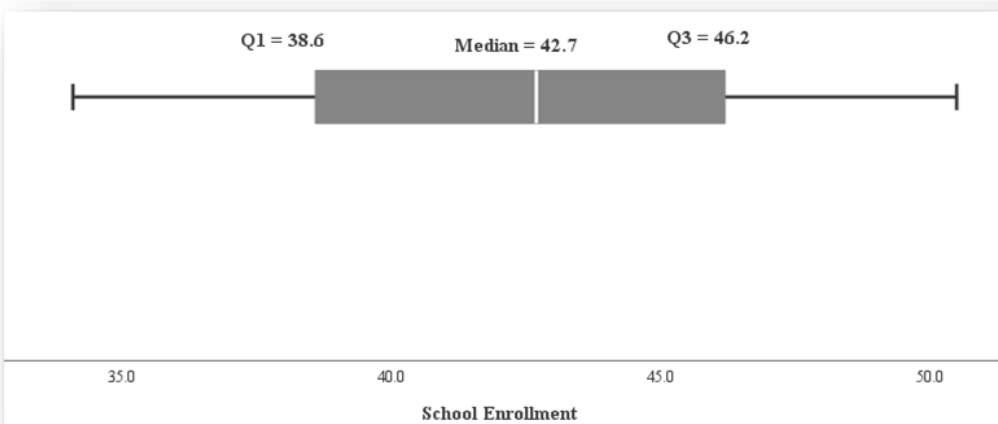
Boxplot Graph Representing a Positively Skewed Distribution of School Enrollment at State-level



Note. No outliers were identified.

Figure 19

Boxplot Graph Representing a Symmetrical Distribution of School Enrollment at Capital-level



Note. No outliers were identified.

Assumptions of Multiple linear Regression

The assumptions tested for the MLR analyses were normality, linearity, homoscedasticity, and multicollinearity. Normality was tested by observing the Normal P-P Plot, linearity and homoscedasticity were tested by observing the scatterplot of residuals, and multicollinearity was tested by observing the variance inflation factor (VIF) obtained from the coefficient table in the SPSS output. Additionally, the independence of residuals was tested by observing the Durbin-Watson value. Outliners were tested by looking for points that were far from the regression line.

Table 12 summarizes the results for the assumptions of the multiple linear regression analyses. Note that the passing criterion set for normality was that residuals should be normally distributed. The passing criterion for linearity was that residuals should have a straight-line relationship with predicted DV scores. The passing criterion for homoscedasticity was that VIF values should be less than 10 (VIF<10). The passing criterion for the independence of residuals is that plotted standardized residuals are randomly scattered. The passing criterion for multicollinearity was that Durbin-Watson values should be between 1.5 – 2.5 (approximate values were acceptable). Overall, the results show that except for linear regression at the region-level, none of the assumptions were violated. Although not relevant, significant results of the analysis at the regional level were still reported in this study.

Table 12

Results of Assumptions of Multiple linear Regression for Manuscript 2

Level	Regression/IV and DV	Normality (passing criterion is that residuals should be normal distributed) **(Pass or fail criterion)**	Linearity (passing criterion is that residuals should have a straight-line relationship with predicted DV scores) **(Pass or fail criterion)**	Homoscedasticity (passing criterion is that plotted standardized residuals are randomly scattered) **(Pass or fail criterion)**	Multicollinearity (passing criterion is VIF <10) **(Pass or fail criterion)**	Independence of Residuals (passing criterion is Durbin-Watson values between 1.5-2.5[1]) **(Pass or fail criterion)**	Outliers (passing criterion is absence of outliers) **(Pass or fail criterion)**
Region	Regression 1 (School Enrollment and COVID-19 mortality)	Fail	Fail	Fail	Fail	Fail	Pass
State	Regression 2 (School Enrollment and COVID-19 mortality)	Pass	Pass	Pass	Pass[2]	Pass	Pass
Capital	Regression 3 (School Enrollment and COVID-19 mortality)	Pass	Pass	Pass	Pass[2]	Pass	Pass

Note. [1]Approximate values are acceptable. [2] Failed for 2 non-significant variables.

Results of Multiple Linear Regression

To approach the association between education and COVID-19 mortality in Brazilian regions, a multiple linear regression analysis was conducted to evaluate the prediction of COVID-19 mortality rate from school enrollment, while controlling for variables related to the population's race, age, sex, and immunization rate (see list of variables included in the Design and Analysis section). Preliminary analyses were conducted to assess the assumptions of normality, linearity, homoscedasticity, multicollinearity, independence of residuals, and outliers. Violations were noted, indicating that the model is invalid in predicting COVID-19 mortality at the regional level (see Table 12). Therefore, regression results for this level were not reported as they cannot be used to interpret the data correctly.

The same approach was taken to evaluate the association between education and COVID-19 mortality in Brazilian states. A multiple linear regression analysis was conducted to evaluate the prediction of COVID-19 mortality rate from school enrollment, while controlling for variable related to the population's race, age, sex, and immunization rate (see control variables listed in Table 13). Preliminary analyses were conducted to assess the assumptions of normality, linearity, homoscedasticity, multicollinearity, independence of residuals, and outliers. No violations were noted (see Table 12). The multiple linear regression analysis results revealed that at the state level, a significant regression equation was found $F(8, 45)=6.574$, $p<0.001$), with an R^2 of 0.539. Table 13 summarizes the results of the regression performed at the state level showing that for each percentage point increase in school enrollment, we can predict a decrease of 9.5 in

COVID-19 death per 100,000 people. School enrollment is a significant predictor of COVID-19 mortality (p<0.001) while controlling for the population's race, age, sex, and immunization rate.

Lastly, to approach the association between education and COVID-19 mortality in Brazilian capitals, a multiple linear regression analysis was conducted to evaluate the prediction of COVID-19 mortality from school enrollment while controlling for variables related to the population's race, age, and sex (see control variables listed in Table 14). Preliminary analyses were conducted to assess the assumptions of normality, linearity, homoscedasticity, multicollinearity, independence of residuals, and outliers. No violations were noted (see Table 12). The multiple linear regression analysis results revealed that at the capital level, a not significant regression equation was found F(7, 46)=1.366, p=0.243), with an R^2 of 0.172. Table 14 summarizes the regression results performed at the capital level showing that school enrollment is not a significant predictor of COVID-19 mortality (p=0.102) while controlling for the population's race, age, and sex.

Table 13

SPSS Output: Regression Analysis of School Enrollment and COVID-19 Mortality at State-level

Coefficients									
		Unstandardized Coeff.		Standardized Coeff.	t	*p*	95% CI for B	
Model		*B*	*SE*	β			*LL*	*UL*
	(Constant)	503.143	785423		0.641	0.525	-1079	2085
	% School enrollment[1]	-9.501	2.436	-0.813	-3.901	<0.001	-14.406	-4.595
	% Male population[2]	7.095	9.645	0.121	0.736	0.466	-12.330	26.521
	% Population between 0-14 years[2]	18.774	6.074	0.959	3.086	0.003	6.510	30.979
1	% Population between 15-49 years[2]	-3.548	5.441	-0.120	-0.652	0.518	-14.507	7.410
	% White population[2]	-779	384	-2.397	-2.030	0.048	-1552	-6.136
	% African origin population[2]	-602	459	-0.298	-1.313	0.196	-1526	321
	% Mixed-race population[2]	-807	386	-2.219	-2.088	0.042	-1585	-28.715
	% Immunized population[2]	3.967	0.740	0.807	5.364	<0.001	2.478	5.457

Note. [1.] Predictor variable [2.] Control variables.

Regression model using the total number of deaths from COVID-19 per 100,000 people at the state-level in Brazil.

Table 14

SPSS Output: Regression Analysis of School Enrollment and COVID-19 Mortality at Capital-level

Coefficients								
		Unstandardized Coeff.		Standardized Coeff.			95% CI for B	
Model		*B*	*SE*	β	t	*p*	*LL*	*UL*
	(Constant)	950	1931		0.492	0.625	-2936	4837
	% School enrollment[1]	-6.732	4.034	-0.310	-1.669	0.102	-14.851	1.388
	% Male population[2]	25.533	14.491	0.391	1.762	0.085	-3.636	54.702
	% Population between 0-14 years[2]	-3.096	9.747	-0.090	-0.318	0.752	-22.715	16.524
1	% Population between 15-49 years[2]	-18.250	8.772	-0.527	-2.080	0.043	-35.908	-0.591
	% White population[2]	-699	1549	-1.282	-0.451	0.654	-3818	2419
	% African origin population[2]	-802	1593	-0.389	-0.503	0.617	-4009	2406
	% Mixed-race population[2]	-405	1553	-0.684	-0.261	0.795	-3531	2721

Note. [1.] Predictor variable [2.] Control variables.

Data on immunization was not available at capital-level; therefore, it was not included in this regression.

Regression model using the total number of deaths from COVID-19 per 100,000 people at the state-level in Brazil.

Distribution of COVID-19 Deaths

Between March 2020 and December 2021, the Secretaria de Vigilância em Saúde [Health Surveillance Secretariat], created by the Ministério da Saúde [Ministry of Health], reported 634,695 deaths by COVID-19, being 212,706 deaths in 2020 and 421 989 deaths in 2021. The number of deaths caused by COVID-19 during these first two years of the pandemic corresponds to 18.8% of all deaths in the country in the same period, and the COVID-19 mortality rate observed was 297.5 deaths per 100,000 people in all age groups. The distribution of the casualties according to age groups, sex, race, and Brazilian region are shown in Table 15, Table 16 and Figures 20-27. According to the distribution, in 2020, the number of deaths across the country was higher among individuals who were 80 years or older (1348.3 deaths/100,000), White (112.3 deaths/100,000), and male individuals (117.5 deaths/100,000). Similarly, in 2021 the number of deaths was still higher among individuals who were 80 years or older (1731 deaths/100,000), White (238.5 deaths/100,000), and male individuals (224.8 deaths/100,000). In 2020 death rate was higher in the Southeast region (109.9 deaths/100,000), conversely, in 2021, the death rate was higher in the South (244.6 deaths/100,000).

Note that although individual-based data was used to show death distribution throughout the first years of the pandemic according to race, age group, sex, and geographic region, this is a population-based study and the data in this section was not used to test the study hypothesis or answer the research question. This death distribution evaluation has been for information only.

Table 15

Distribution of COVID-19 Death in 2020 and 2021 according to Age Group

Demographic	[1]*N* (%) in 2020	[2] Death rate per 100,000	[1]*N* (%) in 2021	[2] Death rate per 100,000
Age				
Infant	278 (0.1)	-	347 (0.1)	-
0-4 years	168 (0.1)	1.1	212 (0.1)	1.4
5-9 years	101 (0.0)	0.7	121 (0.0)	0.8
10-14 years	128 (0.1)	0.9	171 (0.0)	1.2
15-19 years	347 (0.2)	2.2	545 (0.1)	3.5
20-29 years	1809 (0.9)	5.3	5517 (1.3)	16.2
30-39 years	5975 (2.8)	17.5	21155 (5.0)	61.7
40-49 years	13542 (6.4)	46.3	45392 (10.8)	152.0
50-59 years	26600 (12.5)	111.4	76859 (18.2)	317.1
60-69 years	47875 (22.5)	286.1	98362 (23.3)	568.7
79-79 years	55979 (26.3)	620.4	93347 (22.1)	991.3
80-older	59880 (28.2)	1348.3	79925 (18.9)	1731.0
Unknown	24 (0.0)	-	36 (0.0)	-
Total	212706 (100.0)	100.3	421989 (100.0)	197.6

Note. [1]*N* = number of deaths by COVID-19. Data on death were accessed from Secretaria de Vigilância em Saúde [Health Surveillance Secretariat] on 18Sep2022, using indicator ICD-10-CM code B34.2 for Coronavirus infection, unspecified. Data on total population for each age group were accessed from Ministério da Saúde [Ministry of Heath] on 18Sep2022. [2] Death rate per 100,000 was calculated by diving the number of deaths per the total population, times 100,000 (data for infant and unknown were excluded from the calculation).

Table 16

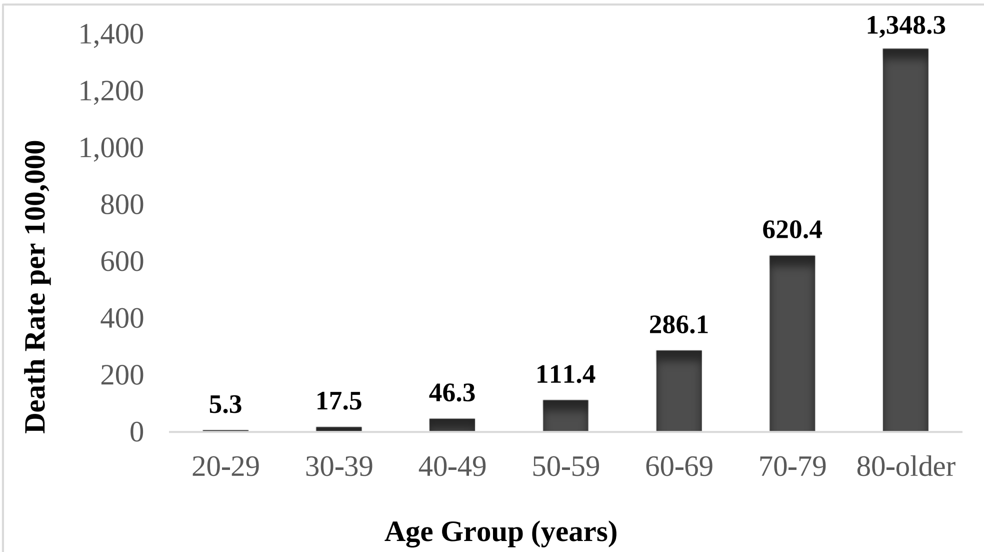
Distribution of COVID-19 Death in 2020 and 2021 according to Sex, Race, and Region

Demographic	[1]*N* (%) in 2020	[2] Death rate per 100,000	[1]*N* (%) in 2021	[2] Death rate per 100,000
Sex				
Male	121641 (57.2)	117.5	234352 (55.5)	224.8
Female	91051 (42.8)	84.1	187588 (44.5)	172.0
Unknown	14 (0.0)	-	49 (0.0)	-
Race				
White	103525 (48.7)	112.3	218852 (51.9)	238.5
African origin	18604 (8.7)	99.7	32149 (7.6)	154.1
Asian origin	1458 (0.7)	-	2591 (0.6)	-
Mixed-race	81572 (38.3)	82.7	139288 (33.0)	142.0
Brazilian Native	972 (0.5)	-	851 (0.2)	-
Unknown	6575 (3.1)	-	28258 (6.7)	-
Region				
Central-West	17381 (8.2)	105.3	39559 (9.4)	236.8
Northeast	54697 (25.7)	95.3	76020 (18.0)	131.8
North	19531 (9.2)	104.6	29861 (7.1)	157.9
Southeast	97798 (46.0)	109.9	202174 (47.9)	225.6
South	23299 (11.0)	77.2	74375 (17.6)	244.6
Total	212706 (100.0)	100.4	421989 (100.0)	197.8

Note. [1]*N* = number of deaths by COVID-19. Data on death were accessed from Secretaria de Vigilância em Saúde [Health Surveillance Secretariat] on 18Sep2022, using indicator ICD-10-CM code B34.2 for Coronavirus infection, unspecified. Data on total population for each sex and region groups were accessed from Ministério da Saúde [Ministry of Heath] on 18Sep2022. Data on total population for each race group were accessed from the Sistema IBGE de Recuperação Automática [IBGE Automatic Recovery System] on 18Sep2022. [2] Death rate per 100,000 was calculated by diving the number of deaths per the total population, times 100,000 (data for unknown, Asian, and Brazilian native population were excluded from the calculation).

Figure 20

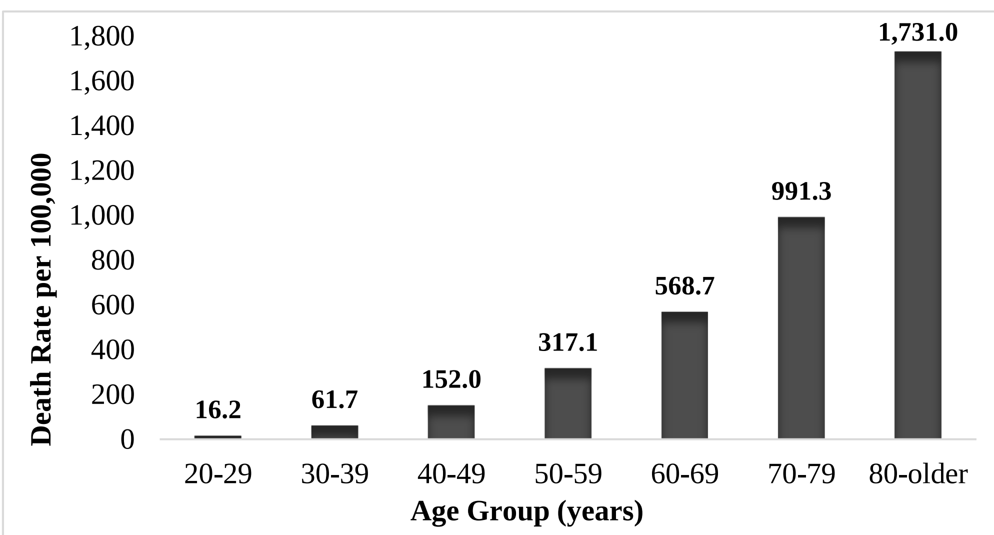
Bar Chart of COVID-19 Death Rate per 100,000 in 2020 by Age Group in Brazil



Note. Only rates >5 are displayed.

Figure 21

Bar Chart of COVID-19 Death Rate per 100,000 in 2021 by Age Group in Brazil



Note. Only rates > 5 are displayed.

Figure 22

Pie Chart of COVID-19 Death Rate pr 100,000 in 2020 by Sex in Brazil

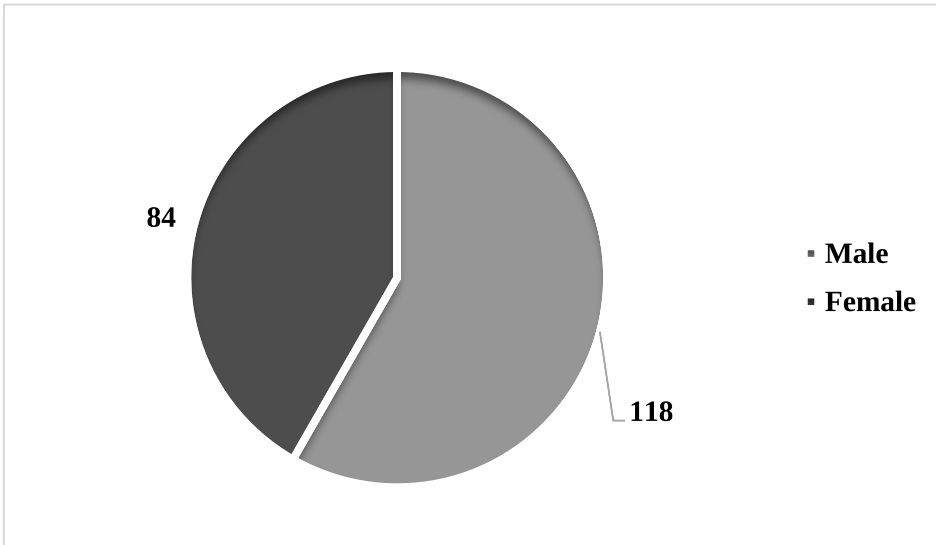


Figure 23

Pie Chart of COVID-19 Death Rate per 100,000 in 2021 by Sex in Brazil

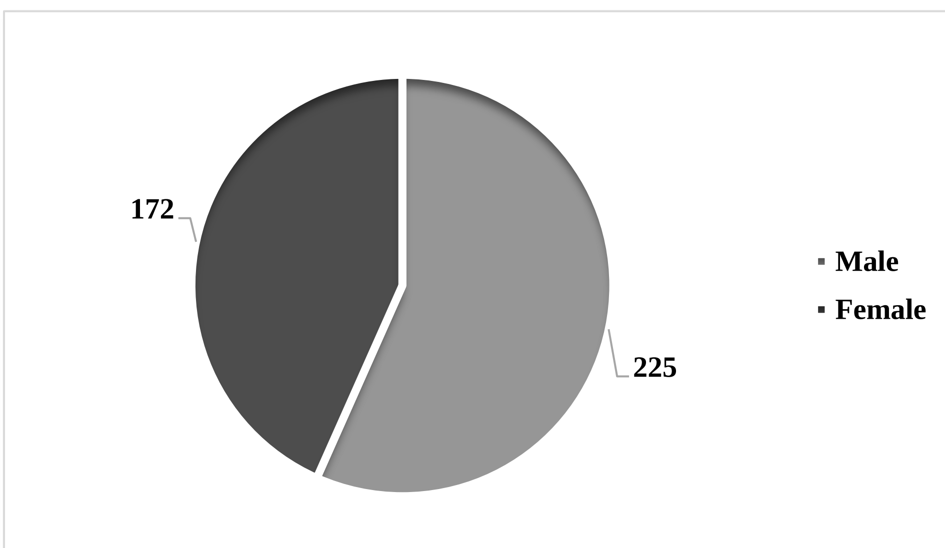


Figure 24

Bar Chart of COVID-19 Death Rate per 100,000 in 2020 by Race in Brazil

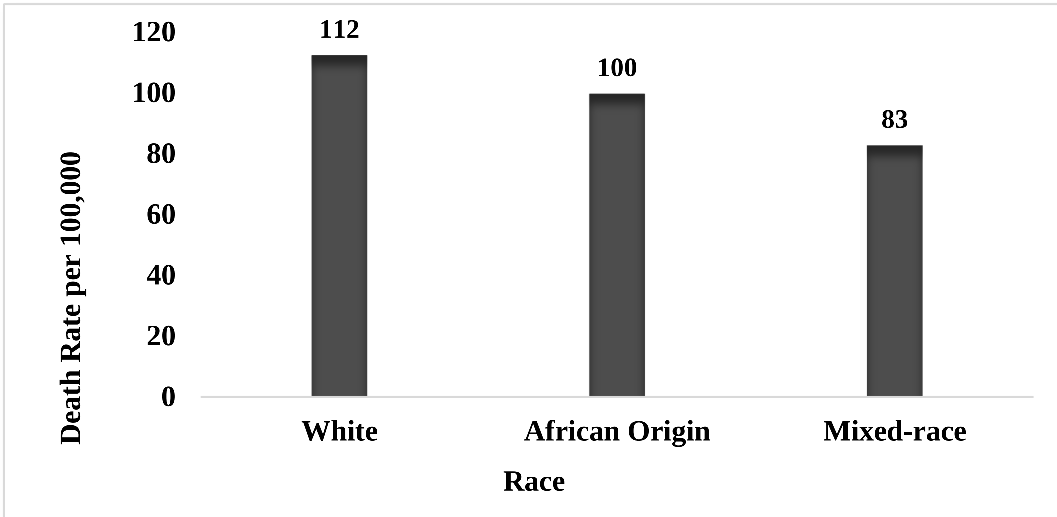


Figure 25

Bar Chart of COVID-19 Death Rate per 100,000 in 2021 by Race in Brazil

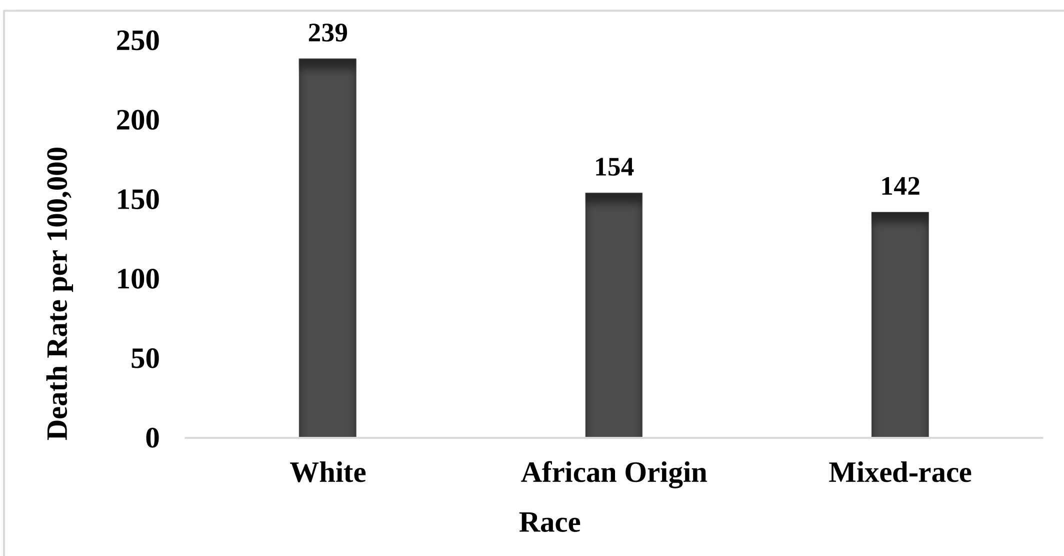


Figure 26

Bar Chart of COVID-19 Death Rate per 100,000 in 2020 by Region

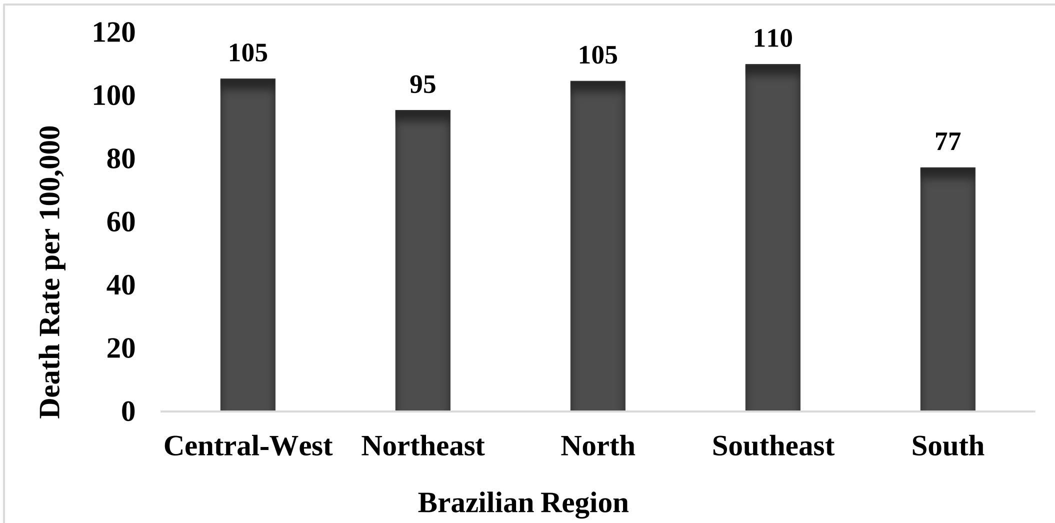
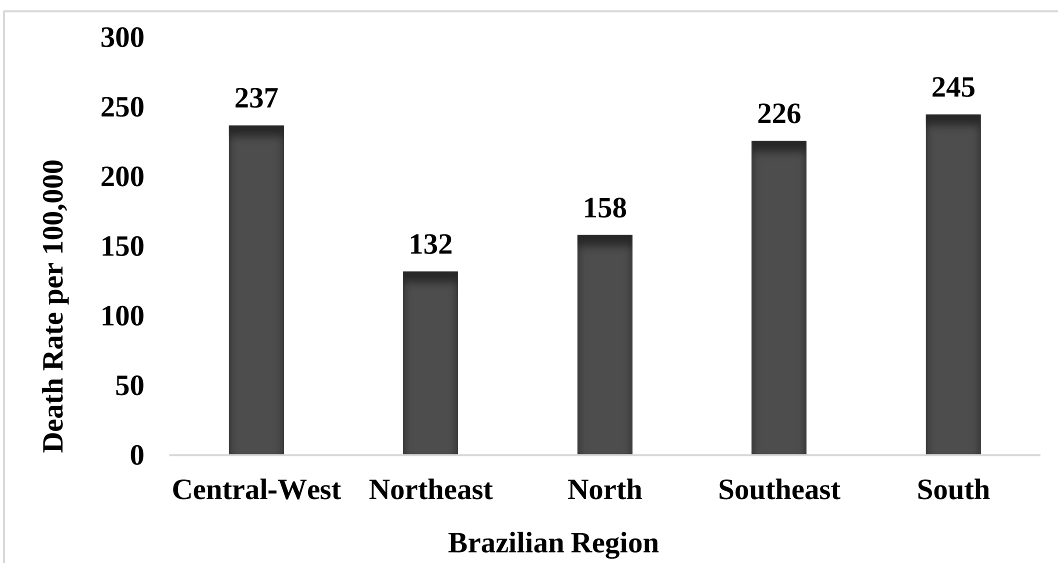


Figure 27

Bar Chart of COVID-19 Death Rate per 100,000 in 2021 by Region



Discussion

Based on the fundamental causes theory, disease spread and mortality are greater among individuals with fewer resources (Link et al., 1995). Therefore, mortality due to COVID-19 is expected to be higher among individuals who are socioeconomically disadvantaged or part of minority groups (Mello, 2020; Ribeiro et al., 2021). During a multiple linear regression analysis to investigate the association between education, measured by school enrollment, and COVID-19 mortality in Brazil while controlling for variable related to the population's race, age, sex, and immunization rate, it was found that school enrollment is a significant predictor of COVID-19 mortality at the state level. Conversely, no statistically significant association was found between school enrollment and COVID-19 mortality rate across Brazilian capitals. Results at the regional level were not reported, as the model was found to be invalid in predicting COVID-19 mortality from income at this level, probably due to the small sample size.

Looking at the distribution of COVID-19 mortality in Brazil during the first two years of the pandemic, it was found that the COVID-19 death rate was higher in the Southeast region in 2020 and higher in the Southern region in 2021, the latter being the most densely populated regions in the country. Based on this observation, we can suggest that the irregularity or heterogeneity of the population's distribution can shape disease spread, as investigated by Martins-Filho et al. (2021) and Miranda et al. (2020). The same analysis found that the COVID-19 death rate was higher among individuals 80 years old or older, male and White, in both years of the pandemic.

As mentioned in the Summary of the Manuscripts section, most of the study limitations were associated with data availability. Two-year data for disease mortality were obtained for each region, state, and capital. Although the data used for the regression analyses were population-based, the data obtained at a national level was converted to an individual-level dataset to investigate death distribution for the first two years of the pandemic. To avoid errors during the manipulation of this large dataset, the data were reviewed a second time to ensure transcription errors were not present. Another limitation was regarding data availability. Since COVID-19 is an ongoing pandemic, many data points were unavailable or ready for analyses, resulting in a need for some manipulation to guarantee dataset completion and readiness before analyses. Data manipulation was required to avoid internal validity issues, but data should be considered reliable as all information was retrieved from governmental resources.

The findings of this study empower public health and political leaders to identify geographic areas in the country that need special attention regarding the development of policies and implementation of public health programs to encourage populations' best behavior and practices concerning preventive actions and vaccination willingness. Furthermore, considering Brazil a country of high economic and demographic disparities, the social change implications of this study may include a potential increase in people's opportunities for socioeconomic equality, particularly a fair income distribution, which could allow individuals to have better access to healthcare, consequently decreasing COVID-19 mortality rates. A decrease in the COVID-19 mortality rate could allow Brazilians to improve individual lives with healthier families and communities.

Based on this study's findings, and to better explore the influences of SES as a risk factor for COVID-19 mortality, further research is recommended with both individual and population-based approaches. Since COVID-19 is an ongoing pandemic, it is likely that more accurate results will only be obtained with post-pandemic data. Lastly, multiple linear regression models may not be the best model for analyzing population-based data considering that sample sizes may be limited compared to individual-based data.

References

Aschengrau, A., & Seage, G. R. (2013). *Essentials of epidemiology in public health*. Jones & Bartlett Publishers.

Brazilian Institute of Geography and Statistics [IBGE] (n.d.). https://www.ibge.gov.br/estatisticas/sociais/educacao/9221-sintese-de-indicadores-sociais.html?edicao=32373&t=downloads.

Brazilian Institute of Geography and Statistics [IBGE] (n.d.). Densidade demográfica [Demographic density]. Retrieved from https://www.ibge.gov.br/geociencias/cartas-e-mapas/sociedade-e-economia/15955-densidade-demografica.html?=&t=acesso-ao-produto

Brazilian Institution of Geography and Statistics [IBGE] (2019). Conheça o Brasil – População – Educação. Retrieved from https://educa.ibge.gov.br/jovens/conheca-o-brasil/populacao/18317-educacao.html

Clouston, S. A. P., Natale, G., & Link, B. G. (2021). Socioeconomic inequalities in the spread of coronavirus-19 in the United States: A examination of the emergence of social inequalities. *Social Science & Medicine*, *268*, 1-6. https://doi.org/10.1016/j.socscimed.2020.113554

Cury, C. R. J. (2008). A educação escolar, a exclusão e seus destinatários. *Educação em Revista*, 205-222. https://doi.org/10.1590/S0102-46982008000200010

Ferreira, V. M., Isolani, L., Diniz, R. C. S., de Oliveira Karklins, E., de Miranda, L. M. P., Ramos, L. H. P., & de Melo, S. N. (2021). Avaliação epidemiológica das regiões do Brasil na pandemia de COVID-19. *Revista Eletrônica Acervo*

Saúde, *13*(4), e7137-e7137. https://doi.org/10.25248/reas.e7137.2021

Governo Brasileiro (1961). Presidencia da República. Casa Civil Subchefia para Assuntos Juridicos. Retrieved from http://www.planalto.gov.br/ccivil_03/leis/L4024.htm.

Ribeiro, V. M., & Vóvio, C. L. (2017). Desigualdade escolar e vulnerabilidade social no território. *Educar em Revista*, 71-87. https://doi.org/10.1590/0104-4060.51372

Secretaria de Vigilância em Saúde [Health Surveillance Secretariat] (2022). Departamento de Análise Epidemiológica e Vigilância de Doenças Não Transmissíveis [Department of Epidemiological Analysis and Surveillance of Noncommunicable Diseases]. Portal do Governo Brasileiro. https://svs.aids.gov.br/daent/centrais-de-conteudos/paineis-de-monitoramento/mortalidade/cid10/

Warner, R. M. (2020). *Applied statistics: From bivariate through multivariate techniques*. Sage.

Manuscript 3: Association of Employment and COVID-19 Mortality in Brazil

[Associação entre a Situação Ocupacional do Brasileiro e a Mortalidade por COVID-19].

Marcelina Machado

Walden University

Outlet for Manuscript

The manuscript on Brazil's employment and COVID-19 mortality will be submitted to the Revista Brasileira de Epidemiologia [Brazilian Journal of Epidemiology] https://www.abrasco.org.br/site/revistas/revista-brasileira-de-epidemiologia/. The journal requires a manuscript with the following contents (1) a cover sheet with titles in English and Portuguese in up to 25 words, the authors' information, acknowledgments in up to 70 words, and information regarding conflict of interest and sponsorship. (2) A structured summary and abstract in up to 250 words containing the study's objectives, methodology, results, and conclusion and four to six keywords will also be included. (3) AMA formatting in Times New Roman with double spacing and font size of 12 is required. The journal also requires abstracts in Portuguese for articles in English, original research results reported in up to 3400 words, and exclusivity in the first publication. The journal's mission is to expand knowledge and development of Epidemiology and related sciences. The journal's mission is aligned with the contents of the present manuscript in a way the results of this study could provide valuable insights for public health practitioners and policymakers, allowing them to evaluate the association between public demographics, employment, and COVID-19 mortality across Brazil.

Abstract

Objective. Investigating the association between employment and COVID-19 mortality in Brazil, while controlling for variables related to race, age, sex, and immunization rate.

Methodology. In this retrospective study, employment rate data were retrieved from the Brazilian Institute of Geography and Statistics, COVID-19 mortality data retrieved from the Health Surveillance Secretariat, population demographics data including race, age, and sex were retrieved from the Ministry of Health, and immunization data were retrieved from the Oswaldo Cruz Foundation. All data were analyzed at region, state, and capital-levels using multiple linear regression on IBM SPSS Statistics version 27.

Results. The findings suggest a statistically significant association between employment and COVID-19 mortality rate at the state level (p=0.023), but not a statistically significant association at the capital level (p=0.088). At the state level, we can predict an increase of three COVID-19 deaths per 100,000 people for each unit increase in employment rate. Variables related to race, age, sex, and immunization were used as control variables in all the regression analyses. Results show that the model is invalid to predict COVID-19 mortality from employment at the regional level.

Conclusion. To better explore the influences of employment on COVID-19 mortality, further research is recommended with both individual and population-based approaches. Since COVID-19 is an ongoing pandemic, it is more likely that more accurate results will be only obtained with post-pandemic data.

***Keywords*:** COVID-19 mortality, employment, Brazil.

Resumo

Objetivo. Investigar a associação entre situação ocupacional e taxa de mortalidade por COVID-19 no Brasil levenado em consideração outros dados demográfico da população incluindo raça, idade e sexo assim como a taxa de immunização.

Metodologia. Nesse estudo quantitativo retrospectivo, dados sobre taxa de emprego foram obtidos do Instituto Brasileiro de Geografia e Estatística, dados sobre mortalidade por COVID-19 foram obtidos da Secretaria de Vigilância em Saúde, dados demográficos populacionais, incluindo raça, idade e sexo, foram retirados do Ministério da Saúde, e dados de imunização foram retirados da Fundação Oswaldo Cruz. Todos os dados foram analisados em niveis de região, estado e capital usando análise de regressão linear multipla no softawre IBM SPSS Statistics versão 27.

Resultados. Os achados sugerem uma associação estatisticamente significativa entre emprego e taxa de mortalidade por COVID-19 no nível estadual (p=0.023), mas não no nível de capital (p=0.088). A nível estadual, podemos prever um aumento de 3 mortes por COVID-19 por 100,000 pessoas por cada aumento unitário da taxa de emprego. Raça, idade, sexo e imunização foram utilizadas como controle. Resultados mostram que o modelo é inválido para prever a mortalidade por COVID-19 em nível regional.

Conclusão. A fim de explorar melhor as influências do SES na mortalidade por COVID-19, recomenda-se que mais pesquisas sejam feitas com abordagens individuais e de base populacional. Como a COVID-19 é uma pandemia em curso, é mais provável que resultados mais precisos só sejam obtidos com dados pós-pandemia.

Palavras-chave: COVID-19, mortalidade, emprego, Brasil.

Introduction

In December 2019, the new coronavirus disease spread in China, becoming a global pandemic on March 11, 2020 (WHO, 2020). Due to the lack of treatment and the unpreparedness to face the rapid growth of cases, many countries found themselves in a desperate situation, seeing the number of deaths multiplying (WHO, 2021). Today in Brazil, the number of fatalities has reached over 600 thousand. This number is concentrated in the country's southeast region, where population density is higher, and virus transmission is potentially more likely to occur (Martins-Filho, 2021). However, besides the country's dimension and its disproportionate population density, Brazil presents high socioeconomic inequalities, with 24.1% of the Brazilian population (about 50 million people) living in poverty, earning an average U$$5.50/day, and 5.7% (about 12 million people) living in extreme poverty, making an average U$$1.90/day (IBGE, 2021). Due to the COVID-19 pandemic in Brazil, the unemployment rate jumped from 11.1% to 14.2% in December 2019 to 2020 as people withdrew from the workforce and lost income (IBGE, n.d.; IBGE, 2021). With the increased unemployment during the pandemic, the government implemented emergency aid to alleviate the financial burden. As vaccines became available in 2021, providing some hope for a return to normalcy, unemployment rates returned to 11.1% in December 2021 (IBGE, n.d.).

The issue that prompted me to search the literature is that increasing COVID-19 mortality rates can negatively affect vulnerable populations and become a severe burden to the public health system of developing nations. There is strong evidence that the high mortality rate in Brazil is associated with the population's demographics and socioeconomic factors (Baqui et al., 2021; Miranda et al., 2020; Rocha et al., 2021). In this study, I focused on employment as a factor that could be associated with the increase in COVID-19 mortality. From March 2020 to December 2021, 634,695 deaths were registered in Brazil (Secretaria de Vigilância em Saúde [Health Surveillance Secretariat], 2022), while the unemployment rate grew considerably (IBGE, n.d.).

The significance of this study lies in the fact that its findings could provide public health leaders and policymakers in Brazil with the information needed to assess how individuals' employment situation impacts COVID-19 mortality rates. By evaluating that, plans could be developed and implemented to either increase employment opportunities for those who need it or provide more financial assistance during public health crises. However, filling this gap does not imply that other socioeconomic factors should be ignored, as studies have shown that multiple factors could directly or indirectly influence COVID-19 mortality.

Although compliance with measures such as social distancing is essential for disease prevention and control, social distance has worsened socioeconomic vulnerability and inequality in countries like Brazil (Figueiredo et al., 2020; Garnier et al., 2021; Natividade et al., 2020). In a developing country where many do not have the luxury of working remotely, staying at home means a reduced income, while working out of the home during the pandemic increases individuals' risk of infection and death. Additionally, Chen et al. (2021), Figueiredo et al. (2020), and Oliveira et al. (2020) explained that COVID-19 incidence and mortality increase vary depending on the population's demographic heterogeneity and other factors concerning the population's socioeconomic conditions, but that would not be all. These studies provide evidence that increased mortality rates from COVID-19 across Brazil can be associated with differences in the country's preparedness, testing, reporting and discrepancies in local policies, medical infrastructure, and healthcare access. For this study, I evaluated the association between COVID-19 mortality in Brazil and employment after controlling for variables related to race, age, sex, and immunization.

Research Question & Hypotheses

This quantitative study addressed the following research question:

Research Question for Manuscript 3 (RQM3): What is the association between employment and COVID-19 mortality rate at different geographic levels across Brazil after controlling for race, age, sex, and immunization rate?

Based on the research question, the following hypotheses were formulated:

Null Hypothesis (H0): There is no statistically significant association between employment and COVID-19 mortality rate at different geographic levels across Brazil after controlling for race, age, sex, and immunization rate.

Alternative Hypothesis (H1): There is a statistically significant association between employment and COVID-19 mortality rate at different geographic levels across Brazil after controlling for race, age, sex, and immunization rate.

The rationale for this research is that, historically, Brazil has suffered from the transmission of numerous infectious diseases concentrated in overpopulated and poor socioeconomic areas (Barreto et al., 2011; Waldman et al., 2016). Throughout the years, the Brazilian government has had many successes and some failures in its trajectory to fight infectious diseases by implementing programs that deal with critical determinants such as sanitation and water treatment as well as vector control and vaccination programs (Barreto et al., 2011; Waldman et al., 2016). However, actions still need to be taken concerning socioeconomic improvements in the country to reduce the risks and likelihood of disease spread and improve the chances of control and prevention (Barreto et al., 2011). To address the research question, this research design includes a retrospective cohort design (Aschengrau et al., 2020) with mortality rates examined between the period of March 2020 to December 2021. This quantitative analysis should help to determine if employment and population demographic characteristics are correlated to COVID-19 mortality in Brazil.

Methodology

Study Population

The target population for this study include Brazilian individuals with different demographic and geographic characteristics, including race (White, African origin, and Mixed), age (0-14, 15-49, and 50-80 years old), sex (male and female), and immunization. Public dataset repositories were used for this research design to obtain data on the COVID-19 mortality population's socioeconomic, demographic, and geographic characteristics. Since this is a population-based study, aggregated data from different geographic levels, including region, state, and capitals, were used for all dependent, independent, and control variables. For this study, convenience sampling was used (Warner, 2013). Based on the data retrieved for each geographic level during the first two years of the pandemic, the convenience sample size for regions is 10, for states is 54, and for capitals is 54. Additionally, information on individual-based data was still used to show death distribution throughout the first years of the pandemic according to race, age group, sex, and geographic region, but were used as information only and not used to test the study hypothesis.

Variables

The variables of this study include the following.

Independent variable (IV):

- Employment: Numbers in percentage indicating individuals who are employed. Actual percentage values from different geographic levels (regional, state, and capital levels) were used as found in the data source.

Dependent variable (DV):

- COVID-19 mortality rate: The mortality rate was calculated for different geographic areas by dividing the number of deaths from COVID-19 in each area by the number of people in the population, then multiplied by 100,000.

Control variables (CV1 – CV8):

- White population: Numbers in percentage. The percentage of the White population was calculated for different geographic areas by dividing the number of White people in each area by the number of people in the population, then multiplied by 100.
- African origin population: Numbers in percentage. The percentage of the African-origin population was calculated for different geographic areas by dividing the

number of African-origin people in each area by the number of people in the population, then multiplied by 100.

- Mixed-race population: Numbers in percentage. The percentage of the Mixed-race population was calculated for different geographic areas by dividing the number of Mixed-race people in each area by the number of people in the population, then multiplied by 100.
- Population between 0-14 years old: Numbers in percentage. The percentage of the 0-14 years old population was calculated for different geographic areas by dividing the number of 0-14 years old people in each area by the number of people in the population, then multiplied by 100.
- Population between 15-49 years old: Numbers in percentage. The percentage of the 15-49 years old population was calculated for different geographic areas by dividing the number of 15-49 years old people in each area by the number of people in the population, then multiplied by 100.
- Population between 50-80 years old: Numbers in percentage. The percentage of the 50-80 years old population was calculated for different geographic areas by dividing the number of 50-80 years old people in each area by the number of people in the population, then multiplied by 100.
- Male population: Numbers in percentage. The percentage of the male population was calculated for different geographic areas by dividing the number of male people in each area by the number of people in the population, then multiplied by 100.

- Immunization rate: Number in percentage. a higher percentage indicates a higher immunization. Actual percentage values from different geographic levels (regional, state, and capital levels) were used as found in the data source There were no codes from missing data.

Sources of Data

Data on employment per geographic location from 2020 and 2021 were accessed on September 18, 2022, from the publicly available dataset repository below:

- Instituto Brasileiro de Geografia e Estatistica (IBGE) [Brazilian Institute of Geography and Statistics].

Data on COVID-19 mortality per geographic location from March 2020 through December 2021 were accessed on September 18, 2022, from the publicly available dataset repository below:

- Secretaria de Vigilância em Saúde [Health Surveillance Secretariat], created by the Ministério da Saúde [Ministry of Health].

Data on the population's demographics, including race, age, and sex per geographic location from 2020 and 2021, were accessed on September 18, 2022, from the publicly available dataset repositories below:

- Instituto Brasileiro de Geografia e Estatistica (IBGE) [Brazilian Institute of Geography and Statistics].
- Tabnet – DATASUS - Ministério da Saúde [Ministry of Health].

Data on the population's immunization (or percentage of population vaccinated with the second dose) per geographic location from March 2021 through September 2022 were accessed on September 18, 2022, from the publicly available dataset repository below:

- Fundação Oswaldo Cruz (Fiocruz) [Oswaldo Cruz Foundation].

Access to public dataset repositories were needed to gather specific data points, including the number of deaths by COVID-19 at region, state, and capital-levels and the population's socioeconomic and demographic information. To minimize potential issues regarding the use of secondary data, datasets were screened for errors, outliers, inconsistencies, and missing values prior to analysis. Identified errors in the datasets were reported in the discussion section.

Validity and Reliability of Data

Internal validity, which is "the degree to which the results of the study can be used to make causal inferences," is often weak in non-experimental studies (Warner, 2013). To increase internal validity by limiting the influence of confounding, a series of control variables, related to race, age, sex, and immunization rate, were included in the analysis. Additionally, regression analyses were performed separately at each geographic level (region, state, and capital) to avoid violation of internal validity. On the other hand, external validity, the "degree to which the study's results can be generalized" (Warner, 2013) could be considered high in this study as the data to be used will represent a large and heterogeneous population (Oliveira, 2020).

The reliability of secondary data depends on the sources from which the data were obtained. Datasets from government sources are often reputable; however, data screening is still recommended to identify errors and inconsistencies (Warner, 2013). The datasets used in this study will be retrieved from Brazilian governmental websites listed in the previous section. Also, data will be reviewed before analysis to ensure the accuracy of data transcription.

Design and Analysis

A multiple linear regression (MLR) analysis was performed using IBM SPSS Statistics version 27 to determine the association between employment and COVID-19 mortality rate at different geographic levels across Brazil while controlling for variables related to race, age, sex, and immunization rate. Multiple linear regression is a statistical technique often used to evaluate the relationship of one continuous dependent variable and multiple continuous or categorical independent variables (Warner, 2013).

Although aggregated data (population-based data) were accessed and used in the primary statistical analysis (MLR) to answer the research question previously elaborated, individual-based data on COVID-19 mortality and population demographics, including race, age, sex, and region, were used to evaluate death distribution in the country during the first two years of the pandemic. Additionally, descriptive statistical analyses were performed for the COVID-19 mortality rate and employment rate variables. Lastly, prior to the performance of the primary analysis, assumptions for MLR were tested. Table 17 and Figure 28 summarize of the proposed analytical test plan.

Table 17

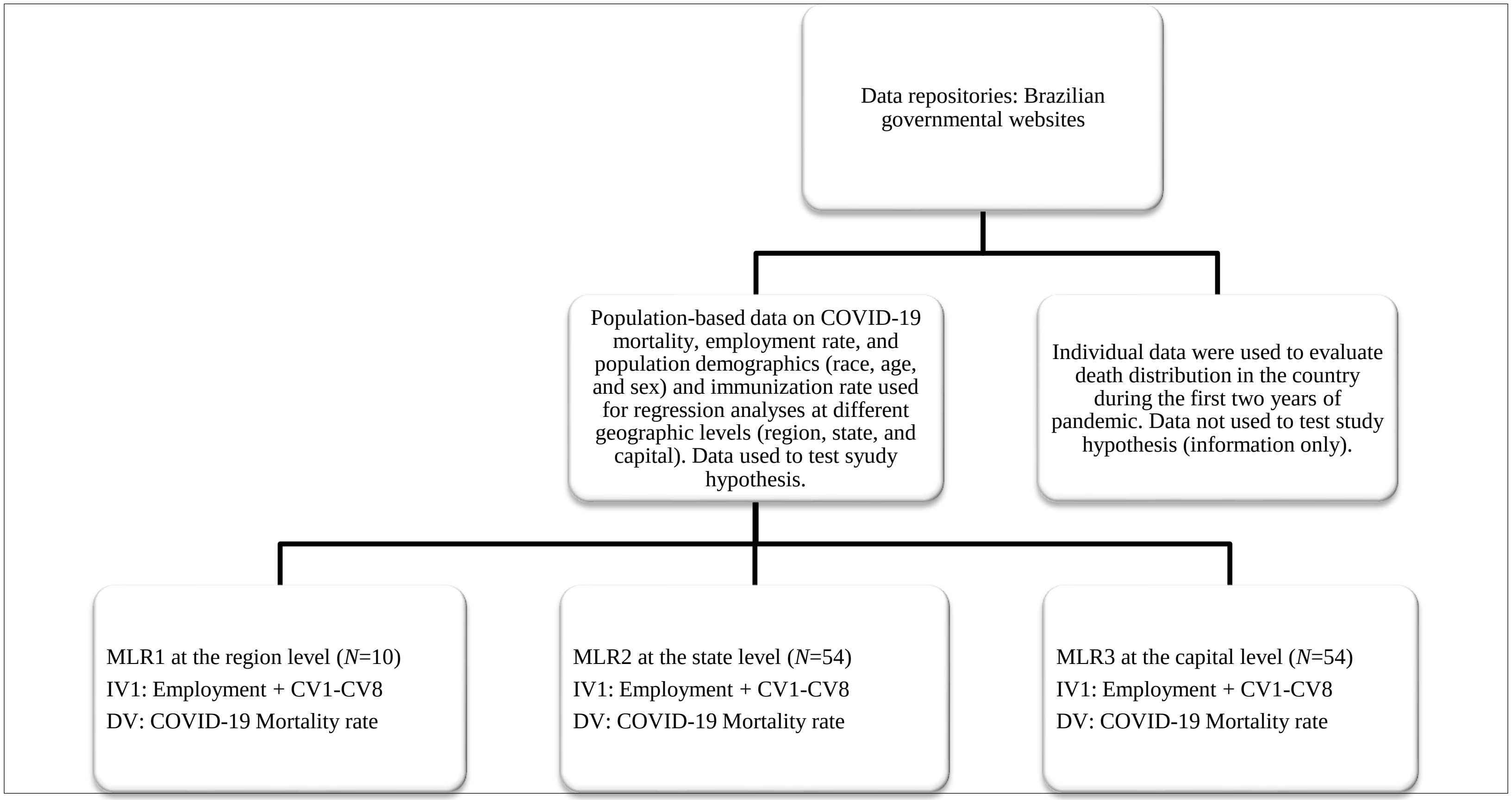
Proposed Analytical Test Plan and Justification

Variable	Type of variable	Multiple Linear Regression (MLR)	Death Distribution Evaluation
COVID-19 Mortality rate [1]	Dependent Variable **DV** (Scale)	Using a population-based dataset and including all variables listed, multiple linear regression analyses were performed to determine the association between school employment and COVID-19 mortality while controlling for the population's demographic characteristics and immunization. MLR is an appropriate test considering that the dependent variable (disease mortality) and the independent variable (employment rate) are continuous variables, and the control variables are scale (Warner, 2013). Assumptions tested: Normality, linearity, homoscedasticity, and multicollinearity. In these analyses, all variables listed were used.	An individual-based dataset was created using only COVID-19 mortality and population demographics, including race, age, sex, and region, to evaluate death distribution in the country during the first two years of the pandemic. Data not used to test study hypothesis. Analyses were done for information only.
Percentage of Employed population [1]	Independent Variables **IV1** (Scale)		
Percentage of White population	Control Variable **CV1** (Scale)		
Percentage of African origin population	Control Variable **CV2** (Scale)		
Percentage of Mixed-race population	Control Variable **CV3** (Scale)		
Percentage of Population between 0-14 Years Old	Control Variable **CV4** (Scale)		
Percentage of Population between 15-49 Years Old	Control Variable **CV5** (Scale)		
Percentage of Population between 50-80 Years Old	Control Variable **CV6** (Scale)		
Percentage of Male Population	Control Variable **CV7** (Scale)		
Percentage of Immunized Population	Control Variable **CV8** (Scale)		

Note. [1] Descriptive statistical analysis were performed on the COVID-19 mortality rate and employment rate variables.

Figure 28

Scheme of Data Flow and Regression Analyses for Manuscript 3



Note. Regression analyses were performed at each geographic level (region, state, and capital) to avoid internal validity.

Results

The study was conducted by retrieving population-based data on COVID-19 mortality, school enrollment, and population demographics, including race, age, and sex, as well as immunization rate from Brazilian governmental websites and analyzed using multiple linear regression model at different geographic levels (see Figure 28) on IBM SPSS software. The population-based data was used to create a dataset containing all information for the regression analyses. The dataset was further converted into an individual-based dataset only on COVID-19 mortality and population demographics to investigate the death distribution during the first two years of the pandemic.

Most data needed for the analyses were retrieved without issues, except for the immunization data, which were not available at the capital-level; therefore, this variable was not included in the regression analysis for this level. Additionally, as vaccination campaigns started on January 2021 in Brazil, the data retrieved for this study are from 2021 and 2022. It is also important to note that the regression at the nation-level was not performed due to the small sample size (N=2).

Descriptive Statistics for COVID-19 Mortality

After running a descriptive statistics analysis on the variable COVID-19 mortality, it was observed that among the 27 states during two years of the pandemic (N= 54), the COVID-19 mortality rate ranged from 0.063% to 0.284% at the state level. The three measures of central tendency for this variable were mean equals 0.14% (M = 0.14), median equals 0.13% (Median = 0.13), and mode equals 0.063% (Mode = 0.063) with a standard deviation equal 0.058 (SD = 0.058). The skewness and kurtosis of this distribution (0.816 and -0.284, respectively) suggest that the distribution deviates from normality and is positively skewed (Table 18 and Figure 29). In Figure 29, the variability or measure of spread distribution of the COVID-19 mortality rate was 0.18% to 0.10%; outliers were not identified.

In a second descriptive statistics analysis, it was observed that among the 27 capitals during the two years of the pandemic (*N*= 54), the COVID-19 mortality rate ranged from 0.107% to 0.519% at the capital level. The three measures of central tendency for this variable were mean equals 0.25% (*M* = 0.25), median equals 0.23% (Median = 0.23), and mode equals 0.107% (Mode = 0.107) with standard deviation equals 0.094 (*SD* = 0.094). The skewness and kurtosis of this distribution (0.325 and -0.886, respectively) suggest that the distribution deviates from normality and is slightly positively skewed (Table 18 and Figure 30). In Figure 30, the variability or measure of spread distribution of the COVID-19 mortality rate was 0.28% to 0.18%; outliers were identified.

Table 18

Statistics of COVID-19 Mortality at State and Capital-levels

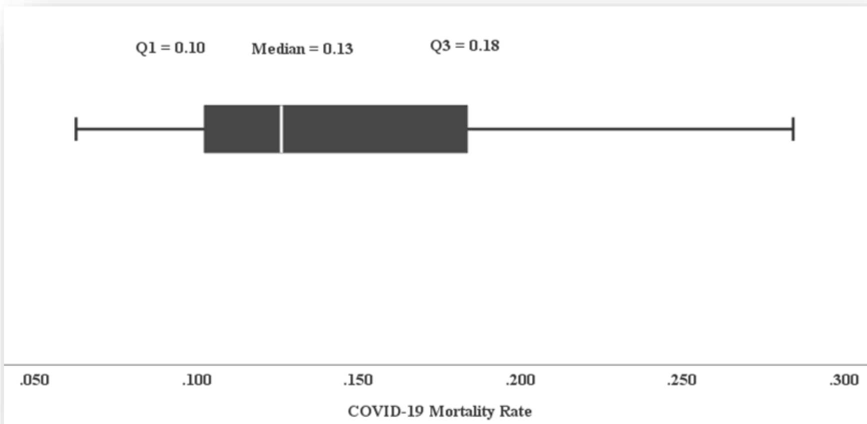
		State-level	Capital-level
N	Valid	54	54
	Missing	0	0
Mean		0.14169	0.24528
Median		0.12622	0.23056
Mode		0.063[1]	0.107 [1]
Std. Deviation		0.058010	0.093895
Variance		0.003	0.009
Skewness		0.816	0.989
Std. Error of Skewness		0.325	0.325
Kurtosis		-0.284	0.886
Std. Error of Kurtosis		0.639	0.639
Range		0.221	0.411
Minimum		0.063	0.107
Maximum		0.284	0.519
Percentiles	25	0.10052	0.18157
	50	0.12622	0.23056
	75	0.18464	0.28142

Note. Statistics were performed at state and capital levels for the first two years of the pandemic. For that, the *Frequencies* function in SPSS was used to observe further Percentile values, Central Tendency, Dispersion, and sample Distribution.

[1] Multiple mode exist. The smallest value is shown.

Figure 29

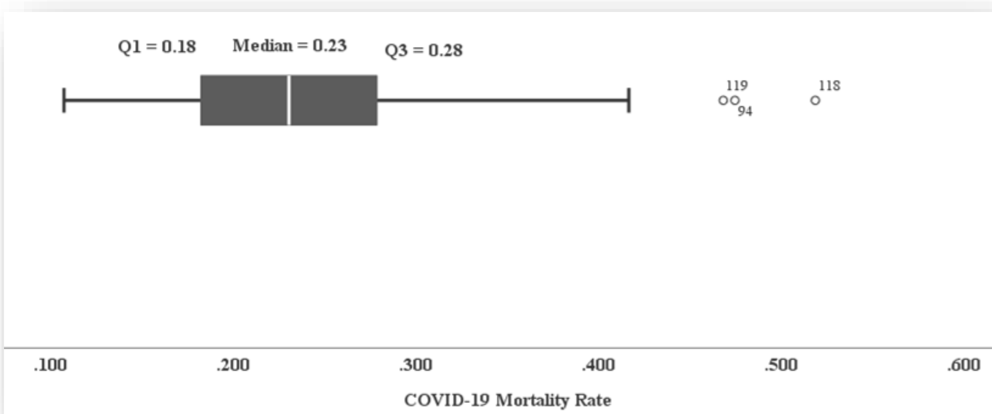
Boxplot Graph Representing a Positively Skewed Distribution of COVID-19 Mortality at State-level



Note. No outliers were identified.

Figure 30

Boxplot Graph Representing a Positively Skewed Distribution of COVID-19 Mortality at Capital-level



Note. Outliers were identified.

Descriptive Statistics for Employment

After running a descriptive statistics analysis, it was observed that among the 27 states during the two years of the pandemic (*N*= 54), the employment rate ranged from 32.1% to 79.4% at the state level. The three measures of central tendency for this variable are mean equals 52.9% (*M* = 52.9), median equals 50.3% (Median = 50.3), and mode equals 43.4% (Mode = 43.4) with standard deviation equals 12.1 (*SD* = 12.1). The skewness and kurtosis of this distribution (0.271 and -0.862, respectively) suggest that the distribution deviates from normality and is positively skewed (Table 19 and Figure 31). In Figure 31, the employment rate's variability or measure of spread distribution was 61.8% to 43.4%; outliers were not identified.

In a second descriptive statistics analysis, it was observed that among the 27 capitals during the two years of the pandemic (*N*=54), the employment rate ranged from 48.9% to 67.5% at the capital level. The three measures of central tendency for this variable are mean equals 57.6% (*M* = 57.6), median equals 57.2% (Median = 57.2), and mode equals 61.1% (Mode = 61.1) with standard deviation equals 3.9 (*SD* = 3.9). The skewness and kurtosis of this distribution (0.188 and -0.424, respectively) suggest a non-normal distribution and slightly positively skewed (Table 19 and Figure 32). In Figure 32, the variability or measure of spread distribution of the school enrollment was 61.1% to 54.6%; outliers were not identified. Note that for both descriptive analyses, the high value of variance (the square root of the standard deviation) may indicate highly scattered data around the mean.

Table 19

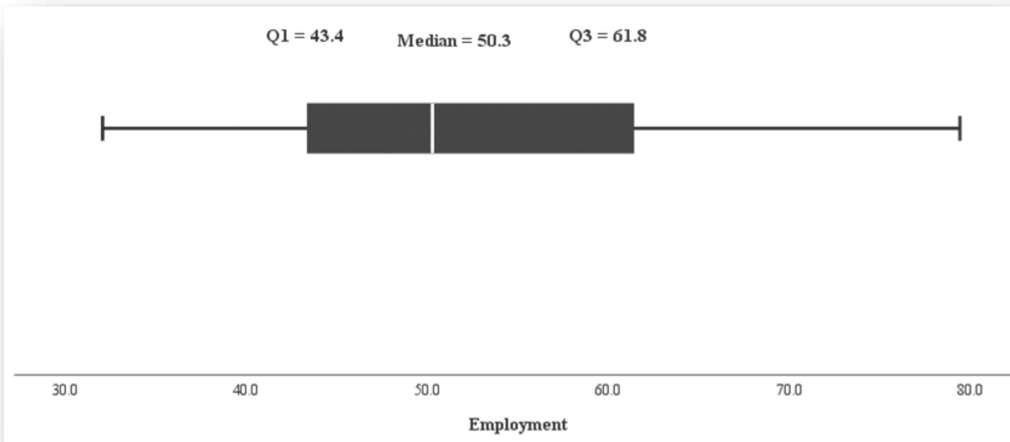
Statistics of Employment at State and Capital-levels

		State-level	Capital-level
N	Valid	54	54
	Missing	0	0
Mean		52.931	57.554
Median		50.300	57.200
Mode		43.4[a]	61.1
Std. Deviation		12.1090	3.9562
Variance [b]		146.628	15.651
Skewness		0.271	0.188
Std. Error of Skewness		0.325	0.325
Kurtosis		-0.862	-0.424
Std. Error of Kurtosis		0.639	0.639
Range		47.3	18.6
Minimum		32.1	48.9
Maximum		79.4	67.5
Percentiles	25	43.400	54.550
	50	50.300	57.200
	75	61.775	61.100

Note. Statistics were performed at state and capital levels for the first two years of the pandemic. For that, the *Frequencies* function in SPSS was used to observe further Percentile values, Central Tendency, Dispersion, and sample Distribution. a. Multiple modes exist. The smallest value is shown. b. Note that for both descriptive analyses, the high value of variance (the square root of the standard deviation) may indicate highly scattered data around the mean.

Figure 31

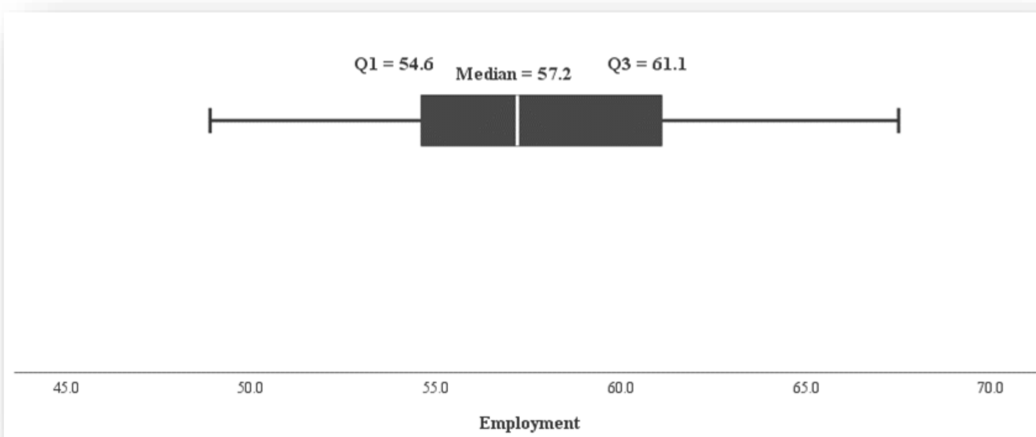
Boxplot Graph Representing the Positively Skewed Distribution of Employment According to State



Note. No outliers were identified.

Figure 32

Boxplot Graph Representing the Positively Skewed Distribution of Employment According to Capital



Note. No outliers were identified.

Assumptions of Multiple linear Regression

The assumptions tested for the MLR analyses were normality, linearity, homoscedasticity, and multicollinearity. Normality was tested by observing the Normal P-P Plot, linearity and homoscedasticity were tested by observing the scatterplot of residuals, and multicollinearity was tested by observing the variance inflation factor (VIF) obtained from the coefficient table in the SPSS output. Additionally, the independence of residuals was tested by observing the Durbin-Watson value. Outliners were tested by looking for points that were far from the regression line.

Table 20 summarizes the results for the assumptions of the multiple linear regression analyses. Note that the passing criterion set for normality was that residuals should be normally distributed. The passing criterion for linearity was that residuals should have a straight-line relationship with predicted DV scores. The passing criterion for homoscedasticity was that VIF values should be less than 10 (VIF<10). The passing criterion for the independence of residuals is that plotted standardized residuals are randomly scattered. The passing criterion for multicollinearity was that Durbin-Watson values should be between 1.5 – 2.5 (approximate values were acceptable). Overall, the results show that except for linear regression at the region-level, none of the assumptions were violated.

Table 20

Results of Assumptions of Multiple linear Regression for Manuscript 3

Level	Regression/IV and DV	Normality (passing criterion is that residuals should be normal distributed) **(Pass or fail criterion)**	Linearity (passing criterion is that residuals should have a straight-line relationship with predicted DV scores) **(Pass or fail criterion)**	Homoscedasticity (passing criterion is that plotted standardized residuals are randomly scattered) **(Pass or fail criterion)**	Multicollinearity (passing criterion is VIF <10) **(Pass or fail criterion)**	Independence of Residuals (passing criterion is Durbin-Watson values between 1.5-2.5[1]) **(Pass or fail criterion)**	Outliers (passing criterion is absence of outliers) **(Pass or fail criterion)**
Region	Regression 1 (Employment and COVID-19 mortality)	Fail	Fail	Fail	Fail	Fail	Pass
State	Regression 2 (Employment and COVID-19 mortality)	Pass	Pass	Pass	Pass[2]	Pass	Pass
Capital	Regression 3 (Employment and COVID-19 mortality)	Pass	Pass	Pass	Pass[2]	Pass	Pass

Note. [1]Approximate values are acceptable. [2] Failed for 2 non-significant variables.

Results of Multiple linear Regression

To approach the association between employment and COVID-19 mortality in Brazilian regions, a multiple linear regression analysis was conducted to evaluate the prediction of COVID-19 mortality rate from employment while controlling for variables related to the population's race, age, sex, and immunization rate (see list of variables included in the Design and Analysis section). Preliminary analyses were conducted to assess the assumptions of normality, linearity, homoscedasticity, multicollinearity, independence of residuals, and outliers. Violations were noted, indicating that the model is invalid in predicting COVID-19 mortality at the regional level (see Table 20). Therefore, regression results for this level were not reported as they cannot be used to interpret the data correctly.

The same approach was taken to evaluate the association between employment and COVID-19 mortality rate in Brazilian states. A multiple linear regression analysis was conducted to evaluate the prediction of COVID-19 mortality from employment while controlling for variables related to the population's race, age, sex, and immunization rate (see control variables listed in Table 21). Preliminary analyses were conducted to assess the assumptions of normality, linearity, homoscedasticity, multicollinearity, independence of residuals, and outliers. No violations were noted (see Table 20). The multiple linear regression analysis results revealed that at the state level, a significant regression equation was found $F(8, 45)=4.620$, $p<0.001$), with an R^2 of 0.451. Table 21 summarizes the results of the regression performed at the state level showing that for each percentage point increase in employment, we can predict an increase of 3 in

COVID-19 deaths per 100,000. Employment rate is a significant predictor of COVID-19 mortality (*p*=0.023) while controlling for variable related to the population's race, age, sex, and immunization rate.

Lastly, to approach the association between employment and COVID-19 mortality in Brazilian capitals, a multiple linear regression analysis was conducted to evaluate the prediction of COVID-19 mortality from employment while controlling for variables related to the population's race, age, and sex (see control variables listed in Table 22). Preliminary analyses were conducted to assess the assumptions of normality, linearity, homoscedasticity, multicollinearity, independence of residuals, and outliers. No violations were noted (see Table 20). The multiple linear regression analysis results revealed that at the capital level, a non-significant regression equation was found F(7, 46)=1.407, *p*=0.226), with an R^2 of 0.176. Table 22 summarizes the regression results performed at the capital level showing that employment is not a significant predictor of COVID-19 mortality (*p*=0.088) while controlling for race, age, and sex.

Table 21

SPSS Output: Regression Analysis of Employment and COVID-19 Mortality at State-level

		Coefficients						
		Unstandardized Coeff.		Standardized Coeff.			95% CI for B	
Model		*B*	*SE*	β	t	*p*	*LL*	*UL*
1	(Constant)	-379	817		-0.463	0.645	-2024	1267
	% Male population[2]	8.792	10.680	0.150	0.823	0.415	-12.719	30.302
	% Population between 0-14 years[2]	8.136	5.522	0.416	1.473	0.148	-2.986	19.258
	% Population between 15-49 years[2]	-2.288	5.919	-0.078	-0.387	0.701	-14.209	9.633
	% White population[2]	-439	413	-1.350	-1.063	0.293	-1270	393
	% African origin population[2]	-201	494	-0.099	-0.408	0.685	-1196	793
	% Mixed-race population[2]	-342	429	-0.940	-0.798	0.429	-1205	521
	Employment (%)[1]	3.055	1.294	0.638	2.360	0.023	0.448	5.662
	% Immunized Population[2]	3.343	0.812	0.680	4.119	<0.001	1.708	4.977

Note. [1.] Predictor variable [2.] Control variables.

Regression model using the total number of deaths from COVID-19 per 100,000 people at the state-level in Brazil.

Table 22

SPSS Output: Regression Analysis of Employment and COVID-19 Mortality at Capital-level

Coefficients

Model	Unstandardized Coeff.		Standardized Coeff.			95% Cl for B	
	B	*SE*	β	t	*p*	*LL*	*UL*
1 (Constant)	588	1962		0.300	0.766	-3360	4537
% of Male Population[2]	14.131	14.316	0.216	0.987	0.329	-14.685	42.947
% of Population between 0-14 Years Old[2]	-6.602	9.234	-0.191	-0.715	0.478	-25.189	11.985
% of Population between 15-49 Years Old[2]	-17.016	8.746	-0.491	-1.945	0.058	-34.622	0.590
% of White Population[2]	-491	1564	-0.901	-0.314	0.755	-3640	2657
% of African Origin Population[2]	-532	1613	-0.258	-0.330	0.743	-3779	2715
% of Mixed-Race Population[2]	-143	1575	-0.241	-0.091	0.928	-3313	3028
Employment (%)[1]	6.657	3.819	0.280	1.743	0.088	-1.031	14.344

Note. [1.] Predictor variable [2.] Control variables.

Data on immunization rate was not available at capital-level; therefore, it was not included in this regression.

Regression model using the total number of deaths from COVID-19 per 100,000 people at the capital-level in Brazil.

Distribution of COVID-19 Death

Between March 2020 and December 2021, the Secretaria de Vigilância em Saúde [Health Surveillance Secretariat], created by the Ministério da Saúde [Ministry of Health], reported 634,695 deaths by COVID-19, being 212,706 deaths in 2020 and 421 989 deaths in 2021. The number of deaths caused by COVID-19 during these first two years of the pandemic corresponds to 18.8% of all deaths in the country in the same period, and the COVID-19 mortality rate observed was 297.5 deaths per 100,000 people in all age groups. The distribution of the casualties according to age groups, sex, race, and Brazilian region are shown in Table 23 – Table 24 and Figure 33 – Figure 40. According to the distribution, in 2020, the number of deaths across the country was higher among individuals who were 80 years or older (1348.3 deaths/100,000), White (112.3 deaths/100,000), and male individuals (117.5 deaths/100,000). Similarly, in 2021 the number of deaths was still higher among individuals who were 80 years or older (1731 deaths/100,000), White (238.5 deaths/100,000), and male individuals (224.8 deaths/100,000). In 2020 death rate was higher in the Southeast region (109.9 deaths/100,000), conversely, in 2021, the death rate was higher in the South (244.6 deaths/100,000).

Note that although individual-based data was used to show death distribution throughout the first years of the pandemic according to race, age group, sex, and geographic region, this is a population-based study and the data in this section was not used to test the study hypothesis or answer the research question. This death distribution evaluation has been for information only.

Table 23

Distribution of COVID-19 Death in 2020 and 2021 according to Age Group

Demographic	[1]*N* (%) in 2020	[2] Death rate per 100,000	[1]*N* (%) in 2021	[2] Death rate per 100,000
Age				
Infant	278 (0.1)	-	347 (0.1)	-
0-4 years	168 (0.1)	1.1	212 (0.1)	1.4
5-9 years	101 (0.0)	0.7	121 (0.0)	0.8
10-14 years	128 (0.1)	0.9	171 (0.0)	1.2
15-19 years	347 (0.2)	2.2	545 (0.1)	3.5
20-29 years	1809 (0.9)	5.3	5517 (1.3)	16.2
30-39 years	5975 (2.8)	17.5	21155 (5.0)	61.7
40-49 years	13542 (6.4)	46.3	45392 (10.8)	152.0
50-59 years	26600 (12.5)	111.4	76859 (18.2)	317.1
60-69 years	47875 (22.5)	286.1	98362 (23.3)	568.7
79-79 years	55979 (26.3)	620.4	93347 (22.1)	991.3
80-older	59880 (28.2)	1348.3	79925 (18.9)	1731.0
Unknown	24 (0.0)	-	36 (0.0)	-
Total	212706 (100.0)	100.3	421989 (100.0)	197.6

Note. [1]*N* = number of deaths by COVID-19. Data on death were accessed from Secretaria de Vigilância em Saúde [Health Surveillance Secretariat] on 18Sep2022, using indicator ICD-10-CM code B34.2 for Coronavirus infection, unspecified. Data on total population for each age group were accessed from Ministério da Saúde [Ministry of Heath] on 18Sep2022. [2] Death rate per 100,000 was calculated by diving the number of deaths per the total population, times 100,000 (data for infant and unknown were excluded from the calculation).

Table 24

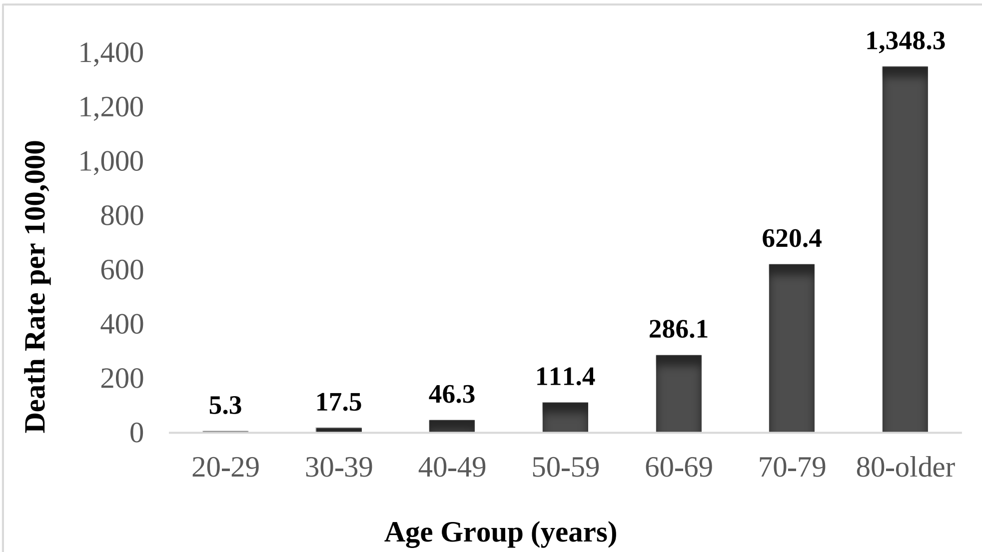
Distribution of COVID-19 Death in 2020 and 2021 according to Sex, Race, and Region

Demographic	[1]*N* (%) in 2020	[2] Death rate per 100,000	[1]*N* (%) in 2021	[2] Death rate per 100,000
Sex				
Male	121641 (57.2)	117.5	234352 (55.5)	224.8
Female	91051 (42.8)	84.1	187588 (44.5)	172.0
Unknown	14 (0.0)	-	49 (0.0)	-
Race				
White	103525 (48.7)	112.3	218852 (51.9)	238.5
African origin	18604 (8.7)	99.7	32149 (7.6)	154.1
Asian origin	1458 (0.7)	-	2591 (0.6)	-
Mixed-race	81572 (38.3)	82.7	139288 (33.0)	142.0
Brazilian Native	972 (0.5)	-	851 (0.2)	-
Unknown	6575 (3.1)	-	28258 (6.7)	-
Region				
Central-West	17381 (8.2)	105.3	39559 (9.4)	236.8
Northeast	54697 (25.7)	95.3	76020 (18.0)	131.8
North	19531 (9.2)	104.6	29861 (7.1)	157.9
Southeast	97798 (46.0)	109.9	202174 (47.9)	225.6
South	23299 (11.0)	77.2	74375 (17.6)	244.6
Total	212706 (100.0)	100.4	421989 (100.0)	197.8

Note. [1]*N* = number of deaths by COVID-19. Data on death were accessed from Secretaria de Vigilância em Saúde [Health Surveillance Secretariat] on 18Sep2022, using indicator ICD-10-CM code B34.2 for Coronavirus infection, unspecified. Data on total population for each sex and region groups were accessed from Ministério da Saúde [Ministry of Heath] on 18Sep2022. Data on total population for each race group were accessed from the Sistema IBGE de Recuperação Automática [IBGE Automatic Recovery System] on 18Sep2022. [2] Death rate per 100,000 was calculated by diving the number of deaths per the total population, times 100,000 (data for unknown, Asian, and Brazilian native population were excluded from the calculation).

Figure 33

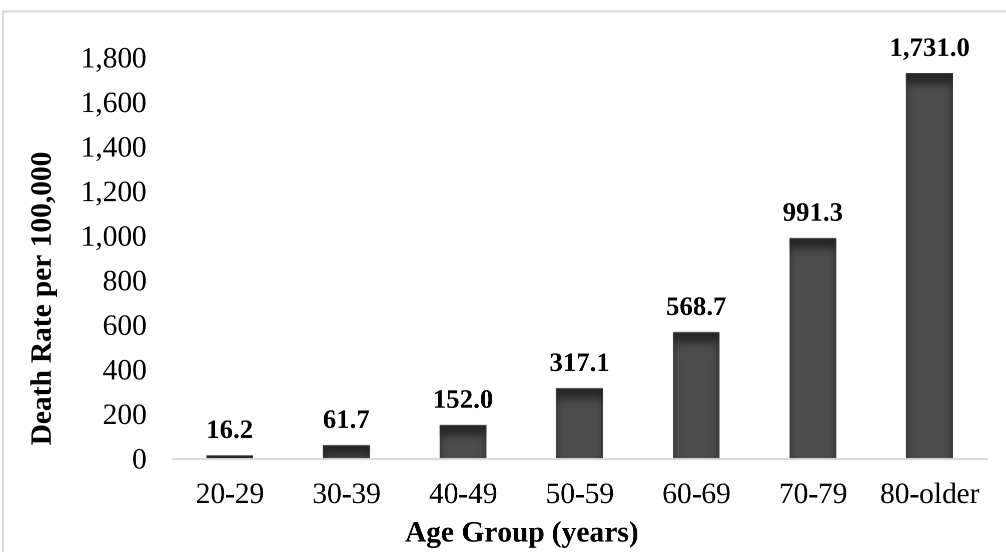
Bar Chart of COVID-19 Death Rate per 100,000 in 2020 by Age Group in Brazil



Note. Only rates > 5 are displayed.

Figure 34

Bar Chart of COVID-19 Death Rate per 100,000 in 2021 by Age Group in Brazil



Note. Only rates > 5 are displayed.

Figure 35

Pie Chart of COVID-19 Death Rate per 100,000 in 2020 by Sex in Brazil

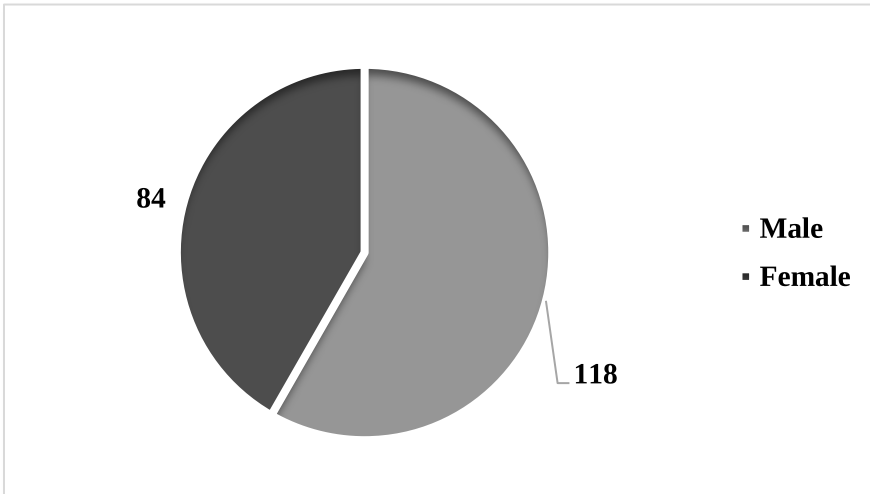


Figure 36

Pie Chart COVID-19 Death Rate per 100,000 in 2021 by Sex in Brazil

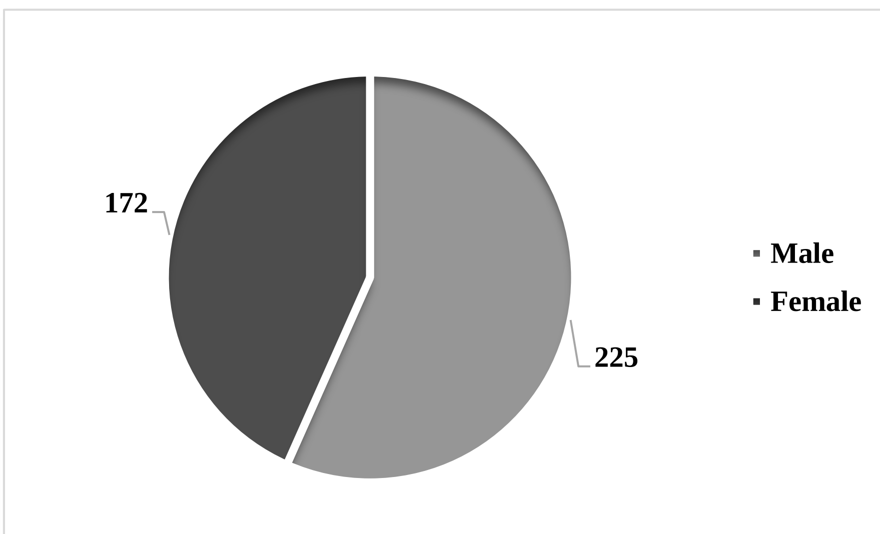


Figure 37

Bar Chart of COVID-19 Death Rate per 100,000 in 2020 by Race in Brazil

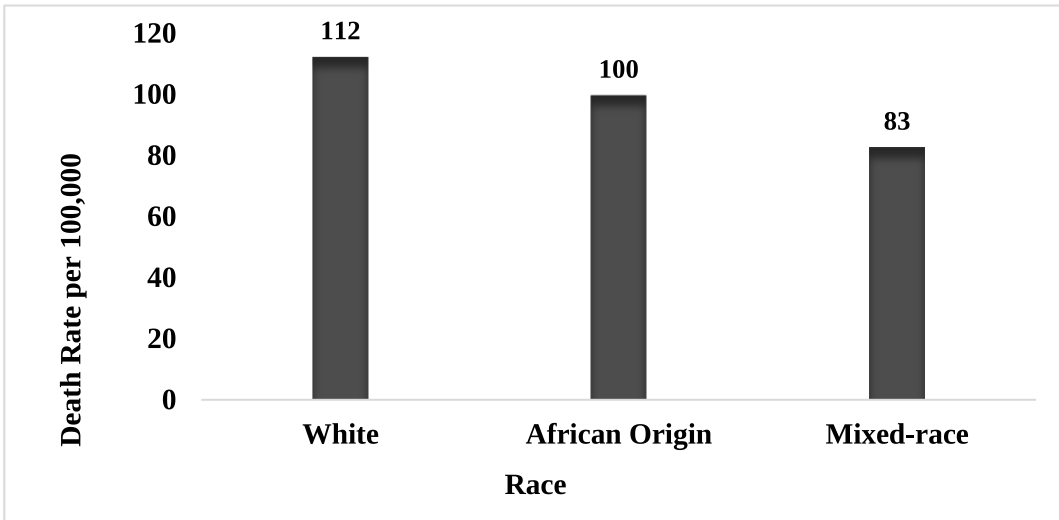


Figure 38

Bar Chart of COVID-19 Death Rate per 100,000 in 2021 by Race in Brazil

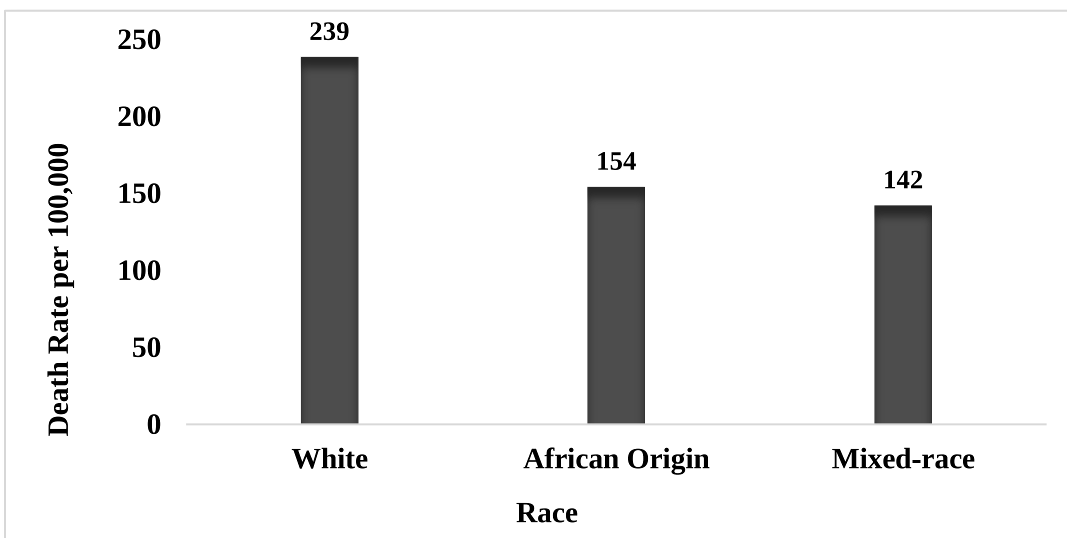


Figure 39

Bar Chart of COVID-19 Death Rate per 100,000 in 2020 by Region

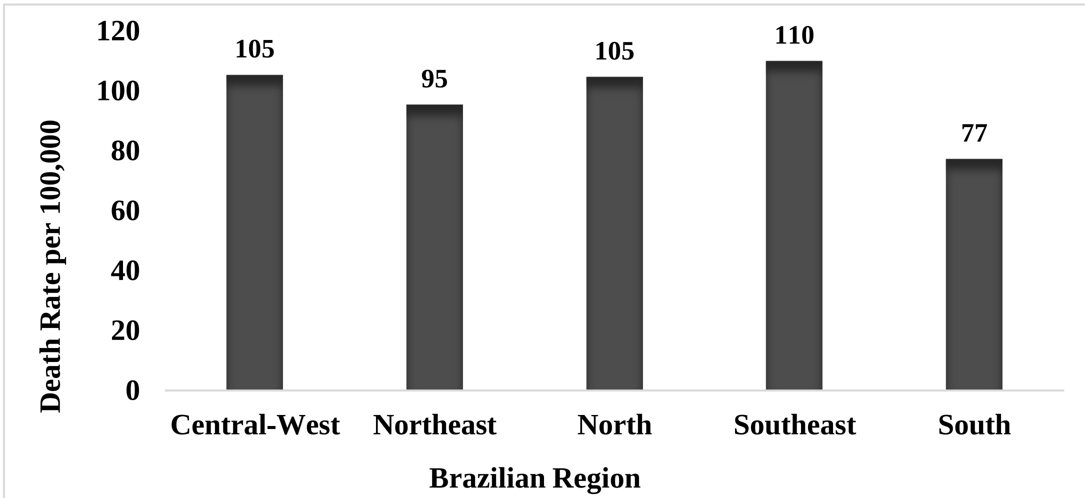
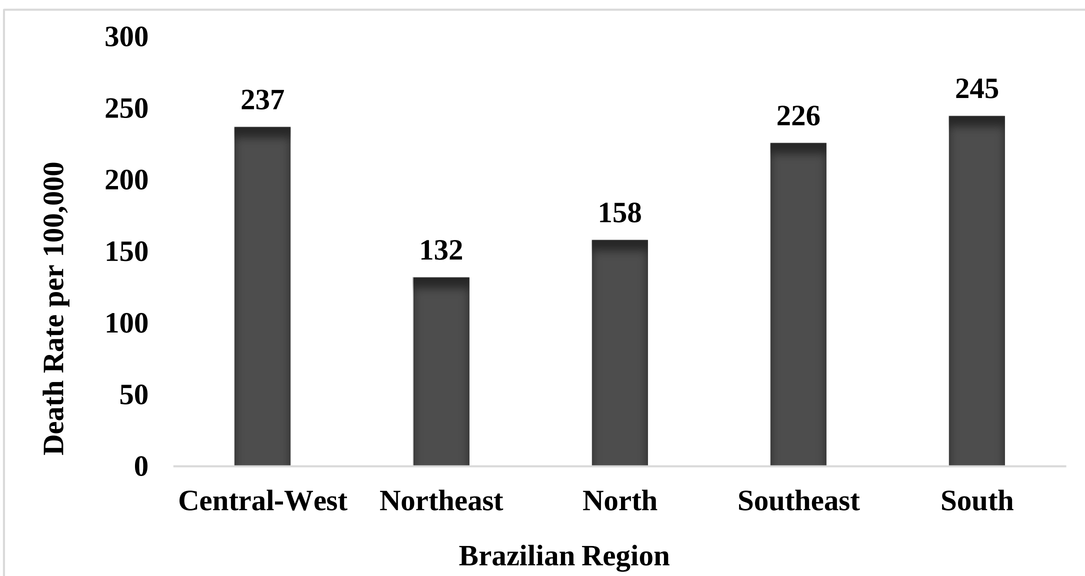


Figure 40

Bar Chart of COVID-19 Death Rate per 100,000 in 2021 by Region



Discussion

Based on the fundamental causes theory, disease spread and mortality are greater among individuals with fewer resources (Link et al., 1995). Therefore, mortality due to COVID-19 is expected to be higher among individuals who are socioeconomic disadvantaged or part of minority groups (Mello, 2020; Ribeiro et al., 2021). During a multiple linear regression analysis to investigate the association between employment and COVID-19 mortality across Brazilian states while controlling for race, age, sex, and immunization rate, it was found that employment is a significant predictor of COVID-19 mortality. Conversely, no statistically significant association was found between employment and COVID-19 mortality rate at the capital level while controlling for race, age, sex, and immunization rate. Results at the regional level were not reported, as the model was found to be invalid in predicting COVID-19 mortality from income at this level, probably due to the small sample size.

Looking at the distribution of COVID-19 mortality in Brazil during the first two years of the pandemic, it was found that the COVID-19 death rate was higher in the Southeast region in 2020 and higher in the Southern region in 2021, the most densely populated regions in the country. Based on this observation, we can suggest that the irregularity or heterogeneity of the population's distribution can shape disease spread, as investigated by Martins-Filho et al. (2021) and Miranda et al. (2020). The same analysis found that the COVID-19 death rate was higher among individuals 80 years old or older, male and White, in both years of the pandemic.

As mentioned in the Summary of the Manuscripts section, most of the study limitations were associated with data availability. Two-year data for disease mortality were obtained for each region, state, and capital. Although the data used for the regression analyses were population-based, the data obtained at a national level was converted to an individual-level dataset to investigate death distribution for the first two years of the pandemic. To avoid errors during the manipulation of this large dataset, the data was reviewed a second time to ensure transcription errors were not present. Another limitation was regarding data availability. Since COVID-19 is an ongoing pandemic, many data points were unavailable or ready for analyses, resulting in a need for some manipulation to guarantee dataset completion and readiness before analyses. Data manipulation was required to avoid internal validity issues, but data should be considered reliable as all information was retrieved from governmental resources.

The findings of this study allow public health and political leaders to identify geographic areas in the country that need special attention regarding the development of policies and implementation of public health programs to encourage populations' best behavior and practices concerning preventive actions and vaccination willingness. Furthermore, considering Brazil a country of high economic and demographic disparities, the social change implications of this study may include a potential increase in people's opportunities for socioeconomic equality, particularly a fair income distribution, which could allow individuals to have better access to healthcare, consequently decreasing COVID-19 mortality rates. A decrease in the COVID-19 mortality rate could allow Brazilians to improve individual lives with healthier families and communities.

Based on this study's findings and to better explore the influences of SES as a risk factor for COVID-19 mortality, further research is recommended with both individual and population-based approaches. Since COVID-19 is an ongoing pandemic, it is more likely that more accurate results will be only obtained with post-pandemic data. Lastly, multiple linear regression models may not be the best model for analyzing population-based data considering that sample sizes may be limited compared to individual-based data.

References

Aschengrau, A., & Seage, G. R. (2013). *Essentials of epidemiology in public health*. Jones & Bartlett Publishers.

Barreto, N. M. P. V., Rios, J. D. C., Ribeiro, E. B., Torreão, P. L., Oliveira, S. A. L., Teixeira, J. R. B., & Santos, N. Q. (2021). Vulnerabilidades sociais relacionadas à infecção e mortalidade por COVID-19: Uma revisão sistemática. *Revista de Saúde Coletiva da UEFS, 11*(2), e6039-e6039. https://doi.org/10.13102/rscdauefs.v11i2.6039

Barreto, M. L., Teixeira, M. G., Bastos, F. I., Ximenes, R. A. A., & Barata, R. B. (2014). Sucessos e Fracassos no Controle de Doenças Infecciosas no Brasil: O Contexto Social e Ambiental, Polıticas, Intervençoes e Necessidades de Pesquisa. *Lancet. Saúde no Brasil 3,* 47–60. https://doi.org/10.1016/S0140-6736(11)60202-X

Chen, Y. Y., & Assefa, Y. (2021). The heterogeneity of the COVID-19 pandemic and national responses: An explanatory mixed-methods study. *BioMed Central Public Health*, *21*(1), 1-15. https://doi.org/10.1186/S12889-021-10885-8

Figueiredo, A. M. D., Figueiredo, D. C. M. M. D., Gomes, L. B., Massuda, A., Gil-García, E., Vianna, R. P. D. T., & Daponte, A. (2020). Social determinants of health and COVID-19 infection in Brazil: An analysis of the pandemic. *Revista Brasileira de Enfermagem, 73* (1-7). http://dx.doi.org/10.1590/0034-7167-2020-0673

Garnier, R., Benetka, J. R., Kraemer, J., & Bansal, S. (2021). Socioeconomic disparities in social distancing during the COVID-19 pandemic in the United States:

Observational study. *Journal of Medical Internet Research*, *23*(1), e24591. http://dx.doi.org/10.2196/24591

Instituto Brasileiro de Geografia e Estatistica [IBGE] (2021). Estudos e Pesquisas. Informacao Demografica e Socioeconomica. Retrieved from https://biblioteca.ibge.gov.br/visualizacao/livros/liv101892.pdf

Instituto Brasileiro de Geografia e Estatistica [IBGE] (n.d.). PNAD Contínua - Pesquisa Nacional por Amostra de Domicílios Contínua. Retrieved from https://www.ibge.gov.br/estatisticas/sociais/trabalho/9173-pesquisa-nacional-por-amostra-de-domicilios-continua-trimestral.html?=&t=series-historicas&utm_source=landing&utm_medium=explica&utm_campaign=desemprego

Martins-Filho, P. R. (2021). Relationship between population density and COVID-19 incidence and mortality estimates: A county-level analysis. *Journal of Infection and Public Health*, *14*(8), 1087. https://doi.org/10.1016/j.jiph.2021.06.018

National Council of Health Secretaries – CONASS (2021). Painel Nacional: COVID-19. Retrieved from https://www.conass.org.br/painelconasscovid19/

Natividade, M. D. S., Bernardes, K., Pereira, M., Miranda, S. S., Bertoldo, J., Teixeira, M. D. G., & Aragão, E. (2020). Social distancing and living conditions in the pandemic COVID-19 in Salvador-Bahia, Brazil. *Ciência & Saúde Coletiva*, *25*, 3385-3392. https://doi.org/10.1590/1413-81232020259.22142020

Oliveira, M., Wong, J., Lippi, G., & Henry, B. M. (2020). Analysis of clinical and demographic heterogeneity of patients dying from COVID-19 in Brazil versus

China and Italy. *The Brazilian Journal of Infectious Diseases*, *24*(3), 273–275. https://doi.org/10.1016/j.bjid.2020.05.002

Waldman, E. A., & Sato, A. P. S. (2016). Trajetória das doenças infecciosas no Brasil nos últimos 50 anos: um contínuo desafio. *Revista de Saúde Pública*, 50, 1-18. https://doi.org/10.1590/S1518-8787.2016050000232

Warner, R. M. (2020). *Applied statistics: From bivariate through multivariate techniques*. Sage.

World Health Organization [WHO] (2020). WHO Director-General's opening remarks at the media briefing on COVI-19 – 11 March 2020. Retrieved from https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-COVID-19---11-march-2020

Part 3: Summary

First Heading

The findings of the three studies work in conjunction to inform how main socioeconomic elements, such as income, education, and employment can work as risk factors that affect COVID-19 mortality in Brazil at different geographic levels. Depending on the geographic level observed, some of the findings do not agree with the theory tested, the fundamental causes theory (Link et al., 1995), which can be explained by a series of other factors, including population density and heterogeneity; shows there are still inconsistencies on how minority groups are at greater risk of disease contamination and death.

During data retrieval for these studies, I encountered unanticipated events regarding data availability and sample size determination, which caused me to deviate from the proposed study. During the proposal stage, the sample size was inadvertently determined based on an individual-based study. However, all data for the dependent, independent, and control variables were aggregated per geographic location at region, state, and capital levels. The total number of aggregated data points accessed and used to create the study dataset was N=118; however, to avoid violation of internal validity during our population-based study, the regression could not be performed using all the combined 118 data points; instead, the data needed to be analyzed at each geographic level of the population (region, state, and capital levels), which shortened the sample size during analyses.

Even with the unanticipated events described, I believe these studies may have positive social change implications, which include a potential increase in people's opportunities for socioeconomic equality and, consequently, a decrease in the COVID-19 mortality rate. This would allow Brazilians to improve individual lives with healthier families and communities. However, since the pandemic is still an ongoing issue, further investigations are needed to obtain post-pandemic data for more accurate information to examine the role of SES in predicting COVID-19 mortality. Therefore, future research is needed to investigate the impact of socioeconomic factors and Brazilian population demographic characteristics on disease spread and mortality.

The overall experience of conducting this research was enriching and allowed me to learn more about the importance of fighting for socioeconomic equality to facilitate families' access to healthcare, which can be particularly scarce during a global pandemic. Political and public health leaders of developing countries, such as Brazil, must constantly remember that improvements in access to education and employment can positively affect the countries' public health. Hopefully, future studies will focus more on /developing countries for better public health preparedness, thus decreasing the chances of disease growth and death that burden disadvantaged populations.

Consolidated References

Aschengrau, A., & Seage, G. R. (2013). *Essentials of epidemiology in public health*. Jones & Bartlett Publishers.

Achdut, N., & Refaeli, T. (2020). Unemployment and psychological distress among young people during the COVID-19 pandemic: psychological resources and risk factors. *International Journal of Environmental Research and Public Health, 17*(19), 7163. https://doi.org/10.3390/ijerph17197163

American Psychological Association [APA] (2022). Socioeconomic Status. Retrieved from https://www.apa.org/topics/socioeconomic-status#:~:text=Socioeconomic%20status%20is%20the%20social,to%20privilege%2C%20power%20and%20control

Bambra, C., Riordan, R., Ford, J., & Matthews, F. (2020). The COVID-19 pandemic and health inequalities. Journal of Epidemiology and Community Health, *74*(11), 964-968. https://doi.org/10.1136/jech-2020-214401

Baqui, P., Marra, V., Alaa, A. M., Bica, I., Ercole, A., & van der Schaar, M. (2021). Comparing COVID-19 risk factors in Brazil using machine learning: the importance of socioeconomic, demographic and structural factors. *Scientific Reports, 11*(1), 1–10. https://doi.org/10.1038/s41598-021-95004-8.

Barberia, L. G., & Gómez, E. J. (2020). Political and institutional perils of Brazil's COVID-19 crisis. *Lancet (London, England)*, *396*(10248), 367. https://doi.org/10.1016/S0140-6736(20)31681-0

Barreto, M. L., Teixeira, M. G., Bastos, F. I., Ximenes, R. A., Barata, R. B., &

Rodrigues, L. C. (2011). Successes and failures in the control of infectious diseases in Brazil: social and environmental context, policies, interventions, and research needs. *The Lancet, 377*(9780), 1877-1889. https://doi.org/10.1016/S0140-6736(11)60202-X

Barreto, N. M. P. V., Rios, J. D. C., Ribeiro, E. B., Torreão, P. L., Oliveira, S. A. L., Teixeira, J. R. B., & Santos, N. Q. (2021). Vulnerabilidades sociais relacionadas à infecção e mortalidade por COVID-19: Uma revisão sistemática. *Revista de Saúde Coletiva da UEFS, 11*(2), e6039-e6039. https://doi.org/10.13102/rscdauefs.v11i2.6039

Bartholo, T. L., Koslinski, M. C., de Andrade, F. M., & de Castro, D. L. (2020). Segregação escolar e desigualdades educacionais no início da escolarização no Brasil. REICE. *Revista Iberoamericana sobre Calidad*, Eficacia y Cambio en Educación, *18*(4), 57-76. https://doi.org/10.15366/reice2020.18.4.003

Bermudi, P. M. M., Lorenz, C., Aguiar, B. S. de, Failla, M. A., Barrozo, L. V., & Chiaravalloti-Neto, F. (2021). Spatiotemporal ecological study of COVID-19 mortality in the city of São Paulo, Brazil: Shifting of the high mortality risk from areas with the best to those with the worst socio-economic conditions. *Travel Medicine & Infectious Disease, 39*, 1-8. N.PAG. https://doi.org/10.1016/j.tmaid.2020.101945

Bravata, D. M., Perkins, A. J., Myers, L. J., Arling, G., Zhang, Y., Zillich, A. J., Reese, L., Dysangco, A., Agarwal, R., Myers, J., Austin, C., Sexson, A., Leonard, S. J., Dev, S., & Keyhani, S. (2021). Association of intensive care unit patient load and demand with mortality rates in US Department of Veterans Affairs hospitals during the COVID-19 pandemic. *Journal of the American Medical Association Network Open*, *4*(1), e2034266-e2034266. https://doi.org/10.1001/jamanetworkopen.2020.34266

Brazilian Institute of Geography and Statistics [IBGE] (2022). Annual average unemployment rate decreases in most states in 2021. https://agenciadenoticias.ibge.gov.br/en/agencia-news/2184-news-agency/news/33054-annual-average-unemployment-rate-decreases-in-most-states-in-2021

Brazilian Institute of Geography and Statistics [IBGE] (n.d.). https://www.ibge.gov.br/estatisticas/sociais/educacao/9221-sintese-de-indicadores-sociais.html?edicao=32373&t=downloads.

Brazilian Institution of Geography and Statistics [IBGE] (2019). Conheça o Brasil – População – Educação. Retrieved from https://educa.ibge.gov.br/jovens/conheca-o-brasil/populacao/18317-educacao.html

Centers for Disease Control and Prevention (n.d.-a). Guidance for Businesses and Employers Responding to Coronavirus Disease 2019 (COVID-19). Retrieved from https://www.cdc.gov/coronavirus/2019-ncov/community/guidance-business-response.html

Chen, Y. Y., & Assefa, Y. (2021). The heterogeneity of the COVID-19 pandemic and national responses: An explanatory mixed-methods study. *BioMed Central Public Health*, *21*(1), 1-15. https://doi.org/10.1186/S12889-021-10885-8

Cifuentes, M. P., Rodriguez-Villamizar, L. A., Rojas-Botero, M. L., Alvarez-Moreno, C. A., & Fernández-Niño, J. A. (2021). Socioeconomic inequalities associated with mortality for COVID-19 in Colombia: A cohort nationwide study. *Journal of Epidemiology and Community Health, 75*(7), 610-615. https://doi.org/10.1136/jech-2020-216275

Chang, V. W., & Lauderdale, D. S. (2009). Fundamental Causes theory, technological innovation, and health disparities: the case of cholesterol in the era of statins. *Journal of Health and Social Behavior, 50*(3), 245-260. https://doi.org/10.1177/002214650905000301

Chowell, G., Bettencourt, L. M., Johnson, N., Alonso, W. J., & Viboud, C. (2008). The 1918–1919 influenza pandemic in England and Wales: spatial patterns in transmissibility and mortality impact. *Proceedings of the Royal Society B: Biological Sciences, 275*(1634), 501-509. https://doi.org/10.1098/rspb.2007.1477

Clouston, S. A. P., Natale, G., & Link, B. G. (2021). Socioeconomic inequalities in the spread of coronavirus-19 in the United States: A examination of the emergence of social inequalities. *Social Science & Medicine, 268*, 1-6. https://doi.org/10.1016/j.socscimed.2020.113554

Cury, C. R. J. (2008). A educação escolar, a exclusão e seus destinatários. *Educação em Revista*, 205-222. https://doi.org/10.1590/S0102-46982008000200010

De Paula, A. M. N. R., Coelho, D., de Oliveira, C. L., & Moura, A. P. V. (2021). Should the convergence between the social and biological vulnerabilities of black communities and the COVID-19 pandemic be considered? *Medical Hypotheses,* *146*, 1-2, 110359. https://doi.org/10.1016/j.mehy.2020.110359

De Souza, C. D. F., Machado, M. F., & do Carmo, R. F. (2020). Human development, social vulnerability and COVID-19 in Brazil: A study of the social determinants of health. *Infectious Diseases of Poverty*, *9*(04), 50-59. https://doi.org/10.1186/s40249-020-00743-x

Díaz-Castro, L., Cabello-Rangel, H., & Hoffman, K. (2021). The impact of health policies and sociodemographic factors on doubling time of the COVID-19 pandemic in Mexico. *International Journal of Environmental Research and Public Health*, *18*(5), 2354. https://doi.org/10.3390/ijerph18052354

Domingues, C., Maranhão, A., Teixeira, A. M., Fantinato, F., & Domingues, R. (2020). The Brazilian National Immunization Program: 46 years of achievements and challenges. 46 anos do Programa Nacional de Imunizações: Uma história repleta de conquistas e desafios a serem superados. *Cadernos de Saude Publica*, *36*Suppl 2(Suppl 2), e00222919. https://doi.org/10.1590/0102-311X00222919

Figueiredo, A. M. D., Figueiredo, D. C. M. M. D., Gomes, L. B., Massuda, A., Gil-García, E., Vianna, R. P. D. T., & Daponte, A. (2020). Social determinants of health and COVID-19 infection in Brazil: An analysis of the pandemic. *Revista Brasileira de Enfermagem*, *73*(1-7). http://dx.doi.org/10.1590/0034-7167-2020-0673

French, G., Hulse, M., Nguyen, D., Sobotka, K., Webster, K., Corman, J., Aboagye-Nyame, B., Dion, M., Johnson, M., Zalinger, B., & Ewing, M. (2021). Impact of hospital strain on excess deaths during the COVID-19 pandemic - United States, July 2020–July 2021. *Morbidity and Mortality Weekly Report, 70*(46), 1613. https://doi.org/10.15585/mmwr.mm7046a5

Garnier, R., Benetka, J. R., Kraemer, J., & Bansal, S. (2021). Socioeconomic disparities in social distancing during the COVID-19 pandemic in the United States: Observational study. *Journal of Medical Internet Research*, *23*(1), e24591. http://dx.doi.org/10.2196/24591

Goutte, S., Péran, T., & Porcher, T. (2020). The role of economic structural factors in determining pandemic mortality rates: Evidence from the COVID-19 outbreak in France. *Research in International Business and Finance, 54*, 101281. https://doi.org/10.1016/j.ribaf.2020.101281

Governo Brasileiro (1961). Presidencia da República. Casa Civil Subchefia para Assuntos Juridicos. Retrieved from http://www.planalto.gov.br/ccivil_03/leis/L4024.htm

Grantz, K. H., Rane, M. S., Salje, H., Glass, G. E., Schachterle, S. E., & Cummings, D. A. (2016). Disparities in influenza mortality and transmission related to sociodemographic factors within Chicago in the pandemic of 1918. *Proceedings of the National Academy of Sciences*, *113*(48), 13839-13844. https://doi.org/10.1073/pnas.1612838113

Greer, S. L., King, E. J., da Fonseca, E. M., & Peralta-Santos, A. (2020). The comparative politics of COVID-19: The need to understand government responses. *Global public health*, *15*(9), 1413-1416. https://doi.org/10.1080/17441692.1783340

Instituto Brasileiro de Geografia e Estatistica [IBGE] (2021). Estudos e Pesquisas. Informacao Demografica e Socioeconomica. Retrieved from https://biblioteca.ibge.gov.br/visualizacao/livros/liv101892.pdf

Instituto Brasileiro de Geografia e Estatistica [IBGE] (n.d.). PNAD Contínua - Pesquisa Nacional por Amostra de Domicílios Contínua. Retrieved from https://www.ibge.gov.br/estatisticas/sociais/trabalho/9173-pesquisa-nacional-por-amostra-de-domicilios-continua-trimestral.html?=&t=series-historicas&utm_source=landing&utm_medium=explica&utm_campaign=desemprego

Khanijahani, A., Iezadi, S., Gholipour, K., Azami-Aghdash, S., & Naghibi, D. (2021). A systematic review of racial/ethnic and socioeconomic disparities in COVID-19. *International Journal for Equity in Health, 20*(1), 1-30. https://doi.org/10.1186/s12939-021-01582-4

Khalatbari-Soltani, S., Cumming, R. C., Delpierre, C., & Kelly-Irving, M. (2020). Importance of collecting data on socioeconomic determinants from the early stage of the COVID-19 outbreak onwards. *Journal of Epidemiology and Community Health*, *74*(8), 620-623. http://dx.doi.org/10.1136/jech-2020-214297

Link, B. G., & Phelan, J. (1995). Social Conditions As Fundamental Causes of Disease. *Journal of Health and Social Behavior*, 80–94. https://doi.org/10.2307/2626958

Lotta, G., Fernandez, M., Kuhlmann, E., & Wenham, C. (2022). COVID-19 vaccination challenge: what have we learned from the Brazilian process? *The Lancet Global Health*. 1-2. https://doi.org/10.1016/S2214-109X(22)00049-3

Luptáčik, M., & Nežinský, E. (2020). Measuring income inequalities beyond the Gini coefficient. *Central European Journal of Operations Research*, *28*(2), 561-578. https://doi.org/10.1007/s10100-019-00662-9

Mackenbach, J. P., Kulhánová, I., Bopp, M., Deboosere, P., Eikemo, T. A., Hoffmann, R., Kulik, M. C., Leinsalu, M., Martikainen, P., Menvielle, G., Regidor, E., Wojtyniak, B., Östergren, O., & EURO-GBD-SE Consortium. (2015). Variations in the relation between education and cause-specific mortality in 19 European populations: A test of the "Fundamental Causes" theory of social inequalities in health. *Social Science & Medicine*, *127*, 51-62. https://doi.org/10.1016/j.socscimed.2014.05.021

Mackenbach, J. P., Kunst, A. E., Cavelaars, A. E., Groenhof, F., Geurts, J. J., & EU Working Group on Socioeconomic Inequalities in Health. (1997). Socioeconomic inequalities in morbidity and mortality in western Europe. *The Lancet*, *349*(9066), 1655-1659. https://doi.org/10.1016/S0140-6736(96)07226-1

Mackenbach, J. P., Looman, C. W. N., Artnik, B., Bopp, M., Deboosere, P., Dibben, C., Kalediene, R., Kovács, K., Leinsalu, M., Martikainen, P., Regidor, E., Rychtaříková, J., & de Gelder, R. (2017). "Fundamental Causes" of inequalities in mortality: an empirical test of the theory in 20 European populations. *Sociology of Health & Illness*, *39*(7), 1117–1133. https://doi.org/10.1111/1467-9566.12562

Mackenbach, J. P., Stirbu, I., Roskam, A. J. R., Schaap, M. M., Menvielle, G., Leinsalu, M., & Kunst, A. E. (2008). Socioeconomic inequalities in health in 22 European countries. *New England Journal of Medicine*, *358*(23), 2468-2481.

https://doi.org/10.1056/NEJMsa0707519

Mackenbach, J. P., Bos, V., Andersen, O., Cardano, M., Costa, G., Harding, S., Reid, A., Hemström, O., Valkonen, & Kunst, A. E. (2003). Widening socioeconomic inequalities in mortality in six Western European countries. *International Journal of Epidemiology, 32*(5), 830-837. https://doi.org/10.1093/ije/dyg209

Martinez-Valle, A. (2021). Public health matters: Why is Latin America struggling in addressing the pandemic? *Journal of Public Health Policy, 42*(1), 27-40. https://doi.org/10.1057/s41271-020- 00269-4

Martins-Filho, P. R. (2021). Relationship between population density and COVID-19 incidence and mortality estimates: A county-level analysis. *Journal of Infection and Public Health, 14*(8), 1087. https://doi.org/10.1016/j.jiph.2021.06.018

Martins-Filho, P. R., de Souza Araújo, A. A., Quintans-Júnior, L. J., & Santos, V. S. (2020). COVID-19 fatality rates related to social inequality in Northeast Brazil: a neighborhood-level analysis. *Journal of Travel Medicine, 27*(7), 1-3. https://doi.org/10.1093/jtm/taaa128.

Martins-Filho, P. R., Quintans-Júnior, L. J., de Souza Araújo, A. A., Sposato, K. B., Souza Tavares, C. S., Gurgel, R. Q., Fontes Leite, D. C., de Paiva, S. M., Santos, H. P., & Santos, V. S. (2021). Socio-economic inequalities and COVID-19 incidence and mortality in Brazilian children: a nationwide register-based study. *Public Health (Elsevier)*, 190, 4–6. https://doi.org/10.1016/j.puhe.2020.11.005

Mateo-Urdiales, A., Fabiani, M., Rosano, A., Vescio, M. F., Del Manso, M., Bella, A., Riccardo, F., Pezzotti, P., Regidor, E., & Andrianou, X. (2021). Socioeconomic patterns and COVID-19 outcomes before, during, and after the lockdown in Italy (2020). *Health & Place*, 71, 102642. https://doi.org/10.1016/j.healthplace.2021.102642

Mello, D. (05 May 2020). *Risco de morrer por coronavírus varia até 10 vezes entre bairros de SP* [Risk of dying from coronavirus varies up to 10 times between neighborhoods in SP] [Internet]. Agência Brasil [Brazil Agency]. Retrieved from Risco de morrer por coronavírus varia até 10 vezes entre bairros (ebc.com.br)

Mena, G. E., Martinez, P. P., Mahmud, A. S., Marquet, P. A., Buckee, C. O., & Santillana, M. (2021). Socioeconomic status determines COVID-19 incidence and related mortality in Santiago, Chile. *Science*, *372*(6545), eabg5298. https://doi.org/10.1126/science.abg5298

Miech, R., Pampel, F., Kim, J., & Rogers, R. G. (2011). The enduring association between education and mortality: the role of widening and narrowing disparities. *American Sociological Review*, *76*(6), 913-934. https://doi.org/10.1177/0003122411411276

Ministry of Health (n.d.). Sistema Único de Saúde (SUS): Estrutura, princípios e como funciona. https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-z/s/sus-estrutura-principios-e-como-funciona

Miranda Demenech, L., de Carvalho Dumith, S., Centena Duarte Vieira, M. E., & Neiva-Silva, L. (2020). Income inequality and risk of infection and death by COVID-19 in Brazil. *Revista Brasileira de Epidemiologia*, 1–12. https://doi.org/10.1590/1980-549720200095

Moore, D. C. B. C., Nehab, M. F., Camacho, K. G., Reis, A. T., de Fátima Junqueira-Marinho, M., Abramov, D. M., & Correia, R. F. (2021). Low COVID-19 vaccine hesitancy in Brazil. *Vaccine*, *39*(42), 6262-6268. https://doi.org/10.1016/j.vaccine.2021.09.013

Murray, C. J., Lopez, A. D., Chin, B., Feehan, D., & Hill, K. H. (2006). Estimation of potential global pandemic influenza mortality on the basis of vital registry data from the 1918–20 pandemic: a quantitative analysis. *The Lancet*, *368*(9554), 2211-2218. https://doi.org/10.1016/S0140-6736(06)69895-4

Nair, S. C., Gasmelseed, H. I., Khan, A. A., Khafagy, I. N., Sreedharan, J., Saleem, A. A., Abdrhman, H. I., Alhosani, A. H., Siddiqua, A. R., Ahmed, A. R., Shubbar, A. I., Aleissaee, A. R., Alanqar, A. W., Hamadeh, A. M., Safdani, F. A., Habbal, F. W., Choker, W. B., Bashir, K. M., Alblooshi, M. A., Farajallah, M. M., Alzaabi, M. N., ShiL, R. S., Alshehhi, S. S., & Douleh, W. F. (2021). Assessment of mortality from COVID-19 in a multicultural multi-ethnic patient population. *BMC Infectious Diseases*, *21*(1), 1-9. https://doi.org/10.1186/s12879-021-06762-9

National Council of Health Secretaries – CONASS (2021). Painel Nacional: COVID-19. Retrieved from https://www.conass.org.br/painelconasscovid19/

Natividade, M. D. S., Bernardes, K., Pereira, M., Miranda, S. S., Bertoldo, J., Teixeira, M. D. G., & Aragão, E. (2020). Social distancing and living conditions in the pandemic COVID-19 in Salvador-Bahia, Brazil. *Ciência & Saúde Coletiva*, *25*, 3385-3392. https://doi.org/10.1590/1413-81232020259.22142020

Ogedegbe, G., Ravenell, J., Adhikari, S., Butler, M., Cook, T., Francois, F., & Horwitz,

L. I. (2020). Assessment of racial/ethnic disparities in hospitalization and mortality in patients with COVID-19 in New York City. *Journal of the American Medical Association, Network Open, 3*(12), e2026881-e2026881. https://doi.org/10.1001/jamanetworkopen.2020.26881

Oliveira, M., Wong, J., Lippi, G., & Henry, B. M. (2020). Analysis of clinical and demographic heterogeneity of patients dying from COVID-19 in Brazil versus China and Italy. *The Brazilian Journal of Infectious Diseases, 24*(3), 273–275. https://doi.org/10.1016/j.bjid.2020.05.002

Organization for Economic Co-operation and Development [OECD] (2020). OECD Economic Surveys. Brazil. December 2020. Overview. https://www.oecd.org/economy/surveys/Brazil-2020-OECD-economic-survey-overview.pdf

Organization for Economic Co-operation and Development [OECD] (2021). OECDiLibrary. Education at a Glance 2021: OECD Indicators. Brazil. https://www.oecd-ilibrary.org/sites/d8d547f5-en/index.html?itemId=/content/component/d8d547f5-en#section-d12020e24001

Ortega, F., & Orsini, M. (2020). Governing COVID-19 without government in Brazil: Ignorance, neoliberal authoritarianism, and the collapse of public health leadership. *Global public health*, *15*(9), 1257-1277. https://doi.org/10.1080/17441692.2020.1795223

Paim, J., Travassos, C., Almeida, C., Bahia, L., & Macinko, J. (2011). The Brazilian health system: history, advances, and challenges. *The Lancet*, *377*(9779), 1778-

1797. https://doi.org/10.1016/S0140-6736(11)60054-8

Ren, X. S., Amick, B. C., & Williams, D. R. (1999). Racial/ethnic disparities in health: the interplay between discrimination and socioeconomic status. *Ethnicity & Disease*, *9*(2), 151-165.

Rattay, P., Michalski, N., Domanska, O. M., Kaltwasser, A., De Bock, F., Wieler, L. H., & Jordan, S. (2021). Differences in risk perception, knowledge and protective behaviour regarding COVID-19 by education level among women and men in Germany. Results from the COVID-19 Snapshot Monitoring (COSMO) study. *Plos One*, *16*(5), e0251694. https://doi.org/10.1371/journal.pone.0251694

Ribeiro, F., & Leist, A. (2020). Who is going to pay the price of COVID-19? Reflections about an unequal Brazil. *International Journal for Equity in Health*, *19*(1), 1–3. https://doi.org/10.1186/s12939-020-01207-2

Ribeiro, V. M., & Vóvio, C. L. (2017). Desigualdade escolar e vulnerabilidade social no território. Educar em Revista, 71-87. https://doi.org/10.1590/0104-4060.51372

Rocha, R., Atun, A., Massuda, A., Rache, B., Spinola, S., Nunes, L., Lago, M., & Castro, M. C. (2021). Effect of socioeconomic inequalities and vulnerabilities on health-system preparedness and response to COVID-19 in Brazil: A comprehensive analysis. *The Lancet Global Health*, *9*(6), e782–e792. 10.1016/S2214-109X(21)00081-4 https://doi.org/10.1016/S2214-109X(21)00081-4

Secretaria de Vigilância em Saúde [Health Surveillance Secretariat] (2022). Departamento de Análise Epidemiológica e Vigilância de Doenças Não Transmissíveis [Department of Epidemiological Analysis and Surveillance of

Noncommunicable Diseases]. Portal do Governo Brasileiro. https://svs.aids.gov.br/daent/centrais-de-conteudos/paineis-de-monitoramento/mortalidade/cid10/

Shadmi, E., Chen, Y., Dourado, I., Faran-Perach, I., Furler, J., Hangoma, P., Hanvoravongchai1, P., Obando, C., Petrosyan, V., Rao, K. D., Ruano, A. L., Shi, L., Souza, L. E., Spitzer-Shohat, S., Sturgiss, E., Suphanchaimat, R., Uribe, M. V., & Willems, S. (2020). Health equity and COVID-19: global perspectives. *International Journal for Equity in Health, 19*(1), 1-16. https://doi.org/10.1186/s12939-020-01218-z

Shkolnikov, V. M., Andreev, E. M., Jdanov, D. A., Jasilionis, D., Kravdal, Ø., Vågerö, D., & Valkonen, T. (2012). Increasing absolute mortality disparities by education in Finland, Norway and Sweden, 1971-2000. *Journal of Epidemiology and Community Health, 66*(4), 372–378. https://doi.org/10.1136/jech.2009.104786

Stojanovic, J., Boucher, V. G., Gagne, M., Gupta, S., Joyal-Desmarais, K., Paduano, S., & Lavoie, K. L. (2021). Global trends and correlates of COVID-19 vaccination hesitancy: Findings from the iCARE study. *Vaccines, 9*(6), 661. https://doi.org/10.3390/vaccines9060661

Strand, B. H., Grøholt, E. K., Steingrímsdóttir, Ó. A., Blakely, T., Graff-Iversen, S., & Næss, Ø. (2010). Educational inequalities in mortality over four decades in Norway: prospective study of middle-aged men and women followed for cause specific mortality, 1960-2000. *British Medical Journal, 340*, 1-10. https://doi.org/10.1136/bmj.c654

Talic, S., Shah, S., Wild, H., Gasevic, D., Maharaj, A., Ademi, Z., & Ilic, D. (2021). Effectiveness of public health measures in reducing the incidence of COVID-19, SARS-CoV-2 transmission, and COVID-19 mortality: Systematic review and meta-analysis. *British Medical Association*, *375*. https://doi.org/10.1136/bmj-2021-068302

U.S. Bureau of Labor Statistics (n.d.). Labor Force Statistics from the Current Population Survey. https://www.bls.gov/cps/

Waldman, E. A., & Sato, A. P. S. (2016). Trajetória das doenças infecciosas no Brasil nos últimos 50 anos: um contínuo desafio. *Revista de Saúde Pública*, *50*, 1-18. https://doi.org/10.1590/S1518-8787.2016050000232

Warner, R. M. (2020). *Applied statistics: From bivariate through multivariate techniques*. Sage.

World Health Organization [WHO] (2020-a). Archived: WHO Timeline - COVID-19. Retrieved from https://www.who.int/news/item/27-04-2020-who-timeline---COVID-19

World Health Organization [WHO] (2020-b). WHO Director-General's opening remarks at the media briefing on COVI-19 – 11 March 2020. Retrieved from https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-COVID-19---11-march-2020

World Health Organization (n.d.-a). Advice for the public: Coronavirus disease (COVID-19). Retrieved from https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public