Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Ian Fowler Date

Analysis of COVID-19 and Chikungunya Co-circulation between 2019 and 2020 in the community-based cohort of Pau da Lima, Salvador, Brazil

By

Ian Fowler
Master of Public Health

Department of Epidemiology

_____

Benjamin Lopman, PhD.
Faculty Advisor

_____

Guilherme Ribeiro, MD PhD.
Field Advisor

Analysis of COVID-19 and Chikungunya Co-circulation between 2019 and 2020 in the community-based cohort of Pau da Lima, Salvador, Brazil

By

Ian Fowler

Biology, B.S.
Georgia State University
2019

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Epidemiology
2023

Abstract

Analysis of COVID-19 and Chikungunya Co-circulation between 2019 and 2020 in the community-based cohort of Pau da Lima, Salvador, Brazil

By Ian Fowler

An outbreak of chikungunya in the Brazilian state of Bahia leading from 2019 into 2020 created a co-circulation event in which the populace was subjected to both the arbovirus as well as the emerging SARS-CoV-2 global pandemic. A cohort study was being conducted in the Pau da Lima neighborhood to measure arbovirus infection rates which was then adapted to capture data regarding SARS-CoV-2 within the population. The spatial distributions of the SARS-CoV-2 status, chikungunya infection status, and other relevant covariates were mapped, and the simultaneous outbreaks were described. An infection rate of approximately 50% for chikungunya virus and 45% for SARS-CoV-2 was identified between 2019 and 2020, which represents significantly high transmission rates for both infections. Moreover, 42.2% of those with incident chikungunya virus infection in 2020 had positive results for SARS-CoV-2 IgG ELISA compared to 47.3% among those with prevalent chikungunya virus infection in 2019 and 46.8% of those with no chikungunya virus infection. No statistically significant association was identified between chikungunya status and SARS-CoV-2 incidence and there was no evidence of a spatially dependent process affecting the relationship. Together this suggests that the high rates of infections resulting from both viruses were independently co-circulating infections that did not interact, spatially or otherwise. Further research would be required to determine if this pattern is consistent across other populations that experienced simultaneous SARS-CoV-2 and arbovirus outbreaks.

Analysis of COVID-19 and Chikungunya Co-circulation between 2019 and 2020 in the community-based cohort of Pau da Lima, Salvador, Brazil

By

Ian Fowler

Biology, B.S.
Georgia State University
2019

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Epidemiology
2023

Acknowledgements

A special thanks to Dr. Guilherme Ribeiro, MD PhD.

who included me in his incredible research,
guided me throughout this process, and
gave me the opportunity to further develop my knowledge and skills

AND

To Dr. Michael Kramer, PhD.

for all his help with the spatial analyses included in this research,
both conceptually and technically

AND

To my friends I made while at Emory

who have also been struggling through their theses
and have made my time at Rollins
significantly better and richer

Table of Contents

1. Introduction

1.1 Chikungunya

Chikungunya virus was first discovered in the 1950s in Africa from which it spread to Asia in the 1960s and the Americas in the 1980s. However, there is evidence of infections dating back to the 1600s or earlier. [1] Initial cases of chikungunya in the Americas were solely imported cases with no local transmission and autochthonous transmission was not identified until 2013 in San Martin. [2] Since then, the range of the *Ae. aegypti* and *Ae. albopictus* mosquitos have continued to expand, and with them, the number of chikungunya epidemics, as well as endemic regions, have increased. Brazil has experienced chikungunya outbreaks consistently for most of the past decade. In 2014, the first two autochthonous outbreaks occurred in states of Amapá and Bahia from which the virus spread primarily within the north and northeastern regions before outbreaks began to be identified throughout the rest of the country. [3] Now Brazil has the highest rate of cases in the world with over 100 per 100,000 persons and with a total of 265,289 confirmed cases in 2022. [4]

1.2 SARS-CoV-2

Severe Acute Respiratory Syndrome Coronavirus 2, or SARS-CoV-2, is a novel human infection first identified in China at the end of 2019. The virus spread internationally throughout 2020 and was declared by the World Health Organization (WHO) as a global pandemic on the 11th of March. As of January 2023, there have been 674,914,950 cases worldwide, leading to over 6 million deaths. [5] Until the end of 2020, there were no available pharmaceutical preventions for SARS-CoV-2 infections and the primary strategies recommended by the WHO or the Center for Disease Control and Prevention (CDC) included face coverings, social distancing, self-isolation, and other non-pharmaceutical interventions (NPIs). [6, 7] These strategies proved mostly successful, especially where rigorously practiced or enforced, however, many situations which served to increase transmission rates also limited the effectiveness of NPIs.

The first case of SARS-CoV-2 was confirmed in Brazil on the 26th of February 2020, having been imported from Europe by upper class Brazilians. [8] By the end of 2020, approximately 7.7 million cases had been identified in Brazil leading to over 190,000 deaths. [9] The SARS-CoV-2 pandemic in Brazil continued to spread, creating a "hotspot" of transmission that placed it just behind the United States and India on the global scale. [9, 10] The rate of infections within the country is compounded by the proportion of the population living in "urban slums or informal settlements," approximately 15%, that are affected by an increased SARS-CoV-2 infection risks and increased risk of negative outcomes; health, social, and economic; associated with the pandemic. [11]

1.3 Co-circulation

While chikungunya and SARS-CoV-2 represent two distinct viral infections with differing causes, patterns, and transmissions, there are notable similarities through which co-circulation events may be produced.

Low socioeconomic neighborhoods within Brazilian cities have become "hotspots" for chikungunya virus infections, as they provide ideal conditions for the mosquito vectors combined with limited access to health care and other necessary resources. [12] With the lack of regular access to drinking water and insufficient sanitation in informal settlements, water must be collected or stored in informal containers providing the mosquitos with ample breeding sites. [13] Furthermore, as the *Ae*. *aegypti* mosquitos thrive in densely populated areas, and the overcrowding of many low-income neighborhoods in Brazilian metropolitan areas, creates additional opportunities for infectious spread.

SARS-CoV-2 infection also disproportionally affects certain groups within vulnerable populations. Individuals between the ages of 20 and 59 years old are generally the most infected and current literature is divided regarding the role of biological sex in both SARS-CoV-2 infections and outcomes. [14, 15] Conversely, there has been strong evidence that SARS-CoV-2 infections are more prevalent in ethnic and racial minorities, which are inherently linked to socioeconomic class standings. [16] Moreover, certain

essential occupations, specifically those in public-facing or crowded settings, increased the risk of SARS-CoV-2 infection. [17] There is additional evidence of socio-economic factors and the associated increase in infection risk. [18] Factors such as economic status, crowded living conditions, poor sanitation, and poor access to health services exacerbated the spread of the pandemic. [17, 18]

The symptomology of both chikungunya and COVID-19 are remarkably similar, creating difficulty for clinical diagnosis. There is evidence suggesting that these similarities have led to the misdiagnosis of COVID-19 as chikungunya, especially early in the pandemic, suggesting that, during concomitant epidemic events, greater attention must be taken when clinically diagnosing diseases and that laboratory confirmation of diagnosis is of increased necessity. [19]

There is little available data estimating the association between chikungunya virus infection and its effect on subsequent SARS-CoV-2 infection. A study conducted in Espírito Santo, Brazil investigating SARS-CoV-2 incidence rates in populations with varying levels of endemic arbovirus infections found that, in areas with higher rates of chikungunya infections, COVID-19 reached incidence levels varying between 1.79 to 3.40% as compared to approximately 2.28% at the state level. [20] A second study conducted in Espírito Santo found that only 2.85% of participants with evidence of dengue or chikungunya infections were positive of SARS-CoV-2 according to chemiluminescent microparticle immunoassay. [19] In another study from Guerrero, Mexico, 10.23% of participants who tested positive for SARS-CoV-2 had evidence of concurrent or recent chikungunya infection. [21] Studies were also identified describing the relationships between SARS-CoV-2 infection and other vector-borne infections. A co-infection rate for dengue was found to be 10.3%, 14.3% for malaria, and 11.4% for overall vector-borne co-infection in a study conducted in Luanda, Angola. [22] A study of Amazonians living in Brazil found that those with prior dengue virus infections were at increased risk for clinically apparent SARS-CoV-2 infection. [23] This finding is corroborated by the research of the WHO while contradicting studies conducted by Nicolelis et al. and

Silvestre et al. which suggest prior dengue infection may have a protective relationship with later SARS-CoV-2 infection. [24 – 26]

1.4 Motivation

The present study seeks to investigate the relationships between chikungunya virus immunological status (no infection, incident infection, or prevalent infection); and incident SARS-CoV-2 infection in a prospective cohort in Brazil. Co-infections between the SARS-CoV-2 and other concurrent endemic or epidemic diseases have been rarely studied, with especially little work dedicated to the interactions between pathogens from entirely different classes. Furthermore, no identified study has attempted to explore these connections while engaging in a spatially dependent analysis.

To this end, data from a cohort study conducted in a selected neighborhood within Pau da Lima, Salvador, Brazil, was analyzed. The study began in 2019 to investigate chikungunya virus and other arboviruses infection rates in the closed population, but as the SARS-CoV-2 pandemic grew in prevalence within Brazil, the focus was shifted to also include the new viral infection, which may mimic many of the symptoms of arbovirus infection. Additionally, since the study period was completed before the advent of the SARS-CoV-2 vaccines, the SARS-CoV-2 infections identified through IgG antibody development represent true incident infections rather than an immune response to the vaccine.

2. Methods

2.1. Community Based Cohort Study

Conducted between 2019 and 2021, the prospective cohort included residents of 252 surveyed households located within the Pau da Lima neighborhood of Salvador, Brazil. The recruitment and initial survey occurred between September and November of 2019 with a follow-up period from October to December of 2020. The primary intention of the study was to estimate the prevalence and incidence of arbovirus

infections; specifically, chikungunya, dengue, and zika; within the vulnerable populations of Salvador, however as the SARS-CoV-2 pandemic emerged in the region at the beginning of 2020, the focus was shifted to also gather data on the novel infection.

2.1.1 Study Setting

Bahia is a Brazilian state located in the northeast region along the coast. It is the fourth largest state and, at the time of the 2022 national census, has a population of 14,659,023 people. The capital of Bahia is the city of Salvador which houses 2,610,987 people, including the approximate 24,000 residents of Pau de Lima. [27] The study was conducted within a selected area of the Pau da Lima neighborhood in Salvador, Bahia, Brazil, and the area was defined by its proximity to the São Marcos Health Unit, as well as its inclusion in previous studies. [28, 29] It included 606 active study participants distributed across 252 households. The Pau da Lima neighborhood is an urban slum according to the definition established by UN-HABITAT; thus, the majority of the participants are at increased risk for infectious diseases as they lack equal access to health care, sanitation, housing, and social or economic resources.

2.1.2 Sample Selection

Participants for the cohort were recruited through random selection of households; volunteers within the community interested in participating were also included. After an initial census of the households in the study area, 200 households were randomly selected and the members of those households, who slept in the residence for at least three nights a week, were invited to participate in the study. Community volunteers who did not live in the selected households were also able to participate in the study. Participants were limited to those over the age of six months old and those who were under 18 years old signed informed assent forms, with parental consent, before participation, while those 18 or older signed informed consent forms. A second round of consent documents was produced and signed upon modification of study procedures to address the SARS-CoV-2 pandemic.

2.1.3 Data Collection

Initial data collection involved both the administration of a standardized questionnaire for all participants as well as the collection of blood for the conduction of enzyme-linked immunoassay (ELISA) analysis to detection of arboviruses and SARS-CoV-2 antibodies in the sera. EUROIMMUN ELISA was used for the SARS-CoV-2 IgG and chikungunya virus IgG analyses while the chikungunya virus IgM ELSIA was provided by InBios International Inc. Samples from every participant were tested at enrollment and during the predetermined follow-up periods for antibodies using ELISA analysis to identify past or present infections.

The questionnaire collected relevant participant demographic information including, but not limited to, questions regarding sex, age, skin color, education level, occupation status, household monthly income, and socio-economic status. The questionnaire also included questions regarding the prior symptoms compatible with chikungunya virus, and later SARS-CoV-2, infection as well as the history of past medical suspicion of chikungunya and potential exposure events.

2.1.4 Case Criteria

A SARS-CoV-2 infection was defined upon the detection of IgG SARS-CoV-2 antibodies in sera obtained between 2019 and the follow-up in 2020. Furthermore, cases were restricted to those with positive results between February 2020, when the first case of SARS-CoV-2 was recorded in Brazil, and the end of December of that same year, before the release of the COVID-19 vaccines. With this temporal window, the likelihood of capturing true cases among those testing positive for SARS-CoV-2 antibodies was greatly increased.

The immunological status in relation to chikungunya virus infection allowed for the classification of the study participants into three categories*: prevalent infection, incident infection,* and *no infection*. All three categories were defined from the ELISA IgG and IgM antibody tests for chikungunya. *Prevalent infections*

or *historic infections* were defined as those with positive ELISA results at the time of study enrollment, suggesting chikungunya virus infection predating the study period. *Incident infections* were defined as positive ELISA results occurring during the follow-up period, between October and December of 2020, among those who were chikungunya negative at enrollment. Finally, those who never returned a positive result for chikungunya antibodies were defined as *no infection*.

2.2. Covariate Distributions

The covariates included in the analyses were selected based upon a review of the literature surrounding SARS-CoV-2 and chikungunya virus infections and due to *a priori* criteria. The distribution of each covariate included in the regression models was mapped using kernel density estimation. These surfaces illustrated the distribution of various factors within the study population and how, visually, they relate to the distribution of chikungunya and SARS-CoV-2 infections.

Demographic characteristics included sex, age, skin color, education history, marital status, employment, family size, family income, and others. Sex (male or female), age (continuous), and skin color (white, black, mixed, Asiatic, or indigenous) were self-reported, and it should be noted that the collection of skin color data is used as a proxy for racial and ethnic groups within the community and specifically acts as a surrogate for socioeconomic status and disparities. Employment was also self-reported as a binary yes or no.

Socio-economic status was defined as the social class according to the Critério Brazil as outlined by the Brazilian Association of Research Companies (ABEP). The population was divided into six class categories: A, B1, B2, C1, C2, and D-E. These classes represent an aggregation of monthly household income combined with other economic resources with Class A representing the highest socioeconomic class and Class E representing the lowest.

2.3. Regression Models

To model the effects of chikungunya virus immunological infection status (incident infection, prevalent infection, or no infection); and the selected covariates on the likelihood of SARS-CoV-2 infection within the study population, two classes of regression models were employed: a linear probability model and a logistic regression. These two classes of models are then further divided between aspatial and spatial models. All the models shared the same potential confounders.

The linear probability and logistic regression models are primarily used to estimate the association between chikungunya infection status and incident SARS-CoV-2 infection. The spatial econometric models, based upon the linear model, and the geographically weighted regression, based upon the logistic regression, are used to investigate potential spatial effects upon the association and spatial heterogeneity throughout the population, respectively. After these regressions are run, additional point estimates are generated, accounting for the spatial processes, which can be compared to the associations determined by the aspatial models. These comparisons are intended to help characterize the relationships between the co-circulating viruses.

2.3.1 Linear Probability Model [30]

Spatial econometric models, which help to describe the relationships between "spatial spillovers" of the exposure, outcome, or both, cannot be run using a logistic regression formula. Thus, to satisfy this requirement, the binary SARS-CoV-2 incidence outcome data was represented using a linear probability model. This model will act as a direct point of comparison for the three econometric models as they share similar mathematical processes.

The linear probability model may violate the assumption that the data is normally distributed due to the binary results of the outcome, however, this violation has a minimal effect on the β estimates for the relationships between the covariates and the outcome. This represents an application of the ordinary least

squares formula to the binary outcome data of SARS-CoV-2 incidence. The estimation represents the covariate conditions associated with a one unit change in the outcome status, in this instance, that represents SARS-CoV-2 infection. The primary discrepancy in the linear probability model as compared to the logistic regression is applied to the standard error resulting from the linear model. While the standard errors from the linear probability model violate the normalcy assumption, the normalcy of the outcome may remain intact as it is conditioned on the value of the covariates. Thus the linear probability model can effectively represent binary outcome data.

An aspatial linear probability model was constructed using the *lm( )* package, which will share its formula with the spatial econometric models, being run in comparison.

2.3.2 Logistic Regression Model

Unlike the spatial econometric models, the geographically weighted regression can run using a logistic regression formula. In order to fit the parameters and create an aspatial comparison model, a logistic regression was also included in the analysis. This regression used the *glm( )* package in *R Studio*. For the comparisons between modeling approaches, the logistic model was designed to be roughly equivalent to the linear probability model, described above.

2.4. Spatial Econometric Models [31]

The spatial economic models, described further below, represent, within a regression formula, the spatial spillover of the exposure, outcome, or both and how it affects the relationships between participants in the study. As opposed to the aspatial models, linear or logistic, the econometric models consider the effect the exposures and outcomes of individuals within the analysis have on one another, whereas the aspatial models only consider the exposure status of the individual. This process involves the direct inclusion of spatial dependence in a regression that estimates the effect of chikungunya status on SARS-CoV-2 incidence while addressing multiple covariates. This is achieved by including a representation of the processes by which

the variables interact through the spill over. These models are especially effective in infectious disease analyses, as is the case with this report, due to its simulation of infectious processes in which the exposure or outcome status of one individual is dependent on the status of those around them.

Spatial econometric models are modifications to linear regressions in which the distribution of the data is assumed to be normal and independent. As linear models generally use rate or risk outcome data and the outcome data in this study is represented as a binary, linear probability models, as discussed above, were employed to allow for the use of the econometric models within these circumstances. The resulting coefficients from these models can be compared to the aspatial model to estimate the effect of the spatial processes on the association of interest.

2.4.1 Spatial Error Model

The spatial error model, using the *errorsarlm( )* function and the inverse distance neighbor weights matrix, represents unmeasured autocorrelation as a result of unknown covariates. This model reflects a spatial dependence among variables not accounted for by including a correlation coefficient for the model's inherent statistical error (λ) which ranges from -1 to 1 with values approaching either extreme representing greater statistical correlation. The weights matrix allows the distance between neighbors to contribute to the likelihood and the intensity of the spatial dependent events.

2.4.2 Spatial Lag Model

Conversely the spatial lag model represents SARS-CoV-2 spatial dispersion events resulting in autocorrelation in the regression model. The model includes a representation of the "spatially lagged value of the neighbor's outcome" and its effect on an individual's outcome as ρ which, like λ, also ranges from -1 to 1. Put simply, the SARS-CoV-2 status of one individual affects the status of their neighbors, just like an infection would spread, and this is considered along with the other variables included in the model. The *lagsarlm( )* function is employed and, as with the other econometric models, the neighbor matrix allows for

the weighting of the spatial dependence of the points as a function of their distance from one another. This represents a spatial contagion parameter and is an effective model of the spread of SARS-CoV-2.

2.4.3 Spatial Durbin Model

Both exposure and outcome spillover events are considered in the spatial durbin model which is an extension of the spatial lag model and also uses the *lagsarlm( )* but sets the *durbin* option to "TRUE." This allows for same process considering spatial spillover of SARS-CoV-2 outcome in the previous model to also apply to the chikungunya status of each participant. Both the measures of the exposure for the given individual and the measures of the exposure and outcome for the individual's neighbors, as defined by the inverse distance matrix, are factored into the association with the outcome. Again, as both SARS-CoV-2 and chikungunya are spread through infectious processes, although with different pathways, this econometric model allows for an effective representation of the physical systems while accounting for spatial autocorrelation.

2.4.4 Model Comparisons

To understand the processes by which the relationship between chikungunya status and SARS-CoV-2 are connected, the abilities of the three econometric models must be compared to determine which has the best "fit" and most effectively explains the autocorrelation seen in the aspatial models. The models are compared using the Akaike Information Criteria, or AIC, which combines the number of independent variables in each model with the maximum likelihood estimates in order to quantify the amount of variation in the data that each model leaves unexplained. Thus, the model with the lowest AIC, among those compared, is the best fit for the data.

2.5. Geographically Weighted Regression Model [32]

To explore variations in parameter values across the study area and to determine the role of spatial non-stationarity in the exposure-outcome relationship, a geographically weighted regression model was

designed for the analysis. The model shares its formula with the aspatial logistic regression and generates an estimate of the association for each point, or in this case each participant, using an iterative process repeating the model with a pre-determined bandwidth. These two models can then be compared to estimate how the odds ratio changes throughout the study area and where it differs from the aspatial model.

The analysis involves the comparison between a global model, in which stationarity is assumed and one value for each parameter is applied to the entire study area, versus a local model, which allows for the testing of non-stationarity by allowing the parameters to vary throughout space. This process describes the distribution of β estimates as normal disease mapping would describe prevalence. Thus, the geographically weighted regression helps to illustrate patterns in the data generating processes underlying the relationship between SARS-CoV-2 and chikungunya status. The model was run using the *GWmodel( )* package.

2.5.1 Optimizing Kernel Bandwidth

In order to set the area in which each run of the geographically weighted regression will include, the kernel bandwidth must be optimized, similar to the process completed during the kernel density estimations. As before, the bandwidth is set to adaptive to best address density variations throughout the study area. This was completed using the *bw.gwr( )* function which accounts for variations associated with all of the included covariates.

3. Results

3.1. Description of Data

3.1.1 Population Characteristics

A sample of 606 individuals across 149 households in Pau da Lima community participated in the study. 374 (61.7%) individuals across 211 unique households were included in this analysis with 232 participants

excluded due to missing data in fields relevant to the regression models and required for some of the complex analytical methods.

Table 1: Description of Study Population

	COVID-19 Positive (N=168)	COVID-19 Negative (N=206)	Overall (N=374)
Chikungunya status			
Incident infection	68 (40.5%)	93 (45.1%)	161 (43.0%)
Prevalent infection	26 (15.5%)	29 (14.1%)	55 (14.7%)
No infection	74 (44.0%)	84 (40.8%)	158 (42.2%)
Sex			
Male	61 (36.3%)	77 (37.4%)	138 (36.9%)
Female	107 (63.7%)	129 (62.6%)	236 (63.1%)
Age (years)			
Mean (SD)	37.5 (16.7)	38.9 (16.5)	38.2 (16.6)
Median [min, max]	37 [13, 78]	38 [13, 98]	37 [13, 98]
Skin color			
White	7 (4.2%)	8 (3.9%)	15 (4.0%)
Black	91 (54.2%)	123 (59.7%)	214 (57.2%)
Mixed	64 (38.1%)	72 (35.0%)	136 (36.4%)
Asian	5 (3.0%)	2 (1.0%)	7 (1.9%)
Indigenous	1 (0.6%)	1 (0.5%)	2 (0.5%)
Socioeconomic class in 2019			
Class A	0 (0%)	0 (0%)	0 (0%)
Class B1	0 (0%)	0 (0%)	0 (0%)
Class B2	10 (6.0%)	6 (2.9%)	16 (4.3%)
Class C1	31 (18.5%)	32 (15.5%)	63 (16.8%)
Class C2	55 (32.7%)	103 (50.0%)	158 (42.2%)
Class D-E	72 (42.9%)	65 (31.6%)	137 (36.6%)
Employment status in 2020			
Employed	70 (41.7%)	95 (46.1%)	165 (44.1%)
Unemployed	98 (58.3%)	111 (53.9%)	209 (55.9%)

168 (44.9%) of the participants tested positive for SARS-CoV-2 by ELISA analysis during the study period and 161 (43%) had incident chikungunya infection in 2020 while 55 (14.7%) had prevalent infection from 2019. Of those involved in the study, 138 (36.9%) were male and the mean age was 38 years old (SD = 16.6). The majority of the study population was self-described as black (57.2%), was of socioeconomic class C2 (42.2%), and was unemployed (55.9%). These values were relatively consistent across SARS-CoV-2 result strata and more detailed descriptions are included in *Table 1*.

3.1.2 SARS-CoV-2 and Chikungunya Prevalence

Table 2: Description of Infection Incidence and Prevalence

	Number with antibodies / total (prevalence of incidence, %)	
	2019	**2020**
Chikungunya virus		
Incidence 2019-2020	–	161 / 319 (50.4)
Cumulative Prevalence 2020	55 / 374 (14.7)	216 / 374 (58.0)
SARS-CoV-2 virus		
Incidence 2019-2020	–	168 / 374 (44.9)
Cumulative Prevalence 2020	1 / 374 (0.5)*	168 / 374 (44.9)
Co-infection		
Incidence for both 2019-2020	–	68 / 347 (19.6)
Cumulative prevalence for both 2020	27 / 374 (7.2)	95 / 374 (25.4)

* There was no recorded transmission of SARS-CoV-2 in Brazil during 2019, the participant detected with SARS-CoV-2 IgG antibodies in 2019 likely represents a false positive result due to prior infection with other viruses in the coronavirus subgroup or an unspecified cross-reaction. This participant was summarily removed from the subgroup followed in 2020 for estimation of incidence and prevalence.

Table 2 shows the incidence and prevalence values of SARS-CoV-2 and chikungunya for the 374 participants included in the analysis as determined by the serosurveys conducted during the study period. Chikungunya antibody prevalence increased from 14.7% in 2019 to 58.0% in 2020 describing a situation in which over half of the participating population tested positive for the arbovirus. Similarly, a significant increase in SARS-CoV-2 infections of 44.9 was also recorded from 2019 to 2020. One individual was found to be positive for SARS-CoV-2 infection in 2019, although this likely represents a false positive in the ELISA analysis as the first case recorded by the Brazilian government did not occur until February of 2020. Of the 161 with incident chikungunya virus infection between 2019 and 2020, 68 (42.2%) tested positive for SARS-CoV-2 and 93 (57.8%) tested negative, whereas 26 (47.3%) tested positive and 29 (52.7%) test negative among the 55 with prior chikungunya virus infection. Among the 158 with no history of infection, 74 (46.8%) tested positive for SARS-CoV-2 and 84 (53.2%) tested negative (*Table 3*). Therefore, no subtantial difference was observed in the risk of SARS-CoV-2 infection in respect to the immunological status of CHIKV infection.

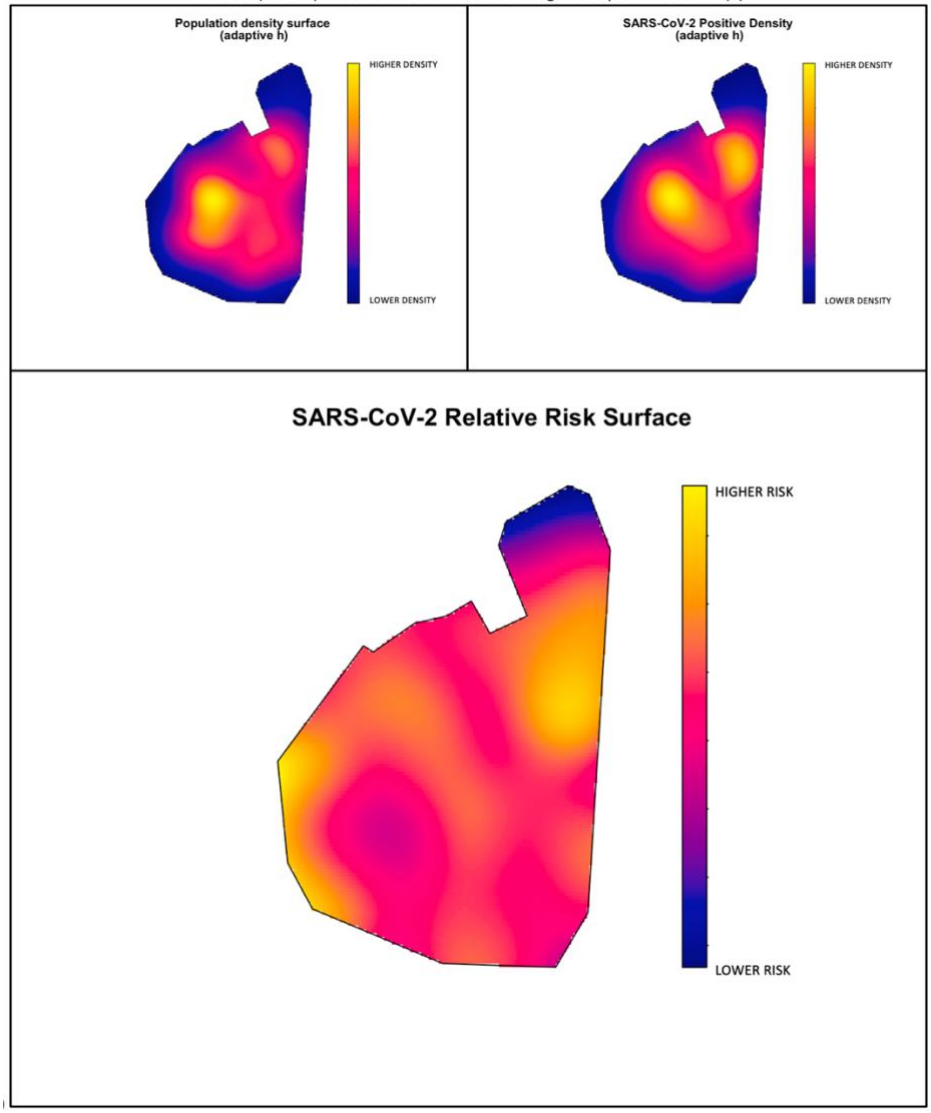
Table 3: SARS-CoV-2 ELISA Results compared to Chikungunya Status

	Chikungunya Status in 2020		
	Incident	**Prevalent**	**No Infection**
SARS-CoV-2 Status			
Positive	68 / 161 (42.2)	26 / 55 (47.3)	74 / 158 (46.8)
Negative	93 / 161 (57.8)	29 / 55 (52.7)	84 / 158 (53.2)

3.2. Distribution Maps

The population distribution within the boundaries of the study site, as represented by the kernel density estimation with an adaptive bandwidth in *Figure 1*, shows that there are two major sites in which the study population is concentrated: one in the southern half, just to the left of center, and one in the lower portion

Fig 1: Map of population density (*top left*), positive for SARS-CoV-2 infection density (*top right*), and relative risk of SARS-CoV-2 infection (*bottom*) with kernel estimation smoothing via adaptive bandwidth (h)



of the northern half, just to the right of center. The site within the lower half expands, to a lesser degree, to the right of center whereas the other site is much more localized.

3.2.1 SARS-CoV-2 Positive and Relative Risk Surfaces

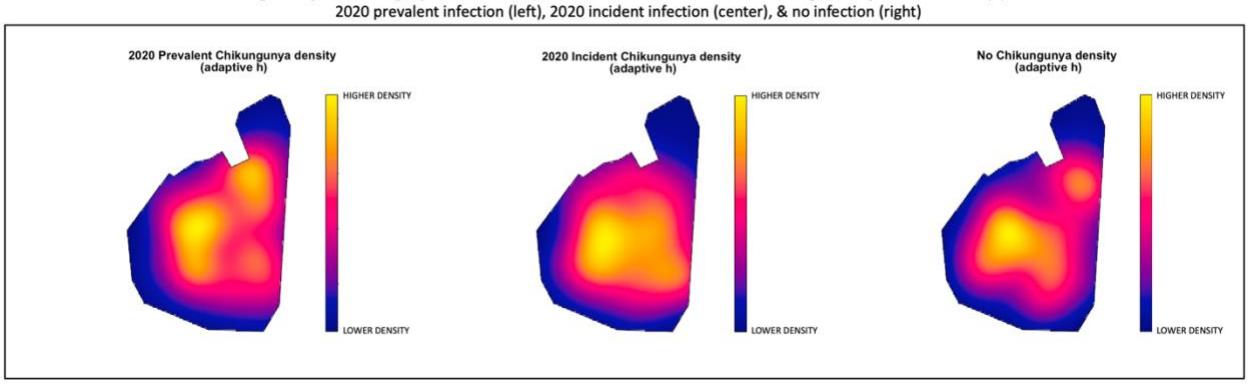
Figure 1 also shows that the distribution of positive SARS-CoV-2 cases follows a similar trend as the population density within the study area although there are larger 'hot spots' for the infection within the sites of higher population values. This suggests, as expected, that the distribution of SARS-CoV-2 cases follows a traditional infectious process which is connected to the density of the population.

The relative risk surface generated by comparing the distribution of SARS-CoV-2 positive cases to the population density surface generated a map showing an even distribution of high risk for SARS-CoV-2 infection throughout the study area. This suggests that all study participants are at a relatively equal risk for SARS-CoV-2.

3.2.2 Chikungunya Exposure Surface

Similarly, in *Figure 4*, the chikungunya cases are centralized around the areas of population density however show a more diffuse pattern, highlighting the mobility of the vector and its distribution throughout the study area. Chikungunya infection rates have also been found to be directly correlated to population density and other conditions present in low-income settings which would be pervasive throughout Pau da Lima.

Fig 4: Maps of chikungunya viral infection status with kernel estimation smoothing via adaptive bandwidth (h)
2020 prevalent infection (left), 2020 incident infection (center), & no infection (right)



3.2.3 Covariate Distributions

The distribution of females (63.1%) within the study area show similarities to the population distribution surface as they are localized within similar sites of higher density. The male population however is primarily focalized within the west half of the central area with significantly less diffusion as compared to the females. This is described visually in *Appendix Figure B1*.

Figure B2 shows that white participants, representing 4% of the total population, were primarily located within the southern half of the study area, black participants (57.2%) within the left southern quadrant; in a similar location as the male distribution, mixed race individuals (36.4%) were relatively evenly distributed throughout the study area, and people of Asian descent (1.9%) were congregated toward the right side of the northern half of the area. The indigenous population only represented 0.5% of the overall population included in the study.

The socioeconomic classes fell into similar distributions as the population density surface and, notably, class C2; which represented the majority of the study population (42,2%); almost exactly matched the population distribution while classes D-E (36.6%) were almost entirely located within the southern half of the study area. Classes A and B1 were not mapped as none of the study participants met the qualifications for either classification (*Figure B3*).

The employed portion of the population is localized within the same area as the male study population whereas the unemployed population is more evenly distributed, although concentrated to the southern half of the study area in *Figure B4*. With a slight minority of the population employed (44.1%) in addition to the location of their distribution, the idea that the male population (36.9%) represent the majority of employed individuals is corroborated.

As a continuous variable, age was not able to be represented using a kernel density estimation surface.

3.3. Aspatial Models

3.3.1 Linear Probability Model

Two linear probability models were implemented: one using the unadjusted formula and the other using the adjusted formula, both of which are described in *Table B1* within the Appendix.

The unadjusted linear probability model, which only considered the relationship between chikungunya status; incident infection, prevalent infection, or no infection; and SARS-CoV-2 status, which estimated a change in chikungunya infection status as resulting in no change in the risk of SARS-CoV-2 infection (β = -0.04 [95% CI: -0.15, 0.07] and β = -0.01 [CI 95%: -0.16, 0.15]). The adjusted model: which controlled for sex, age, skin color, socioeconomic status, and employment status; estimated very similar point values with similar 95% confidence intervals. However, socioeconomic status was found to be significantly associated with the SARS-CoV-2 status with a p-value of 0.004.

3.3.2 Logistic Regression Model

The aspatial logistic regression employed the same formulas as the linear probability model. In the unadjusted model, participants with incident chikungunya virus infection had similar odds of developing SARS-COV-2 infection (OR = 0.83 [95% CI: 0.53 to 1.29]) as compared to those with evidence of prevalent infections (OR = 1.02 [95% CI: 0.55 and 1.88]) and those negative for chikungunya virus antibodies. Once again, the odds ratios resulting from the adjusted model were effectively the same as those of the unadjusted model. Similarly to the linear model, socioeconomic status was significantly associated in the adjusted regression analysis (p-value = 0.004).

Table 4: Results of Aspatial and Spatial Regression Models

	Linear Probability Model			Logistic Regression Model			Spatial Error Model			Spatial Lag Model			Spatial Durbin Model		
	Beta	95% CI	p-value	OR	95% CI	p-value	Beta	Std. Error	p-value	Beta	Std. Error	p-value	Beta	Std. Error	p-value
Chikungunya status 1*	-0.04	-0.15, 0.07	0.47	0.84	0.53, 1.33	0.47	-0.02	0.05	0.68	-0.03	0.05	0.57	-0.03	0.05	0.57
Chikungunya status 2*	-0.01	-0.16, 0.15	0.90	0.96	0.50, 1.82	0.90	-0.01	0.07	0.83	–	0.07	0.97	–	0.07	0.97
Sex	0.01	-0.10, 0.12	0.89	1.03	0.65, 1.63	0.88	-0.02	0.05	0.73	-0.01	0.05	0.85	-0.01	0.05	0.85
Age	0	0	0.24	0.99	0.98, 1.01	0.23	–	–	0.14	–	–	0.41	–	–	0.41
Skin color															
White	0.03	-0.72, 0.78	0.94	1.12	0.04, 34.7	0.94	0.05	0.35	0.89	-0.07	0.36	0.84	-0.07	0.36	0.84
Black	-0.01	-0.71, 0.70	0.99	0.97	0.04, 26.7	0.98	0.11	0.33	0.74	-0.04	0.34	0.90	-0.04	0.34	0.90
Mixed	0.04	-0.66, 0.75	0.90	1.20	0.04, 33.3	0.90	0.14	0.33	0.68	-0.01	0.34	0.99	-0.01	0.34	0.99
Asian	0.30	-0.49, 1.1	0.45	3.68	0.10, 145	0.44	0.40	0.37	0.29	0.27	0.38	0.47	0.27	0.38	0.47
Socioeconomic Class			0.004			0.004									
A	–	–		–	–		–	–	–	–	–	–	–	–	–
B1	–	–		–	–		–	–	–	–	–	–	–	–	–
B2	–	–		–	–		–	–	–	–	–	–	–	–	–
C1	-0.14	-0.42, 0.14		0.56	0.17, 1.74		-0.23	0.21	0.28	-0.03	0.14	0.86	-0.03	0.14	0.86
C2	-0.30	-0.56, -0.04		0.29	0.09, 0.85		-0.37	0.20	0.06	-0.11	0.13	0.40	-0.11	0.13	0.40
D-E	-0.11	-0.37, 0.15		0.63	0.20, 1.87		-0.18	0.20	0.37	0.05	0.13	0.69	0.05	0.13	0.69
Employment status	-0.05	-0.16, 0.06	0.36	0.81	0.51, 1.26	0.35	-0.02	0.05	0.64	-0.04	0.05	0.44	-0.04	0.05	0.44
Lambda (λ)	–	–	–	–	–	–		5.25e-5							
Rho (ρ)	–	–	–	–	–	–					4.36e-5			5.16e-5	

* Chikungunya status 1 (chik_stat1) and Chikungunya status 2 (chik_stat2) represent nominal variables created for the analysis which describe the three statuses in which participants can fall. A value of 1 for Chikungunya status 1 with a value of 0 for Chikungunya status 2 represents an incident infection. Conversely, a value of 0 for Chikungunya status 1 and 1 for Chikungunya status 2 represents a prevalent infection. A value of 0 for both Chikungunya status 1 and 2 represents a participant with no chikungunya virus infection.

3.3.3 Aspatial Model Comparisons

The Akaike Information Criteria (AIC) value for the adjusted linear probability model was estimated as 563.37 which when compared to the value for the adjusted logistic regression, at 521.94, suggests that the logistic model is a better fit for the data set used in the analysis.

3.4. Spatial Econometric Models

The spatial error model estimates were found to be statistically insignificant and effectively zero. The value of λ, which is the correlation coefficient for the model's statistical error, was found to be 5.25e-5 and the AIC of the model was 519.85. The λ value approaching zero and the absence of statistically significant associations within the model suggests that there is no evidence of spatial dependence among variables that are not included in the model.

The lag model estimated a β value for incident chikungunya virus infection compared to no infection of -0.03 (p-value of 0.57), a ρ value of 4.36e-5; which represents the spatial dependence, and an AIC of 525.89. With no statistically significant association between variables in the model and a ρ value of 4.36e-5, the

model showed no evidence of "spatial spillover" between the outcome values of neighbors within the study area.

Similarly, the spatial durbin model estimated the same β value as the lag model with the same p-value, however the ρ value and AIC were slightly higher at 5.16e-5 and 536.02 respectively. There was still no evidence of spatial processes affecting the relationship between the exposure and the outcome.

3.4.4 Econometric Model Comparisons

The AIC value of the aspatial model was found to be 563.37 which can be compared to the 519.85 of the error model, the 525.89 of the lag model, and the 536.02 of the durbin model. While this suggests that the spatial econometric models are a better fit for the data, as opposed to the aspatial model, the lack of significant associations generated by either class of model likely invalidates any conclusions drawn from these comparisons.

4. Discussion

4.1. Principal Findings

4.1.1 Infection Prevalence

During the observation period, over half of the studied population developed chikungunya virus infection (50.4%), according to ELISA analysis, and just under half were infected by SARS-CoV-2 (44.9%). Approximately one quarter of the population were infected by both chikungunya and SARS-CoV-2 (25.4%). These values represent a significantly higher incidence rate as compared to the national value for Brazil at approximately 3.5% for SARS-CoV-2 and <0.1% for chikungunya. [9, 33] Due to the small size of the population, the accuracy of extrapolations to that level must be considered. Nonetheless, this corroborates the idea that co-circulation events between SARS-CoV-2 and other viruses; such as

chikungunya, dengue, zika, and malaria; did occur and that the shift in public health efforts to focus on the SARS-CoV-2 pandemic may have masked these events.

4.1.2 No Evidence of Interaction

The results of the regression analyses suggest that, within this study population, the odds of SARS-CoV-2 infection among those with incident or prevalent chikungunya infection are no different from those with no history of chikungunya infection. This indicates that the two infections spread independently throughout the population and the status of one infection did not affect the status of the other. Similar results were generated with both linear and logistic analyses. Additionally, there was very little variation in SARS-CoV-2 infection across strata divided by sex, skin color, or employment. Where the literature is mostly divided regarding the association between arbovirus infection status and subsequent SARS-CoV-2 infection, with about half finding a protective effect and half finding a harmful effect, this study found evidence suggesting that both viruses did in fact circulate simultaneously within the cohort but did not affect one another. It should be noted that the similarity in the values resulting from the linear probability model and logistic regression model suggest that the linear regression adequately represented the binary outcome data with minimal distortion to the point estimates.

4.1.3 Spatial Effects

The spatial econometric models estimated similar, nonsignificant values for the exposure status and the covariate associations as the linear probability model. This suggests that “spatial spillover” involving exposure or outcome status did not affect the results of the regressions and that there is spatial independence between the values of individuals included in the analysis. This further supports the findings that the two co-circulating outbreaks were not connected and that the spatial distributions of SARS-CoV-2 infection was not impacted by that of chikungunya infection. Additionally, the geographically weighted regression was unable to run, using either optimized or self-generated bandwidth values, and thus no results were included in the analysis. This indicates that there is not a significant amount of spatial heterogeneity within

the study population regarding these regression coefficients. The relative risk surface generated through kernel density estimation further supports this conclusion as the risk of SARS-CoV-2 infection is evenly distributed throughout the study bounds.

4.2. Limitations

The study was primarily limited by the size of the analyzed population. Only 606 participants were included in the initial study and of those only 374 were able to be included in the analysis. The relatively small sample size likely influenced the ability to detect associations between chikungunya status and SARS-CoV-2 incidence. Furthermore, regarding the spatial analyses, the size of the study population and study bounds may have similarly affected attempts to identify spatial dependence and heterogeneity. This is possibly an example of the modifiable areal unit problem, which may be avoided by varying the level of scale at which the data is analyzed. Additional research would need to be conducted to determine the effect the population size had on the values or significance of these results.

Moreover, the nature of the participant sampling may have introduced selection bias. While participant residences were randomly selected from those within the bounds of the study area, participation was based upon voluntary involvement in the study, both from the selected households and from the larger community. This may have introduced bias into the data collection as there may have been differences in response rates and in loss to follow-up between those with recent or current infections and those with no history of infection affecting exposure and outcome distributions. Additional bias may have been introduced during the data collection stage.

The limited amount of previous data regarding the association between previous or concurrent chikungunya infection and its effect on incident SARS-CoV-2 infection posed a challenge in finding comparisons for the current research and so it is difficult to draw wider conclusions based on this data. This limitation will be discussed in more detail below.

4.3. Comparisons

When compared to other studies examining SARS-CoV-2 and arbovirus co-infection, the prevalence values found during the current analysis were substantially higher with more individuals testing positive for both chikungunya and SARS-CoV-2 IgG during the study period. The calculated prevalence value for chikungunya and SARS-CoV-2 co-infection of 25.4% is approximately 10 times that found by Vicente et al. at 2.28% or Stringari et al. at 2.85%. [19, 20] Nunez-Avellaneda et al. found an approximate co-infection prevalence of 10.23% within their study population, still less than half of the value determined during the present study. [21] It should be noted that in the three comparison studies, the population size was greater than 1,000 participants, as opposed to the 374 included in this analysis. Moreover, the present study was conducted in a region of significantly higher rates of both chikungunya and SARS-CoV-2 infections, potentially reflecting some of the differences in the chikungunya prevalence and SARS-CoV-2 incidence values.

No studies were identified which estimated a measure of association between chikungunya infection and SARS-CoV-2 infection, however, examples were found discussing the harmful or protective benefits of prior dengue infection on COVID-19 incidence. The studies were even distributed between those who found an increased risk of infection among those with evidence of dengue antibodies and those which found a decreased risk. [22 – 26] The present study suggests no association between chikungunya status and SARS-CoV-2 infection.

4.4. Implications

This study was intended as a descriptive, exploratory analysis investigating the relationships between chikungunya status (*incident infection, prevalent infection,* or *no infection*) and SARS-CoV-2 incidence within a closed population. The analysis found little to no differences across age, sex, racial/ethnic, or infection status strata in terms of the odds of incident SARS-CoV-2 infection. To this end, it was determined

that there was no statistically significant relationship between chikungunya infection and SARS-CoV-2 infection, including when spatially dependent processes were considered, and that there was little crossover regarding the factors that influence chikungunya virus infection and those that influence SARS-CoV-2 infection. Socioeconomic status was found to be a risk factor for SARS-CoV-2 infection, but the risk of SARS-CoV-2 infection was evenly distributed across the other factors considered in the analysis, which contrasts previous research on the factors associated with infection risks for these viruses.

Due to the disparate finding within the literature regarding the associations between vector-borne diseases and subsequent SARS-CoV-2 infection, the results of this analysis should not be seen as conclusive. The relatively small size of the population, despite the ideal timing of the cohort study, may have reduced the magnitude and the precision of any estimated measures of association, although the significant number of SARS-CoV-2 infection events suggests precision. Further research including different, larger study populations would be needed to better investigate the relationship. Additionally, differing study populations may better elucidate any potential spatial relationships concerning chikungunya and SARS-CoV-2 co-circulation events.

5. References

1. Weaver SC, Charlier C, Vasilakis N, et al. Zika, Chikungunya, and Other Emerging Vector-Borne Viral Diseases. *Annu Rev Med*. 2018;69:395–408.

2. Weaver SC. Arrival of Chikungunya Virus in the New World: Prospects for Spread and Impact on Public Health. *PLOS Neglected Tropical Diseases*. 2014;8(6):e2921.

3. Nunes MRT, Faria NR, de Vasconcelos JM, et al. Emergence and potential for spread of Chikungunya virus in Brazil. BMC Medicine. 2015;13(1):102.

4. Chikungunya worldwide overview. (https://www.ecdc.europa.eu/en/chikungunya-monthly). (Accessed February 15, 2023)

5. COVID Live - Coronavirus Statistics - Worldometer. (https://www.worldometers.info/coronavirus/). (Accessed January 30, 2023)

6. Lamb YN. BNT162b2 mRNA COVID-19 Vaccine: First Approval. *Drugs*. 2021;81(4):495–501.

7. Flaxman S, Mishra S, Gandy A, et al. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature*. 2020;584(7820):257–261.

8. Shadmi E, Chen Y, Dourado I, et al. Health equity and COVID-19: global perspectives. *International Journal for Equity in Health*. 2020;19(1):104.

9. Mathieu E, Ritchie H, Rodés-Guirao L, et al. Coronavirus Pandemic (COVID-19). *Our World in Data* [electronic article]. 2020; (https://ourworldindata.org/coronavirus/country/brazil). (Accessed February 5, 2023)

10. Silva SJR da, Pena L. Collapse of the public health system and the emergence of new variants during the second wave of the COVID-19 pandemic in Brazil. *One Health*. 2021;13:100287.

11. Fofana MO, Jr NN, Ticona JPA, et al. Structural factors associated with SARS-CoV-2 infection risk in an urban slum setting in Salvador, Brazil: A cross-sectional survey. *PLOS Medicine*. 2022;19(9):e1004093.

12. Magalhaes T, Chalegre KDM, Braga C, et al. The Endless Challenges of Arboviral Diseases in Brazil. *Tropical Medicine and Infectious Disease*. 2020;5(2):75.

13. Philbert A. Preferred breeding habitats of Aedes Aegypti (Diptera- Culicidae) Mosquito and its public health implications in Dares Salaam, Tanzania Anitha Philbert1* and Jasper. N. Ijumba2. *E3 Journal of Environmental Research and Management 2141-7466*. 2013;4.

14. Salzberger B, Buder F, Lampl B, et al. Epidemiology of SARS-CoV-2. *Infection*. 2021;49(2):233–239.

15. Kopel J, Perisetti A, Roghani A, et al. Racial and Gender-Based Differences in COVID-19. *Frontiers in Public Health* [electronic article]. 2020;8. (https://www.frontiersin.org/articles/10.3389/fpubh.2020.00418). (Accessed December 20, 2022)

16. Estrela FM, Soares CFS e, Cruz MA da, et al. Pandemia da Covid 19: refletindo as vulnerabilidades a luz do gênero, raça e classe. *Ciênc. saúde coletiva*. 2020;25:3431–3436.

17. Chen JT, Krieger N. Revealing the Unequal Burden of COVID-19 by Income, Race/Ethnicity, and Household Crowding: US County Versus Zip Code Analyses. *Journal of Public Health Management and Practice*. 2021;27(Supplement 1):S43.

18. Koelle K, Martin MA, Antia R, et al. The changing epidemiology of SARS-CoV-2. *Science*. 2022;375(6585):1116–1121.

19. Stringari LL, Souza MN de, Junior NF de M, et al. Covert cases of Severe Acute Respiratory Syndrome Coronavirus 2: An obscure but present danger in regions endemic for Dengue and Chikungunya viruses. *PLOS ONE*. 2021;16(1):e0244937.

20. Vicente CR, Silva TCC da, Pereira LD, et al. Impact of concurrent epidemics of dengue, chikungunya, zika, and COVID-19. *Rev. Soc. Bras. Med. Trop.* [electronic article]. 2021;54. (http://www.scielo.br/j/rsbmt/a/xLBTRGsz8fwzWXf433DMLrJ/?lang=en). (Accessed October 10, 2022)

21. Nunez-Avellaneda D, Villagómez FR, Villegas-Pineda JC, et al. Evidence of Coinfections between SARS-CoV-2 and Select Arboviruses in Guerrero, Mexico, 2020–2021. *Am J Trop Med Hyg*. 2022;106(3):896–899.

22. Sebastião CS, Gaston C, Paixão JP, et al. Coinfection between SARS-CoV-2 and vector-borne diseases in Luanda, Angola. *Journal of Medical Virology*. 2022;94(1):366–371.

23. Nicolete VC, Rodrigues PT, Johansen IC, et al. Interacting Epidemics in Amazonian Brazil: Prior Dengue Infection Associated With Increased Coronavirus Disease 2019 (COVID-19) Risk in a Population-Based Cohort Study. *Clin Infect Dis*. 2021;73(11):2045–2054.

24. Pan American Health Organization/World Health Organization. Epidemiological update: Dengue in the context of COVID-19. 3 December 2020, Washington, D.C. PAHO/WHO. 2020. https://iris.paho.org/bitstream/handle/10665.2/53174/EpiUpdate3December2020_eng.pdf?sequence=1&isAllowed=y. Accessed 15 March 2021.

25. Nicolelis MAL, Raimundo RLG, Peixoto PS, et al. How super-spreader cities, highways, hospital bed availability, and dengue fever influenced the COVID-19 epidemic in Brazil. 2020;2020.09.19.20197749. (https://www.medrxiv.org/content/10.1101/2020.09.19.20197749v1). (Accessed April 5, 2023)

26. Silvestre OM, Costa LR, Lopes BVR, et al. Previous dengue infection and mortality in COVID-19. *Clin Infect Dis*. 2020; ciaa1895.

27. 2022 Census | IBGE. (https://www.ibge.gov.br/en/statistics/social/education/22836-2020-census-censo4.html?=&t=resultados). (Accessed February 15, 2023)

28. Kikuti M, Cunha GM, Paploski IAD, et al. Spatial Distribution of Dengue in a Brazilian Urban Slum Setting: Role of Socioeconomic Gradient in Disease Risk. *PLoS Negl Trop Dis*. 2015;9(7):e0003937.

29. de Aguiar DF, de Barros ENC, Ribeiro GS, et al. A prospective, multicentre, cohort study to assess the incidence of dengue illness in households from selected communities in Brazil (2014-2018). *Int J Infect Dis*. 2021;108:443–453.

30. Linear Probability Model. (https://murraylax.org/rtutorials/linearprob.html). (Accessed April 6, 2023)

31. Kramer M. Week 11 Spatial Regression II: Spatial econometric regression | EPI 563: Spatial Epidemiology, Fall 2022. (Accessed April 6, 2023).(https://mkram01.github.io/EPI563-SpatialEPI/spatial-regression-ii-spatial-econometric-regression.html). (Accessed April 6, 2023)

32. Kramer M. Week 12 Spatial Regression III: Geographically Weighted Regression | EPI 563: Spatial Epidemiology, Fall 2022. (Accessed April 6, 2023).(https://mkram01.github.io/EPI563-SpatialEPI/spatial-regression-iii-geographically-weighted-regression.html). (Accessed April 6, 2023)

33. Gutiérrez LA, https://www.facebook.com/pahowho. PAHO/WHO Data - Weekly Report | PAHO/WHO. *Pan American Health Organization / World Health Organization*. 2019;(https://www3.paho.org/data/index.php/en/mnu-topics/chikv-en/550-chikv-weekly-en.html). (Accessed April 17, 2023)

34. RStudio Team (2022). RStudio: Integrated Development Environment for R. RStudio, PBC, Boston, MA URL http://www.rstudio.com/.

35. Spatial Weights as Distance Functions. (https://spatialanalysis.github.io/lab_tutorials/Spatial_Weights_as_Distance_Functions.html). (Accessed April 6, 2023)

36. Kramer M. Week 7 Disease Mapping IV: Kernel Density Estimation | EPI 563: Spatial Epidemiology, Fall 2022. (Accessed April 6, 2023).(https://mkram01.github.io/EPI563-SpatialEPI/disease-mapping-iv-kernel-density-estimation.html). (Accessed April 6, 2023)

6. Appendices

6.A. Additional Methods

6.A.1 Data Preparation

The data collected from the study was imported in *R Studio* [34] as Excel files. Variables were translated from Portuguese into English and observation values were converted into the proper formats, such as dates or binary variables being converted from character variables to date or integer variables, respectively. Additionally, nominal and categorical variables were created from the available data where applicable for future analyses. Summation and aggregate variables were also created before duplicated or unnecessary observations were removed from the working data set. A dictionary for the data set used in the analysis is included in the appendices.

6.A.2 Study Bounds

The boundary of the neighborhood in which the study occurred was imported into R Studio as a shape file with the *ESPG 32724* coordinate system. The 32724 ESPG is a projection of the World Geodetic System 1984 (WGS 84) coordinate reference system that creates equal area projections for the eastern coast of Brazil between 36° and 42°W which includes all of Salvador, Bahia.

6.A.3 Establishing Spatial Geometry

The coordinates for the residence of each of the participants in the study were included in the data and the data set was converted into a spatial data frame using the *sf package*. To match the coordinate system of the study bounds, the Coordinate Reference System (CRS) of the spatial data was also set to *ESPG 32724*. As multiple participants were recruited from the same household, and therefore shared latitude and longitude points, the coordinates for every participant were randomly "jittered" by 0.05km to create completely unique values for development of the neighbor matrix.

6.A.4 Spatial Neighbors Matrix [35]

To represent the infectious network between individuals within the study, and how that has affected incidence rates, the spatial relationships of the participants must be quantified. This is achieved through the development of a spatial neighbor matrix in which the relationships between individuals are weighted in accordance with a spatial parameter. Participants in this study were assigned "neighbors" using the inverse distance continuity in which the furthest distance between two closest points was used to assign neighbor relationships. This defined the spatial relationships as a function of the "nearness" to each participant in so far as the closer two points are the greater the effect they have on one another.

The nearest neighbor for each point was found using the *knn2nb* function and the maximum value from that dataset was set as the critical threshold. A neighbor distance bandwidth was then set between zero and the value defined as the critical threshold. Distances and inverse distances between each point were calculated using the bandwidth, which allowed for the creation of the neighbor weights matrix using *nb2listw* and *nb2mat* functions.

The previously complete spatial "jittering" was necessary for the execution of the inverse distance neighbor weights matrix as identical points would have a neighbor distance of zero creating an undefinable inverse distance. The resulting matrix included 94.19% non-zero linkages with a minimum number of 1 neighbor, a maximum of 504 neighbors, and an average value of 479.

6.A.5 Kernel Density Estimations [36]

Kernel density estimations with adaptive bandwidths were used to describe the distributions of participants across various factors and characteristics. The kernel density estimation acts as a non-parametric smoothing of the local intensity of point values into a summary surface describing the entire area. The adaptive bandwidth allowed for smoothing of the point values to better represent the actual distribution by continuously modifying the size of the measured kernel in relation to the density of the points.

To create the estimation surfaces, the outcome of interest and the coordinates of the participant residences were isolated in temporary data sets that were then converted into spatial point files, which includes the number of points and the spatial intensity, using the *ppp( )* function. The spatial point conversion process includes the limitation of the data to the study boundary using the previously defined study bounds object. Similarly, the reference bandwidth was generated using an optimization function on the spatial point file as part of the *OS( )* function. The resulting surface was then mapped using the *plot( )* function.

6.A.6 Relative Risk Surfaces

Relative risk surfaces, representing the risk of chikungunya virus and SARS-CoV-2 infection across the study area, were generated using the distributions of positive cases as compared to the population distribution. The spatial point files for positive cases and for the total population were converted into a risk surface using the *risk( )* function, while maintaining the adaptive bandwidth, which could then be mapped with the *plot( )* function.

6.B. Additional Tables & Figures

Table B1: Model Formulas

Unadjusted	$sars.stat20 \sim \alpha + \beta_1 chik.stat1 + \beta_2 chik.stat2$
Adjusted	$sars.stat20 \sim \alpha + \beta_1 chik.stat1 + \beta_2 chik.stat2 + \gamma_1 sex + \gamma_2 age + \gamma_3 skin.color1 + \gamma_4 skin.color2 + \gamma_5 skin.color3 + \gamma_6 skin.color4 + \gamma_7 seclass1 + \gamma_8 seclass2 + \gamma_9 seclass3 + \gamma_{10} seclass4 + \gamma_{11} seclass5 + \gamma_{12} employment$

Table B2: Summary of Inverse Distance Neighbor Weights Matrix

Number of individual regions	374
Number of nonzero links	130954
Percentage of nonzero weights	93.62
Average number of links	350

	Number of links	Number of regions
Number of least connected regions	1	1
Number of most connected regions	369	31

Fig B1: Map of biological male population (left) and female (right) with kernel estimation smoothing via adaptive bandwidth (h)

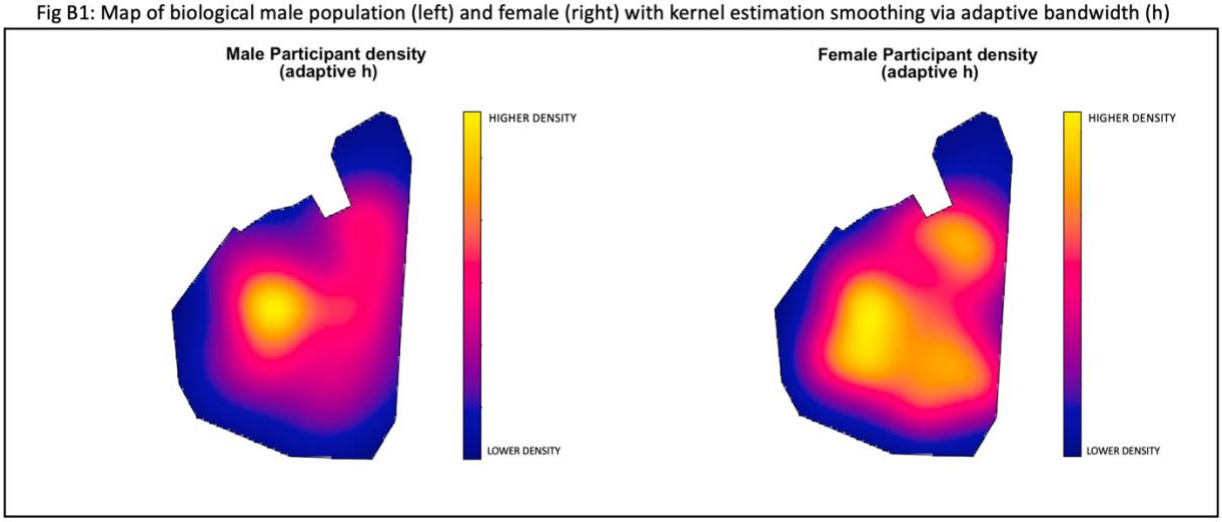


Fig B2: Map of self-assigned skin color with kernel estimation smoothing via adaptive bandwidth (h)
White participants(upper left), Black participant (upper right)
Mixed participants (lower left), Asian participants (lower right)

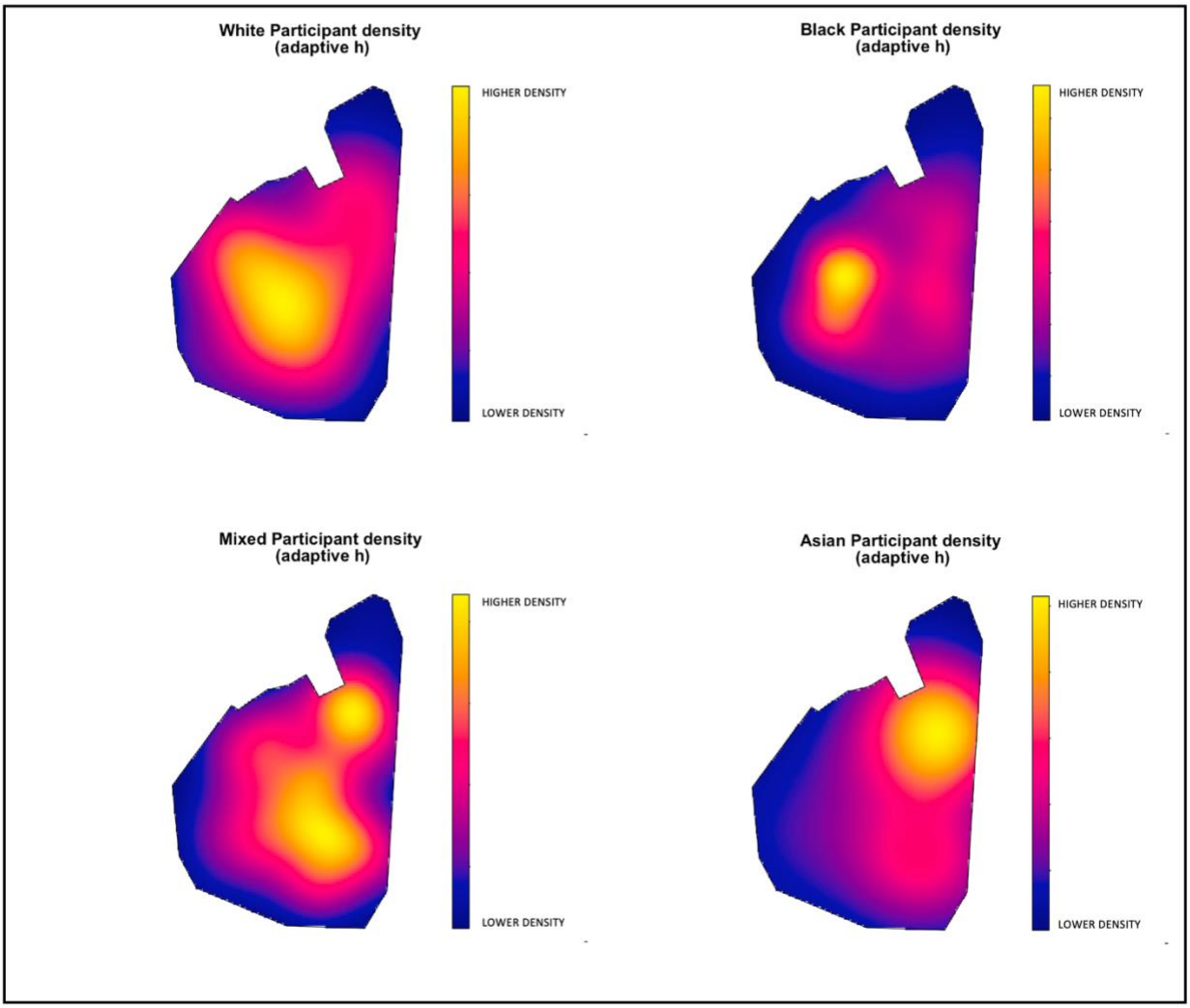


Fig B3: Map of socioeconomic class with kernel estimation smoothing via adaptive bandwidth (h)
Class B2 (upper left), Class C1 (upper right)
Class C2 (lower left), Class D-E (lower right)

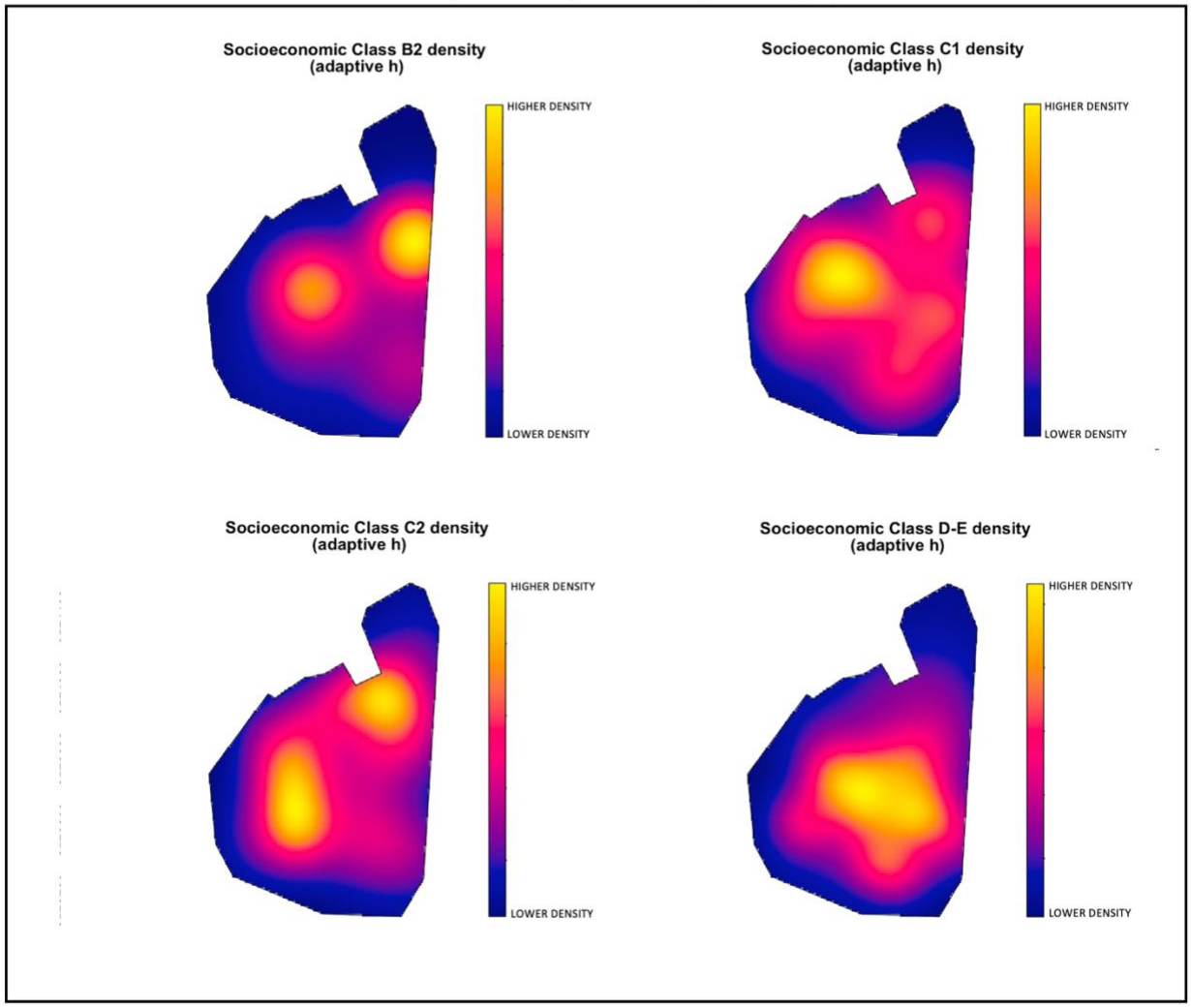
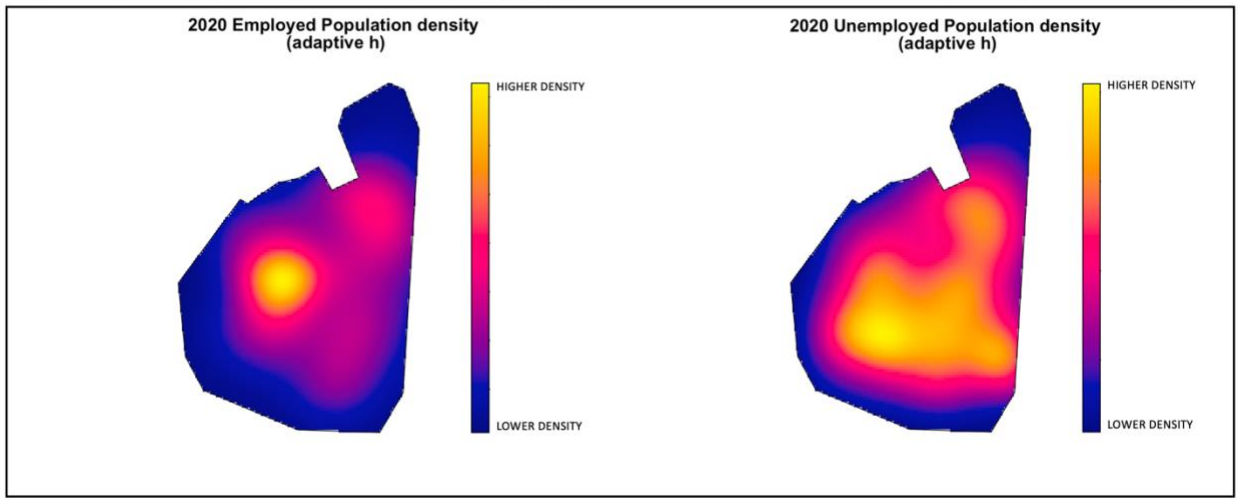


Fig B4: Map of employed (left) versus unemployed (right) populations with kernel estimation smoothing via adaptive bandwidth (h)



6.C. Additional Results

6.C.1 Inverse Distance Neighbor Matrix

The inverse distance neighbor weights matrix generated 130954 linkages between the 374 individuals included in the study representing spatial relationships within the bandwidth and each assigned a weight based upon the distance between points in the links. The average number of links for each point was 350 with only one point having one link and 31 points having a maximum of 369 links each. More information regarding the construction of the matrix can be found in section *6.A.4* and the output is further described in *Table B2* in the Appendix. The matrix was then employed for all of the following spatial analyses.

6.C.2. Geographically Weighted Regression

The geographically weighted regression was set up using an adaptive bandwidth established by the *bw.gwr()* function which would shift the bandwidth according to the number of points included within each search window, however the *gwr.basic()* function, and in turn the regression, was unable to run due to the large size of the bandwidth generated which included the majority of the study population. Additional fixed and adaptive bandwidths were attempted using generated and hand-picked values with no change to the execution of the geographically weighted regression.