DISASTER IMPACTS ON THE HEALTH OF THE POPULATION AFFECTED BY THE BRUMADINHO DAM COLLAPSE, BRAZIL

Abstract

This study examines the impact of the 2019 Brumadinho dam collapse on hospitalization rates in municipalities within the Paraopeba River Basin. Using a Differences-in-Differences (DiD) model, it evaluates the disaster's effect over time by comparing affected municipalities with a control group. The results indicate significant changes in hospitalization rates in the exposed municipalities relative to the control group, highlighting the disaster's negative impact on local public health. The analysis also incorporates additional variables, including COVID-19 mortality rates, GDP, and sex ratio, to provide a more comprehensive understanding of the influencing factors.

Keywords: Paraopeba River Basin; Differences in Differences; Public Health

Introduction

The collapse of Dam I at the "Córrego do Feijão" Mine in Brumadinho, Minas Gerais, Brazil, in January 2019, was one of the worst technological disasters in Latin America. Its immediate effects included the deaths of 270 people, with three bodies still missing, and the release of 13 million cubic meters of toxic tailings that traveled over 300 km, contaminating rivers and affecting 26 municipalities.

However, beyond these immediate and tragic consequences, the Brumadinho disaster must be understood as a prolonged social and environmental crisis with lasting repercussions. Critical disaster studies emphasize that disasters are not simply isolated events but socially constructed processes that unfold over time, reproducing and deepening existing inequalities (Siena, Valencio, 2009). This perspective is essential in contexts marked by historical socioeconomic disparities and uneven access to essential services.

In the affected region, the Paraopeba River Basin plays a vital role in supplying water to nearly half of the Belo Horizonte metropolitan region and is characterized by intensive mining and steelmaking activities that have shaped patterns of settlement, employment, and environmental degradation. The toxic mudflow not only contaminated water sources but also disrupted livelihoods and damaged local infrastructure, with potential long-term demographic impacts affecting fertility, mortality, and migration dynamics (Hogan, 1991).

Health impacts from disasters are multifaceted. They may include direct injuries and fatalities as well as indirect consequences such as increased rates of infectious diseases, mental health disorders, and the disruption of routine health services. Importantly, these effects are unequally distributed. Literature highlights that disasters often intensify social inequalities, disproportionately affecting women, racialized populations, and economically marginalized groups. For instance, Freitas (2010) notes that women face increased unpaid care work, heightened vulnerability to sexual and domestic violence, and greater mental health burdens in the aftermath of disasters, especially in the Global South, where systemic inequalities limit access to resources and protection.

Moreover, the COVID-19 pandemic, which overlapped with the years following the dam collapse, adds an additional layer of complexity. It represents an exogenous shock that could both obscure and amplify disaster-related health outcomes, making it critical to consider pandemic dynamics when evaluating the disaster's impact.

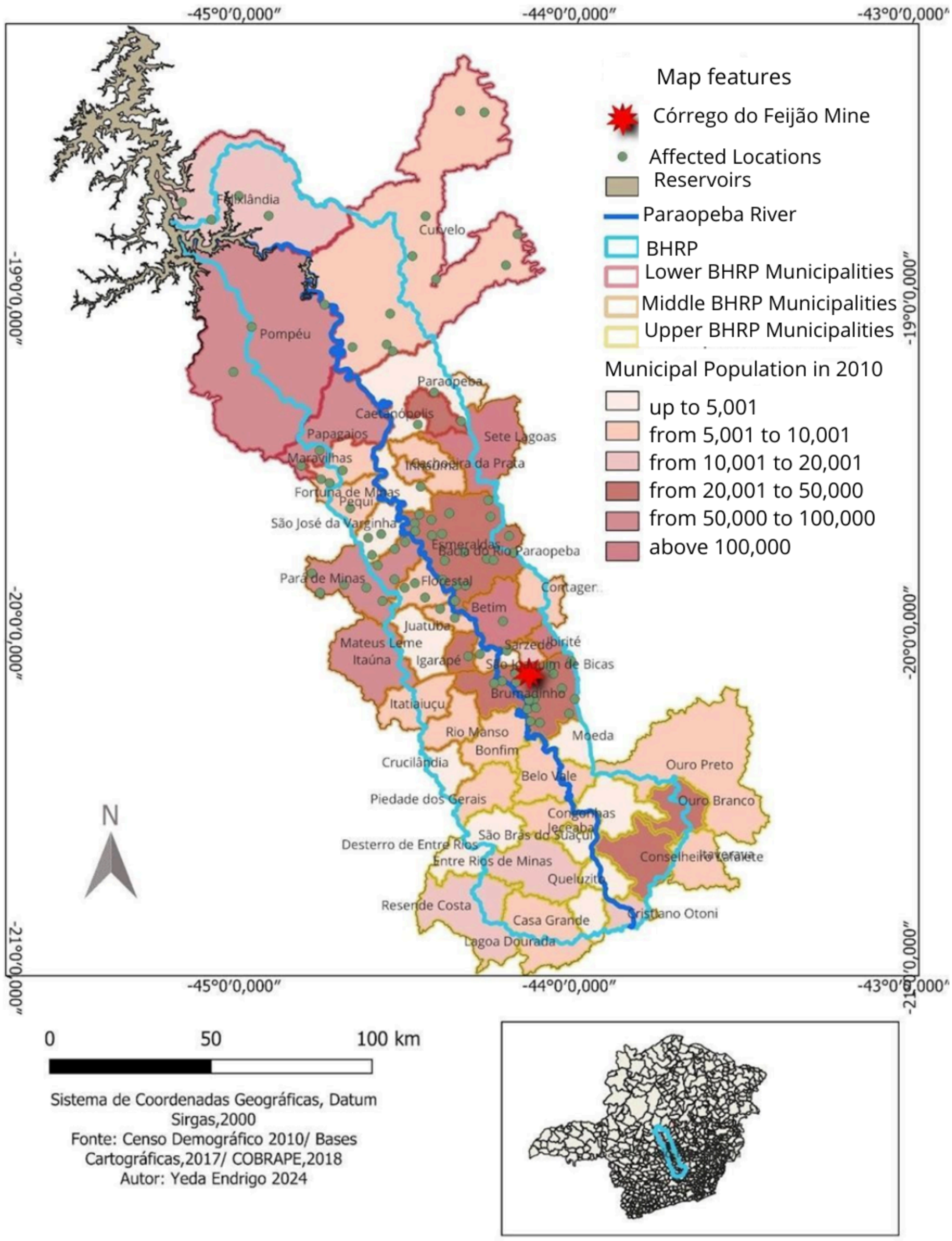
Given these challenges, quantitative method**s** that adjust for confounding variables are necessary to evaluate disaster impacts on health accurately. This study applies a Differences-in-Differences (DiD) approach to analyze changes in hospitalization rates in municipalities of the Paraopeba River Basin, comparing those identified as exposed to the disaster with carefully matched control municipalities. By including covariates such as GDP per capita, sex ratio, and COVID-19 mortality rates, this analysis seeks to provide a nuanced understanding of how the disaster affected health outcomes in a region marked by profound socioeconomic and demographic heterogeneity.

Study Area

The Paraopeba River Hydrographic Basin (BHRP) is located in the central metropolitan mesoregion of Minas Gerais, Brazil, and spans a drainage area of 12,091 km². It is one of the main tributaries of the São Francisco River (MATOS & DIAS, 2011; IGAM, 2024). Economic activities in the region are dominated by mining and steelmaking (COBRAPE, 2018), and the basin supplies water to 47% of the population of the Belo Horizonte metropolitan region (ARSAE, 2013).

The BHRP comprises 48 municipalities, classified as small (up to 10,000 inhabitants, n=21), medium (10,001–50,000, n=17), and large (over 50,001, n=10) according to IBGE standards. Together, these municipalities cover a total area of 21,493.8 km² and have a combined population of 2,632,575 residents.

Paraopeba River Basin (BHRP): Population (2010), Dam Location and Affected Areas



The Paraopeba River Hydrographic Basin (BHRP) is located in the central metropolitan mesoregion of Minas Gerais and comprises a drainage area of 12,091 km². It is considered one of the main tributaries of the São Francisco River (MATOS, DIAS, 2011; IGAM, 2024). Activities related to mining and steelmaking stand out (COBRAPE, 2018), in addition to supplying 47% of the population of the metropolitan region of Belo Horizonte (ARSAE, 2013).

The BHRP is divided into upper, middle, and lower sections, as shown in Figure 1.

The BHRP is composed of 48 municipalities that are divided between high, medium and low. Of these, 21 municipalities are small (up to 10,000 inhabitants), 17 municipalities are medium-sized (between 10,001 and 50,000 inhabitants) and 10 municipalities are large (more than 50,001 inhabitants) according to IBGE parameters.

The total area of the 48 municipalities together corresponds to 21,493.8 kilometers and a population of 2,632,575 people.

Methodology

This is an exploratory study based on secondary data from the Hospital Information System (SIH), made available by DATASUS. The SIH records hospitalizations in facilities that are part of Brazil's Unified Health System (SUS), including public and private hospitals with agreements. Data are reported through the Hospital Admission Authorization (AIH) system.

Table 1.Variables Recorded in the SIH/SUS Hospital Information System

Category	Variables	Description
Hospital	Code, Municipality, Legal Status.	Identifying and administrative data about the hospital.
Patient	Sex, Date of Birth, Age, Municipality of Residence, Postal Code, Occupation, Economic Sector, Nationality.	Demographic and economic information of the patients.
Hospitalization	Specialty, Type of Admission (Emergency, Elective, etc.), Admission Date, Discharge Date, Days of Stay, etc.	Details about the hospitalization episode, including duration and nature of treatment.
Hospital Costs	Hospital Services, Professional Services, Diagnostic and Therapy Services, Neonatal Care, etc.	Costs associated with hospitalization, broken down by type of service.

Source: Authors

Selection of Control Municipalities: Propensity Score Matching (PSM)

To assess changes in hospitalization rates before and after the disaster, it was necessary to define the treatment and control groups for comparison. The treatment group was defined based on official classifications of municipalities directly affected by the Brumadinho dam rupture and mudflow in 2019. These classifications were obtained from government reports and impact assessments identifying the exposed municipalities.

To ensure that the control group had similar socioeconomic and demographic characteristics to those of the exposed municipalities, Propensity Score Matching (PSM) was applied among the remaining municipalities within the Paraopeba River Basin. Originally proposed by Paul Rosenbaum and Donald Rubin (1983), PSM estimates the probability of a unit receiving treatment based on observed characteristics. This estimated value, known as the propensity score (PS), enables the matching of units in the treatment and control groups with similar scores. After matching, outcomes can be compared while minimizing selection bias (CHIAVEGATTO et al., 2013).

In this study, four socioeconomic and demographic variables were used to calculate the propensity score: average income, percentage of households with access to piped water (2010 Census), formal employment rate, and crude birth rate. These variables were selected for their ability to capture structural conditions of the municipalities that could influence the observed public health impacts after the disaster.

After matching based on the propensity score, it was possible to select a set of control municipalities more comparable to the exposed municipalities, thereby increasing the robustness of the Difference-in-Differences (DiD) analysis.

Difference-in-Differences (DiD) Model

The Difference-in-Differences (DiD) approach compares changes in hospitalization rates between treatment and control groups before and after the 2019 dam collapse to identify the event’s causal impact. The method relies on the parallel trends assumption, which posits that if both groups were following similar trends before the event, any divergence observed after can be attributed to the disaster.

The general DiD regression model is expressed as:

$$Yit = \alpha + \beta Tt + \gamma Gi + \delta TtGi + \varepsilon it$$

where:

Y_it = hospitalization rate for municipality *i* at time *t*.

T_t = dummy variable indicating the post-event period (1 = post; 0 = pre).

G_i = dummy variable indicating group membership (1 = treatment; 0 = control).

T_t × G_i = interaction term representing exposure to the event in the post-treatment period.

β_3 = coefficient of interest capturing the causal effect of the disaster.

For the model estimation, the variables presented in Table 2 were used.

Table 2. Variables Used in the Analysis

Variable	Type	Description
Y	Dependent	Hospitalization rate
Group2	Independent	Dummy: 1 if municipality is exposed; 0 otherwise
Treatment	Independent	Dummy: 1 for post-rupture period; 0 for pre-rupture period
did	Independent	Interaction term: Group2 × Treatment
taxa_mortalidade	Control	Standardized COVID-19 mortality rate in the municipalities
GDP	Control	Municipal per capita Gross Domestic Product
sex ratio	Control	Proportion of men relative to women in the municipalities

Source: Authors

Table 3 presents the estimated coefficients from six Difference-in-Differences (DiD) regression models assessing the impact of the Brumadinho dam disaster on hospitalization rates in municipalities of the Paraopeba River Basin. Below, we provide an interpretation of these results.

Table 3. Results of Difference-in-Differences (DiD) Regression Models (2010–2021)

Model	Variable	Estimate	Standard Error	t-value	p-value
1 – Up to 2021	Intercept	0,001847	0,00005376	34,353	< 2e-16 ***
	Group2	0,001398	0,00005818	24,026	< 2e-16 ***
	Treatment	0,0001088	0,0001095	0,994	0,32
	Did	-0,0004957	0,0001177	-4,21	2,55e-05 ***
	Adjusted R^2	0,0009	-	-	-
2 – Up to 2021 with GDP	Intercept	0,002503	0,00005786	43,248	< 2e-16 ***
	Group2	0,001305	0,00005823	22,406	< 2e-16 ***
	Treatment	0,0002488	0,0001095	2,271	0,0231 *
	Did	-0,0004084	0,0001177	-3,47	0,000521 ***
	GDP	-0,00001782	0,000000583	-30,525	< 2e-16 ***
	Adjusted R^2	0,0021	-	-	-
3 – Up to 2021 with Sex Ratio	Intercept	0,01891	0,000121	156,235	< 2e-16 ***
	Group2	0,002129	0,00005748	37,05	< 2e-16 ***
	Treatment	0,00009904	0,0001078	0,919	0,358
	Did	-0,0006703	0,0001159	-5,782	7,38e-09 ***
	Sex Ratio	-0,0001918	0,000001224	-156,763	< 2e-16 ***
	Adjusted R^2	0,0021	-	-	-
4 – Up to 2021 with GDP and Sex Ratio	Intercept	0,01961	0,0001229	159,535	< 2e-16 ***
	Group2	0,002034	0,00005751	35,367	< 2e-16 ***
	Treatment	0,0002435	0,0001078	2,258	0,024 *
	Did	-0,0005805	0,0001159	-5,009	5,47e-07 ***
	GDP	-0,00001839	0,000000574	-31,997	< 2e-16 ***
	Sex Ratio	-0,0001921	0,000001223	-157,065	< 2e-16 ***
	Adjusted R^2	0,03272	-	-	-
5 – Up to 2019 (Pre-Pandemic)	Intercept	0,0400662	0,0004013	99,836	< 2e-16 ***
	Group2	0,0120239	0,0005676	21,185	3,93e-13 ***

	Treatment	-0,0023623	0,0012691	-1,861	0,08117 .
	Did	0,0057769	0,0017948	3,219	0,00536 **
	Adjusted R²	0,9669	-	-	-
6 – Up to 2021 Excluding COVID-19 Mortality	Intercept	0,0400662	0,000569	70,416	< 2e-16 ***
	Group2	0,0120239	0,0008047	14,942	2,58e-12 ***
	Treatment	-0,0049355	0,001138	-4,337	0,00032 ***
	Did	0,0052302	0,0016094	3,25	0,00401 **
	Adjusted R²	0,9432	-		

Source: Authors

Model 1 serves as the baseline specification with basic controls. The interaction term (Did) is negative and statistically significant, suggesting a reduction in hospitalization rates in exposed municipalities after the disaster. However, the model's explanatory power is very low (Adjusted R^2 = 0.0009), indicating it does not capture much of the variability in the data and may be omitting important factors.

Model 2 introduces GDP per capita as a control variable. Results show that GDP is significantly associated with reduced hospitalization rates, while the interaction term remains negative and significant. This supports the interpretation of lower hospitalization rates in exposed municipalities in the period following the critical event. Nonetheless, the overall model fit remains limited (Adjusted R^2 = 0.0021).

Model 3 includes the sex ratio as an additional covariate. The coefficient on sex ratio is negative and significant, indicating that municipalities with a higher proportion of men tend to have lower hospitalization rates overall. The interaction term (Did) remains negative and significant, providing further evidence of reduced hospitalization rates in exposed municipalities in the period following the critical event in exposed municipalities. The model's explanatory power remains low (Adjusted R^2 = 0.0021).

Model 4 combines GDP per capita and sex ratio as control variables. The interaction term remains negative and significant, indicating a consistent association with reduced hospitalization rates in exposed municipalities after the disaster. Both GDP and sex ratio maintain significant effects in the expected direction (reducing hospitalization rates). While this model shows a slight improvement in fit (Adjusted $R^2 = 0.03272$), it still explains only a modest portion of the variability observed.

Models 5 and 6 use a more traditional DiD specification, comparing periods before and following the critical event without including additional socioeconomic controls". In both models, the interaction term (Did) is positive and statistically significant. This indicates that when no socioeconomic covariates are included, there is an observed increase in hospitalization rates in exposed municipalities relative to the control group after the disaster.

Model 5, which is restricted to the pre-pandemic period (up to 2019), shows a Did coefficient of approximately 0.00578, indicating a meaningful increase in hospitalization rates in exposed municipalities in the short term following the dam collapse. The model's very high Adjusted R^2 (0.9669) suggests that it explains almost all the variability observed. However, the absence of socioeconomic controls may mean that this estimated effect also captures unobserved structural differences or time trends between the groups.

Model 6 extends the analysis through 2021 but excludes the standardized COVID-19 mortality rate as a control to better isolate the disaster effect. Again, the interaction term is positive (0.00523) and significant, indicating a continued increase in hospitalization rates in exposed municipalities relative to the control group. The Adjusted R^2 remains high (0.9432), reflecting strong overall fit to the time series data but also raising concerns about potential overfitting or omitted variable bias.

In summary, these results demonstrate that the inclusion of socioeconomic controls substantially influences the estimated disaster effect. Models with controls (1 to 4) suggest reductions in hospitalization rates in exposed municipalities in the period following the critical event, while simpler models without controls (5 and 6) indicate increases. This contrast highlights the importance of careful model specification and controlling for confounding factors when evaluating the health impacts of disasters.

Discussion

This study analyzed the impact of the Brumadinho dam disaster on hospitalization rates in municipalities of the Paraopeba River Basin using Difference-in-Differences (DiD) models with varying specifications. The results show that including socioeconomic control variables such as GDP per capita and sex ratio changes the estimated disaster effect.

Models with controls consistently showed negative and significant interaction terms, suggesting reductions in hospitalization rates in exposed municipalities compared to controls in the period following the critical event. This pattern may reflect multiple mechanisms, such as changes in healthcare-seeking behavior, reallocation of local health resources, or adaptations in the organization of health services, including the prioritization of emergency cases and postponement of elective admissions—effects that may have been intensified by the COVID-19 pandemic context.

Conversely, simpler models without controls showed positive and significant interaction terms, indicating increases in hospitalization rates after the disaster in exposed municipalities relative to controls. These specifications likely fail to account for structural differences between municipalities or broader temporal trends unrelated to the disaster. The contrast between specifications reinforces the importance of carefully specifying models to avoid biased interpretations of disaster impacts on health outcomes.

Additionally, even models with socioeconomic controls displayed very low Adjusted R^2 values, indicating that much of the variability in average hospitalization rates per municipality remains unexplained. This highlights the complexity of isolating the effect of a disaster on health outcomes measured at the municipal level, especially in the context of pre-existing socioeconomic inequalities, differences in health system access, and overlapping exogenous shocks such as the pandemic.

Conclusion

The results of this study demonstrate that estimates of the Brumadinho dam disaster's impact on hospitalization rates depend strongly on model specification and the inclusion of socioeconomic control variables. Models with controls suggest reductions in hospitalization rates in exposed municipalities following the critical event period, while simpler before-and-after comparisons indicate significant increases.

These findings highlight the need for careful, context-sensitive, and theoretically grounded analyses when assessing the health impacts of disasters. Policymakers and public health planners should interpret simple before-and-after comparisons cautiously, recognizing that failing to account for structural inequalities between exposed and non-exposed regions can lead to misleading conclusions and potentially ineffective interventions.

One of the main limitations of this study is the use of municipality-level aggregated hospitalization rates, without access to individual-level data that would allow for the capture of detailed demographic characteristics, pre-existing conditions, or personal healthcare trajectories. Additionally, the analysis did not differentiate between causes of hospitalization (e.g., ICD-10 chapters), which may obscure important variations in the types of diseases most affected by the disaster.

Future research should aim to use more granular data to investigate impacts by ICD-10 chapter and to analyze inequalities by race/ethnicity and gender. Such an approach is essential to understand how the effects of environmental disasters are distributed unequally across population subgroups and to support the development of more equitable public health policies that are sensitive to historical and social vulnerabilities.

Moreover, the results underscore the importance of preparing health systems and risk management policies for scenarios involving the succession and overlap of disasters, as occurred with the dam rupture unfolding in the broader context of the COVID-19 pandemic. Recognizing the possibility of multiple, interconnected critical events is essential for strengthening community resilience and reducing inequalities in access to and outcomes from health care.

Bibliography

Abadie, A. (2005). Semiparametric Difference-in-Differences Estimators. *Review of Economic Studies*, 72(1), 1–19. https://doi.org/10.1111/0034-6527.00321

Chiavegatto Filho, A. D. P., et al. (2013). Como incluir características dos distritos do município de São Paulo em estudos epidemiológicos?: Análise da desigualdade de renda pelo uso do propensity score matching. *Saúde e Sociedade*, 22, 1145–1153.

Freitas, R. C. M. (2010). The construction of an agenda concerning gender, socio-environmental disasters, and development. *Estudos Feministas*, 5(SE), 0–0.

Freitas, C. M. de, Barcellos, C., Asmus, C. I. R. F., Silva, M. A. da, & Xavier, D. R. (2019). From Samarco in Mariana to Vale in Brumadinho: Mining dam disasters and Public Health. *Cadernos de Saúde Pública*, 35(5).

Hogan, D. J. (1991). Crescimento demográfico e meio ambiente. *Revista Brasileira de Estudos de População*, 8(1/2), 61–71.

Instituto Brasileiro de Geografia e Estatística (IBGE). (n.d.). Informações sobre Brumadinho. Retrieved from https://www.ibge.gov.br/novo-portal-destaques/23639-informacoes-brumadinho.html

Siena, M., & Valencio, N. (2009). Gênero e desastres: Uma perspectiva brasileira sobre o tema. In N. Valencio et al. (Eds.), *Sociologia dos Desastres: Construção, interfaces e perspectivas no Brasil* (pp. 58–66). São Carlos: RiMa Editora.