medRxiv preprint doi: https://doi.org/10.1101/2024.01.26.24301823; this version posted January 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

An ecological study of COVID-19 outcomes among Florida counties

Sobur Ali[1], Taj Azarian[1*]

[1] Burnett School of Biomedical Sciences, University of Central Florida, Orlando, Florida, USA.

*Corresponding author: Taj Azarian, University of Central Florida (taj.azarian@ucf.edu)

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2024.01.26.24301823; this version posted January 26, 2024. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Abstract

Background: During the COVID-19 pandemic, Florida reported some of the highest number of cases and deaths in the US; however, county-level variation in COVID-19 outcomes has not been comprehensively investigated. The present ecological study aimed to assess corelates of COVID-19 outcomes among Florida counties that explain variation in case rates, mortality rates, and case fatality rates (CFR) across pandemic waves.

Method: We obtained county-level administrative data and COVID-19 case reports from public repositories. We tested spatial autocorrelation to assess geographic clustering in COVID-19 outcomes: case rate, mortality rate, and CFR. Stepwise linear regression was employed to test the association between case, death, and CFR and 18 demographic, socioeconomic, and health-related county-level predictors.

Results: We found mortality rate and CFR were significantly higher in rural counties compared to urban counties, among which significant differences in vaccination coverage was also observed. Multivariate analysis found that the percentage of the population aged over 65 years, the percentage of the obese people, and the percentage of rural population were significant predictors of COVID-19 case rate. Median age, vaccination coverage, percentage of people who smoke, and percentage of the population with diabetes were significant influencing factors for CFR. Importantly, vaccination coverage was significantly associated with a reduction in case rate ($R = -0.26$, $p = 0.03$) and mortality ($R = -0.51$, $p < 0.001$). Last, we found that spatial dependencies play a role in explaining variations in COVID-19 CFR among Florida counties.

Conclusion: Our findings emphasize the need for targeted, equitable public health strategies to reduce disparities and enhance population resilience during public health crises. We further inform future spatial-epidemiological analyses and present actionable data for policies related to preparedness and response to current and future epidemics in Florida and elsewhere.

Keyword: COVID-19, Florida, Socioeconomic, Regression Analysis, CFR

medRxiv preprint doi: https://doi.org/10.1101/2024.01.26.24301823; this version posted January 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

1. Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), responsible for the pandemic of the coronavirus disease 2019 (COVID-19), has posed an unprecedented crisis to public health worldwide. As of December 2023, COVID-19 has resulted in 103 million cases and 1.14 million deaths in the US. As the pandemic unfolded, it became increasingly evident that demographic, socioeconomic, and health-related risk factors play a crucial role in influencing the COVID-19 outcome among communities in the US (Grasselli et al., 2020b; Onder et al., 2020).

Socioeconomic predictors, including income, education, employment, and access to healthcare services, have long been recognized as critical determinants of health outcomes (Braveman and Gottlieb, 2014). These factors may be associated with an individual's capacity to adhere to public health guidelines, access healthcare resources, and implement preventive measures, which can impact the spread and severity of infectious diseases such as COVID-19. Further, individuals residing in rural counties, characterized by socioeconomic challenges such as higher poverty rates, limited access to healthcare services, and lower educational attainment, have been found to be particularly vulnerable to COVID-19 (Marmot and Allen, 2020; Khanijahani et al., 2021). Recent studies in US have highlighted that counties with higher social deprivation and income inequality experienced higher COVID-19 associated morbidity and mortality (Ossimetha et al., 2021; Tan et al., 2021). In particular, African-Americans residing in low-income neighborhoods as well as counties with a higher percentage of Hispanic residents reported disproportionately higher rates of COVID-19 and more severe outcomes (Egede and Walker, 2020; Kim and Bostwick, 2020; Islam et al., 2021). However, variations in study timelines, population demographics, risk factors assessed, and endpoints has resulted in mixed findings among studies investigating the influence of demographic and socioeconomic risk factors on COVID-19 incidence rates (Martins-Filho et al., 2021).

Florida, one of the most populous states in the US, is notable for its unique demographic composition, high rates of immigration, recognition as a popular tourist destination, and climate. The diverse population, ranging from densely populated urban

medRxiv preprint doi: https://doi.org/10.1101/2024.01.26.24301823; this version posted January 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

centers to sparsely populated rural counties, contributes to a complex web of potential influences on disease transmission and outcomes (Reyes et al., 2013; Neiderud, 2015). Considering the socioeconomic variation among counties in Florida, it is crucial to examine how these factors may have influenced COVID-19 outcomes. Understanding the relationship between demographic characteristics, socioeconomic factors, health indicators, and COVID-19 metrics in the context of Florida's counties may provide valuable insights for public health interventions and policies to mitigate the pandemic's impact and to structure the public health response in the post-pandemic phase.

Nevertheless, there remains a substantial gap in existing research, particularly in examining the extent to which rural counties in Florida have been disproportionately affected by COVID-19. In this study, we first investigated geographic clustering of COVID-19 outcomes. We then analyzed variation in case rate, mortality rate, and CFR across epidemic waves and assessed the association of demographic, socioeconomic, and health-related factors, and COVID-19 outcomes. Our results show that demographic, socioeconomic, and health-related differences explain a significant proportion of the variation in COVID-19 outcomes and that those inequalities disproportionately affected rural counties in Florida.

medRxiv preprint doi: https://doi.org/10.1101/2024.01.26.24301823; this version posted January 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

2.Methods

2.1 Map data

Florida Rural counties designations were obtained from the Florida Department of Health (FDOH) (FloridaHealth, 2023). The State of Florida defines rural as (i) a county with a population of 75,000 or less or (ii) a county with a population of 125,000 or less which is contiguous to a county of 75,000 or less or (iii) any municipality within a county as described above (FloridaHealth, 2023). County-level cartographic boundary Shapefiles were downloaded from the United States Census Bureau TIGER Geodatabase (USCensus, 2022).

2.2 COVID-19 Outcome Data

In this study, case rate, mortality rate and CFR were treated as dependent variables. County-level confirmed COVID-19 case and death data were extracted from the GitHub repository of The New York Times (https://github.com/nytimes/covid-19-data), based on reports from state and local health agencies (The New York Times, 2021). Population statistics were obtained from the United States Census Bureau (https://www2.census.gov/geo/docs/reference/ua/). Case rate represents number of cumulative cases per 100,000 population and mortality rate the number reported cumulated deaths from COVID-19 per 100,000 population as of December 2022. Each county's CFR was determined by dividing the cumulative deaths associated with COVID-19 by cumulative COVID-19 cases (Cao et al., 2020). Because the CFR is a ratio rather than an incidence or mortality rate, it is less affected by differences in testing rates between locations (Kathe and Wani, 2021).

For our analysis, we categorized the SARS-CoV-2 pandemic into five epidemic waves, each coinciding with the emergence of new variants. The delineation of these waves was based on the epidemic curve and prevalence of notable variants of concern during specific date ranges. The initial wave occurred from June 1, 2020, to September 30, 2020, corresponding to the onset of the pandemic and the original strain. Subsequent waves included the second wave from October 1, 2020, to May 30, 2021, associated with the Alpha variant; the third wave from July 1, 2021, to October 31, 2021, marked by

medRxiv preprint doi: https://doi.org/10.1101/2024.01.26.24301823; this version posted January 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

the Delta variant; the fourth wave from December 1, 2021, to February 28, 2022 characterized by the emergence of the Omicron variant (mainly BA.1, BA.2 subvariant); and the fifth wave spanned from April 1, 2022, to October 31, 2022, comprised largely of Omicron sub-variants BA.4 and BA.5.

2.3 Covariates

Eighteen demographics, socioeconomic, health related factors were included in the study (**Table 1**). Data were collected from the Florida Community Health Assessment Resource Tool Set (CHARTS, www.flhealthcharts.gov). CHARTS is developed and maintained by the FDOH, Division of Public Health Statistics and Performance Management, Bureau of Community Health Assessment. Different data population estimates, behavioral risk factors, health care providers, reportable diseases, deaths, and more. CHARTS is updated continuously throughout the year as new data becomes available. It offers an inventory of data sources together with an estimate of when each type of data will be updated. It also has an update log that records changes and additions to the platform in chronological order. (FLHealthCHART, 2023).

Data for the percent of the population in the county within rural blocks in Florida was collected from the United States Census Bureau. COVID-19 vaccination data for Florida counties were accumulated from the Center for Disease Control and Prevention (CDC). Information from multiple vaccine partners is included in this data set, including federal entity facilities, retail pharmacies, long-term care facilities, dialysis centers, Federal Emergency Management Agency, and Health Resources and Services Administration partner sites (CDC, 2023). All data used in this study were retrieved from publicly available databases, and no patient information was involved. Age-adjusted deaths from all causes per 100,000 population were collected from Florida CHARTS. Age-adjusted rates, which accounts to variation in the age distribution within the population of interest, were calculated as described here (https://www.flhealthcharts.gov/charts/OpenPage.aspx?tn=665).

medRxiv preprint doi: https://doi.org/10.1101/2024.01.26.24301823; this version posted January 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

2.4 Statistical Analysis

All variables were reported as mean and standard deviation. Coefficient of variation (CV) of each variable was calculated to assess the relative dispersion or variability. Spread and collinearity of the predictor variables were assessed through histograms, bivariate scatterplots, and Spearman correlation coefficients using R-package *psych* v2.3.6 (Revelle, 2023). We used the Shapiro–Wilks test method, Q–Q graph, to evaluate the normality. Correlation coefficients were calculated for each predictor with the outcome using Spearman or Pearson correlation according to data type and distribution. Florida county-level case rates, mortality rates, and CFR were visualized in the map using ggplot2 v3.4.3 (Wickham et al., 2016) and *sf* v1.0.14 (Pebesma and Bivand, 2023) package in R v4.3.1.

We used multivariate ordinary least squares (OLS) regression model to examine the relationship of predictor variables and COVID-19 outcomes. Then, we built multivariate models for each outcome using a stepwise backward procedure, starting with the fully saturated model. Stepwise regression approaches include an excellent screening mechanism for independent variables, which helps to reduce the impact of statistically insignificant factors on the regression equations (Silhavy et al., 2017). Multiple models are built in each iteration by dropping each variable simultaneously. The Akaike Information Criteria (AIC) of the models is also computed, and the model that yields the lowest AIC is retained for the next iteration. AIC is used to assess model fit and lower AIC indicates better model quality for prediction. For the final model, R^2 was calculated. The value of R^2 represents a model's ability to explain the variance in the dependent variable and therefore a higher R^2 value indicates the better performance of the model. This value is calculated based on a comparison between the predicted and observed values of the dependent variables (Middya and Roy, 2021).

For multivariate analysis, we used the Variance Inflation Factor (VIF) to assess multicollinearity. The VIF was computed for each covariate in the final model to assess the degree of collinearity and improve regression model accuracy (Middya and Roy, 2021). All statistical analysis and visualization were performed using R v4.3.1 (Team, 2021) with Rstudio v2023.06.1 (RStudio, 2020) statistical software.

medRxiv preprint doi: https://doi.org/10.1101/2024.01.26.24301823; this version posted January 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

2.5 Spatial autocorrelation

We calculated the global Moran *I* statistic to examine whether COVID-19 case rate, mortality rate, and CFR were spatially autocorrelated at the county level (Getis, 2009). The global Moran I statistic estimates the degree of geographically clustering among COVID-19 outcomes. The Moran I statistic value varies from -1 to 1. Positive values suggest positive spatial autocorrelation, indicating that counties with similar outcomes are in proximity. Negative values indicate negative spatial autocorrelation or dispersion, which means that counties with similar outcomes are far apart and values close to 0 indicate a random distribution (Bilal et al., 2021) . For spatial autocorrelation analysis, we utilized first-order Queen contiguity weights to determine which counties were neighbors. Counties were considered neighbors when they share one common point or vertex. (Ramírez-Aldana et al., 2020). Queen contiguity binary neighbor list and spatial weights matrix were constructed using R package *spdep* v1.2.8 (Pebesma and Bivand, 2023)

We measured local spatial autocorrelation to identify spatial clusters for COVID-19 outcomes. Local spatial autocorrelation identifies four distinctive categories: High-High, Low-Low, High-Low, and Low-High. High-High regions denoted counties with a high value surrounded by neighboring counties sharing a similarly high value. Conversely, Low-Low areas indicated spatial homogeneity characterized by low-value counties encircled by neighboring counties with similarly low incidence. High-Low demonstrated a spatial relationship between high and low-value counties, while Low-High portrayed low-value counties bordered by high-value counties. Local spatial clusters of COVID-19 outcome was measured using GeoDa v1.22 software (Anselin et al., 2009)

We compared the final selected ordinary least squares (OLS) model and two kinds of spatial regression models. A spatial lag model (SLM) examines how the adjacent counties influence the COVID-19 outcome in a county. The spatial lag parameter (r) is an estimate of the relationship between the average logged period prevalence in nearby counties and the logged period prevalence in a single focus county. In contrast, in a spatial error model (SEM), the parameter (l) determines how well a county's ordinary least squares (OLS) residual correlates with the residuals of its neighbors. This spatial

medRxiv preprint doi: https://doi.org/10.1101/2024.01.26.24301823; this version posted January 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

error parameter (l) represents the strength of the association between the average residuals or errors in nearby counties and the residual or error of the current county (Sun et al., 2020). We used the R "lagsarlm" and "errorsarlm" functions from *spatialreg* package to perform spatial regression analysis (Bivand et al., 2013).

***Table 1**: Definition and Summary statistics of different variables and COVID-19 outcomes in 67 countries in Florida. Mean, standard deviation (SD) and Coefficient of Variation (CV) of variables are listed.*

Variable	Description	Mean	SD	CV
Response Variables (COVID-19 outcomes)				
Case rate	Confirmed cases per 100,000 population	30368	7163	23.59
Mortality rate	Reported deaths per 100,000 population	460.30	147.47	32.04
Case fatality rate	Case Fatality Rate in percentage	1.55	0.49	31.79
Demographic and socioeconomic factors				
Median age	Population median age	44.03	6.54	14.85
Aged	% of people age over 65 years	22.96	7.93	34.54
Income	Median household income in Doller	51290	10300	20.08
Diploma	% Individuals with a high school diploma (aged 25 years and older)	33.25	7.20	21.65
Rural population	% of the population in the county within rural blocks	37.50	32.26	86.02
Unemployment rate	Unemployment rate, % of labor force	6.64	1.53	23
Poverty rate	% of total population, poverty rate	15.18	5.10	33.6
Below poverty level	% of total population, individuals below poverty level	14.96	5.03	33.59
Health and behavioral factors				
Hospital beds	Hospital beds, rate per 100,000 population	209.14	141.73	67.77
Acute care beds	Acute care beds, rate per 100,000 population	173.37	111.53	64.33
Health Insurance	% of population, with any type of health care insurance coverage	82.61	4.31	5.22
Vaccinated one dose	% of population who received one dose of vaccine	68.45	17.37	25.38
Fully vaccinated	% of population fully vaccinated	58.24	14.95	25.67
Obesity	% of total population, who are obese	32.46	6.07	18.69
Diabetes	% of total population, who have ever been told they had diabetes	13.37	3.09	23.11
Smoking	% of total population, who are current smokers	19.14	5.08	26.55
Race and Ethnicity				
Black	% of Population Black (of total population)	14.86	9.38	48.24
Hispanic	% of Population Hispanic (of total population)	14.59	13.09	89.75

medRxiv preprint doi: https://doi.org/10.1101/2024.01.26.24301823; this version posted January 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

3. Results

3.1 Descriptive mapping of COVID-19 outcomes in Florida

Florida comprises 67 counties with a population of 21 million, where a higher percentage of the state's population (87.7%) lives in urbanized areas. By December 2022, Florida reported around seven million confirmed cases and around 84,000 COVID-19 associated deaths. Visualization of cumulative rates across the five waves showed geographical heterogenicity in COVID-19 case rate, mortality rate, and CFR across Florida counties. Seminole, Miami-Dade, and Broward Counties had the highest case rate, while Glades, St. Johns, and Sumter Counties exhibited lower rates (**Figure 1A**). Union, Suwannee, Citrus Counties had the highest death rate, whereas Orange, Monroe, and St. Johns Counties had the lowest mortality rate (**Figure 1B**). Citrus, Highlands, and Union counites had highest CFR and Leon, Orange, Monroe counties had lowest CFR (**Figure 1C**).

Table 1 presents the summary and descriptive statistics of COVID-19 outcome variables and independent variables, and the last column includes the Coefficient of Variation (CV), which informs dispersion or variability of the data. Mortality rate and CFR were more dispersed than case incidence rate. Demographic, socioeconomic, and health-related risk factors vary widely among counties. The percentage of people in the rural areas, hospital bed rate, Hispanic population, and poverty rate are the most heterogeneous factors.

medRxiv preprint doi: https://doi.org/10.1101/2024.01.26.24301823; this version posted January 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

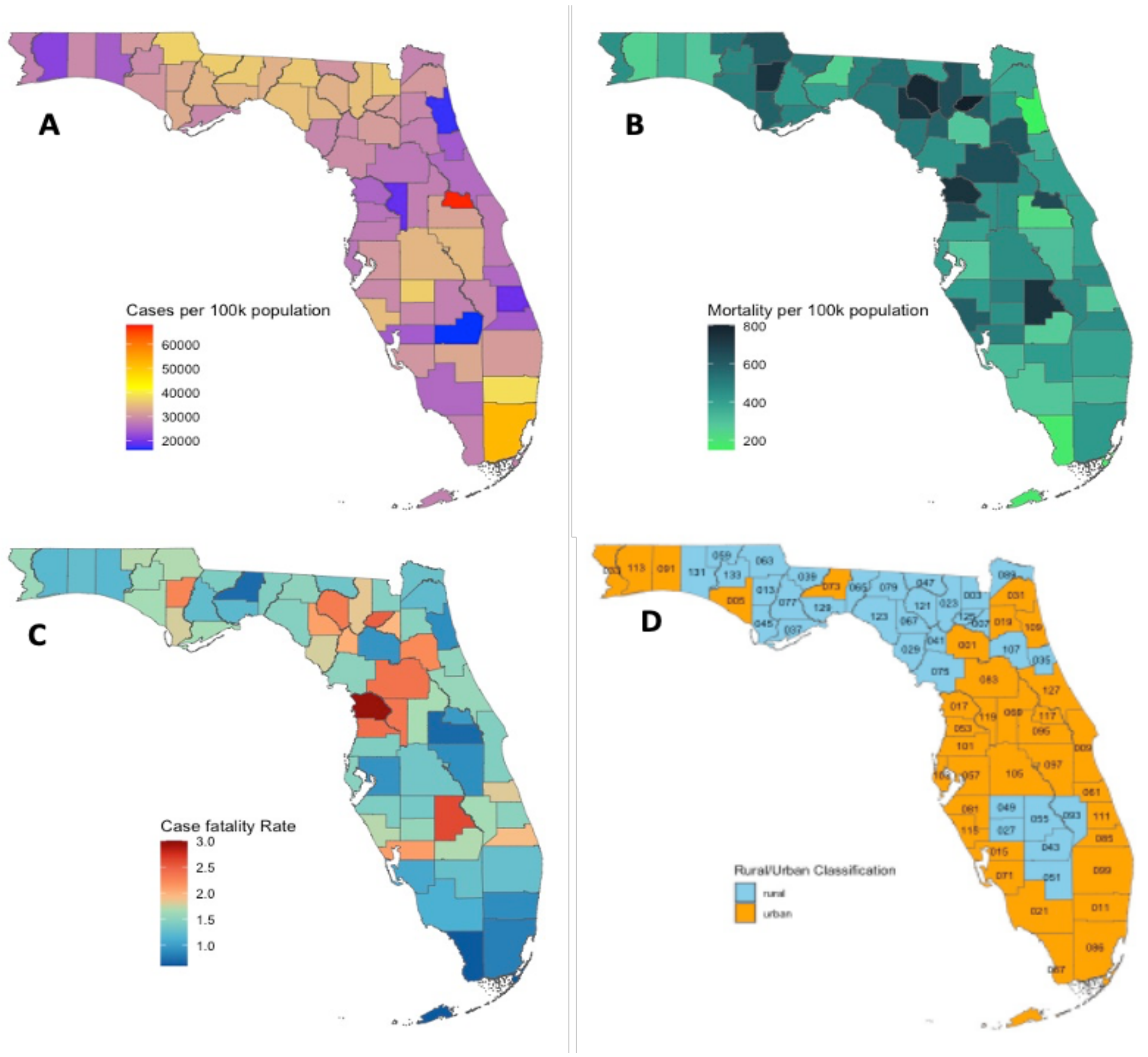


Figure 1: Statewide distribution of county-level COVID-19 confirmed cases per 100,000 population (A), mortality per 100,000 population (B), case fatality rate (C) in Florida from March 2020 to December 2022. (D) Florida map showing geographic distribution of urban and rural counties.

3.2 Spatial autocorrelation of COVID-19 outcome between Florida counties

Moran's I statistics were calculated to assess spatial autocorrelation in COVID-19 outcome variables among Florida counties. The results indicated significant but weak positive spatial autocorrelation for case rate (I = 0.12, p = 0.037), mortality Rate (I = 0.22, p =0.001), and CFR (I = 0.3, p < 0.00).

medRxiv preprint doi: https://doi.org/10.1101/2024.01.26.24301823; this version posted January 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

The Local Moran test was conducted to examine the spatial clustering in more detail, the "High-High" category signifies areas with high values surrounded by other counties with high values. For case rate High-High cluster in red, indicating Orange County with significant high case rate surrounded by counties with similarly high rates. Conversely, we observed a Low-Low cluster consisted of five counties including Flagler, Putnam, Marion, Citrus and Hernando County. This cluster and other Low-Low counties indicate that these counties with low case rate surrounded by counties with similarly low case rate (**Figure 2A**). Madison, Suwannee, Lafayette, Gilchrist and Columbia Counties with high mortality rate were surrounded by neighboring counties with high mortality rate. Counties with low mortality rate surrounded by counties with low mortality rate were Okaloosa, Collier, and Miami-Dade Counties (**Figure 2B**). For CFR we identified one High-High cluster consist of four counties: Marion, Citrus, Sumter, and Hernando counties with significantly high CFR surrounded by similar CFR and a Low-Low cluster consisted of four counties (**Figure 2C**).

medRxiv preprint doi: https://doi.org/10.1101/2024.01.26.24301823; this version posted January 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

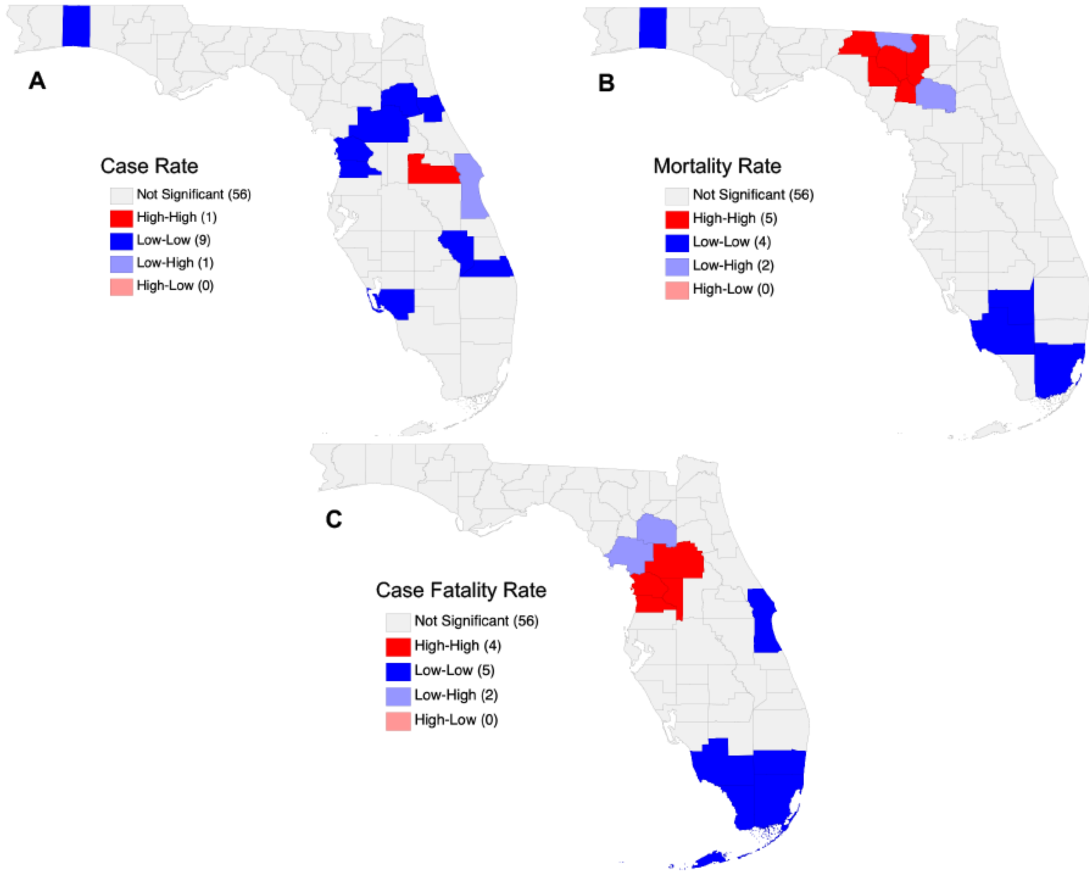


***Figure 2**: Spatial autocorrelation of COVID-19 outcomes by Florida county considering queen contiguity weights. Significant spatial clustering for case rate (B), mortality rate (C), and CFR (D) was observed. The High-High cluster (red) indicates counties with high values of a variable that are significantly surrounded by regions with similarly high values.* The *"Low-Low" cluster (blue) refers to counties where low COVID-19 outcome is surrounded by other counties with low COVID-19 outcome.*

medRxiv preprint doi: https://doi.org/10.1101/2024.01.26.24301823; this version posted January 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

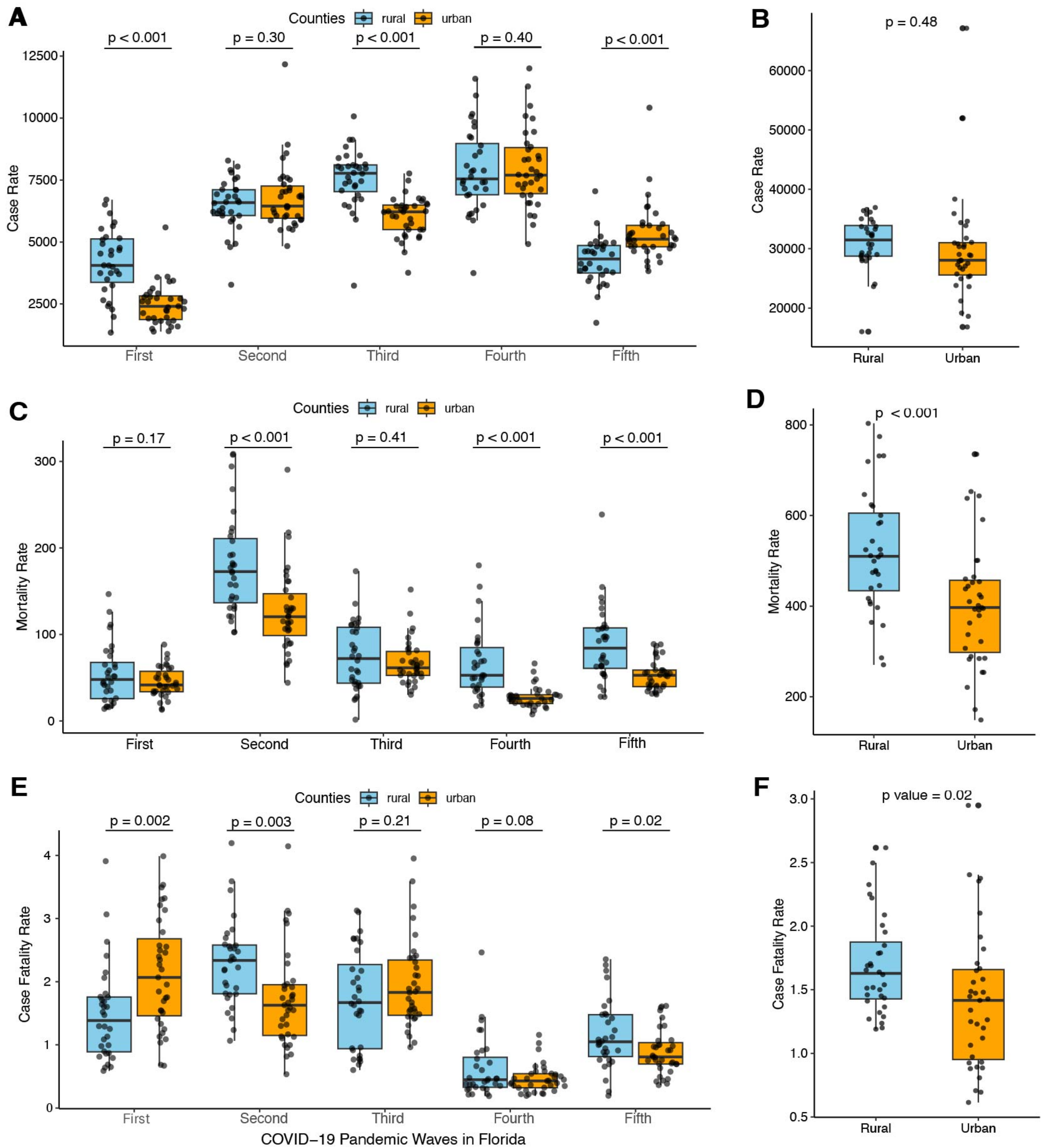


Figure 3: *Variation in case rate (top) and mortality rate (bottom) among urban (orange) and rural (blue) Florida counties across epidemic waves. Significance assessed using a t-test. Data are presented as box plots (center line at the median, upper bound at 75th percentile, lower bound at 25th percentile). Each dot point indicates the data of each county. Not significant=NS, ***P <0.001.*

medRxiv preprint doi: https://doi.org/10.1101/2024.01.26.24301823; this version posted January 26, 2024. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

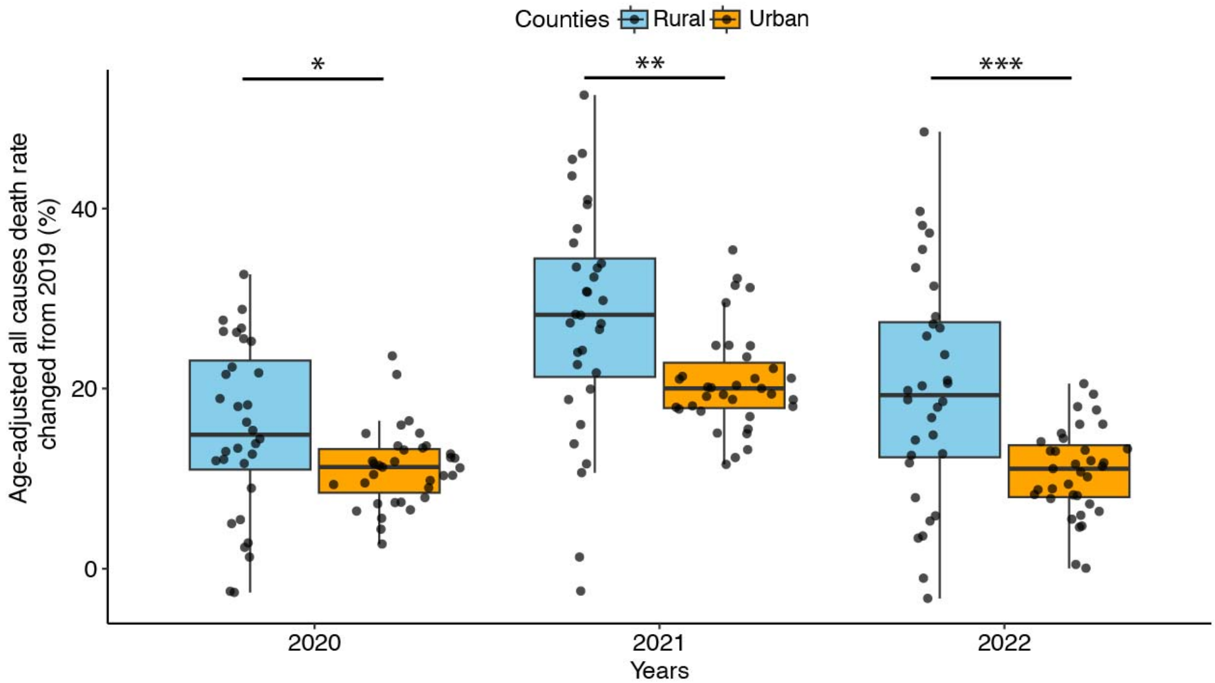
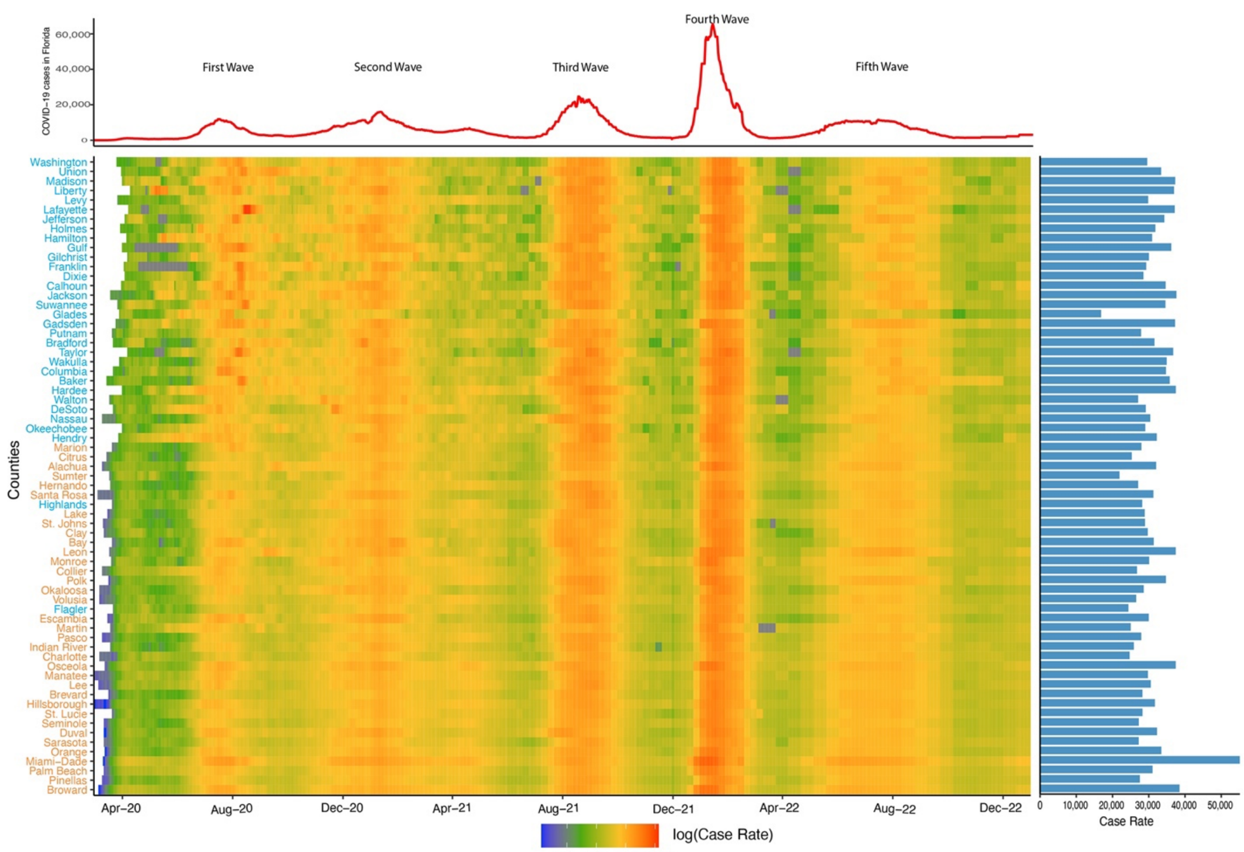


***Figure 4**: Increased age-adjusted all-cause mortality rates (per 100k population) in counties during pre-pandemic and pandemic times in Florida*



medRxiv preprint doi: https://doi.org/10.1101/2024.01.26.24301823; this version posted January 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Figure 5: *Trend in COVID-19 case rate among Florida counties. The heat map illustrates daily COVID-19 cases per 100,000 population with red coloration indicating higher values, and blue coloration indicating lower values. Counties are sorted by rural population percentage, with higher percentages at the top and lower percentages at the bottom. County names are color-coded, with orange indicating urban counties and blue representing rural counties, as defined by the Florida Health Department. The bar plot on the right side of the figure shows each county's COVID-19 case rate per 100,000 population from the pandemic's beginning to December 2022. At the top, a trend line displays the 7-day moving average of daily COVID-19 cases in Florida, with annotations marking five significant pandemic waves as of December 2022.*

3.3 Trends in COVID-19 case and mortality rates

We examined the case and mortality rates of COVID-19 Florida counties from the beginning of the pandemic through December 2022, focusing on variation between urban and rural counties. Our analysis revealed a notable pattern - higher case rates were observed in the first and third waves in rural counties compared to urban counties. However, no significant difference in COVID-19 case rates was observed between urban and rural counties for the second and fourth waves (***Figure 3A***). Furthermore, we observed a noteworthy temporal difference between urban and rural counties. Over the five waves, rural areas appeared to experience a more acute increase in COVID-19 case incidence, while urban areas faced a more protracted impact over time (***Figure 5***). Our analysis indicates that, despite the differences in epidemic trajectories, cumulative COVID-19 case rates did not differ significantly between urban and rural counties (***Figure 3B***).

medRxiv preprint doi: https://doi.org/10.1101/2024.01.26.24301823; this version posted January 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

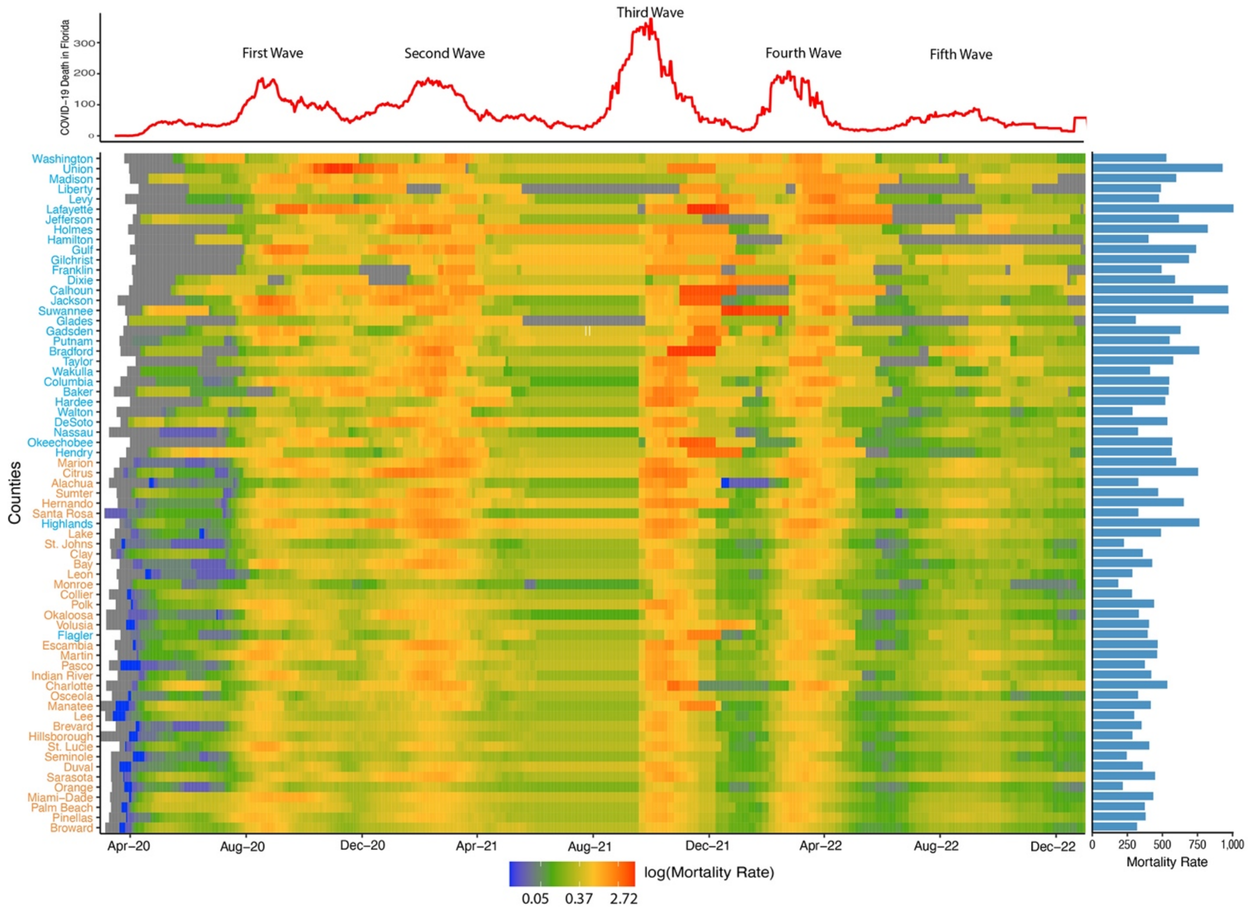


Figure 6: *Trend in COVID-19 mortality rates among Florida counties. The heat map illustrates daily COVID-19 deaths per 100,000 population with red coloration indicating higher values, and blue coloration indicating lower values. Counties are sorted by population percentage within rural blocks, with the county having the highest percentage at the top. County names are color-coded, with orange indicating urban counties and sky blue representing rural counties, as defined by the Florida Health Department. At the top, a trend line displays the 7-day moving average of daily COVID-19 associated deaths in Florida, with annotations marking five significant pandemic waves. The right side presents the county-wise mortality rate as of December 2022.*

We also analyzed mortality rates in Florida counties across five pandemic waves spanning from the onset of the COVID-19 pandemic to December 2022. We examined the variations in mortality patterns between urban and rural counties and variation in the impact of SARS-CoV-2 variants. We found that the mortality rate in rural counties was significantly higher than in urban counties during the second, third, and fifth waves of

medRxiv preprint doi: https://doi.org/10.1101/2024.01.26.24301823; this version posted January 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

the COVID-19 pandemic ***(Figure 3C, Figure 6***). While the difference in mortality rate between urban and rural counties during the first and fourth waves were not significant, the overall mortality rate was significantly higher among rural counties. We further observed variation in the case and mortality rate trends across the fives waves of the pandemic in Florida. In particular, the third wave (Delta variant) exhibited the highest mortality rate, while the fourth wave (Omicron variant) showed a significantly higher case rate. Next, we compared differences in CFR among urban and rural counties. CFR was significantly higher (p = 0.02) in rural counites than urban **(*Figure 3F*)**. However, when we assessed CFR difference during the five waves, we found inconsistent pattern among urban and rural counties. In the first wave, CFR was significantly higher in urban counties. Whereas CFR was higher in the urban counties in the second and fifth waves (**Figure 3E**). As decreased access to SARS-CoV-2 testing among rural counties may have resulted in an under estimation of cases rates, we conducted a sensitivity analysis finding that even after accounting for a 2.6% underreporting of case among rural counties, the CFR remained significantly higher (p = 0.04). Last, in our analysis of COVID-19's impact in Florida counties, we extended our investigation to include the age-adjusted all-cause mortality rate, comparing urban and rural counties before and after the pandemic. We calculated the percentage change in age-adjusted all-cause mortality for 2019 (pre-pandemic) and subsequent years (2020, 2021, 2022), finding a consistent and statistically significant trend: rural counties experienced a higher mean percentage change in age-adjusted all-cause mortality than urban counties across the analyzed years (**Figure 4**).

medRxiv preprint doi: https://doi.org/10.1101/2024.01.26.24301823; this version posted January 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

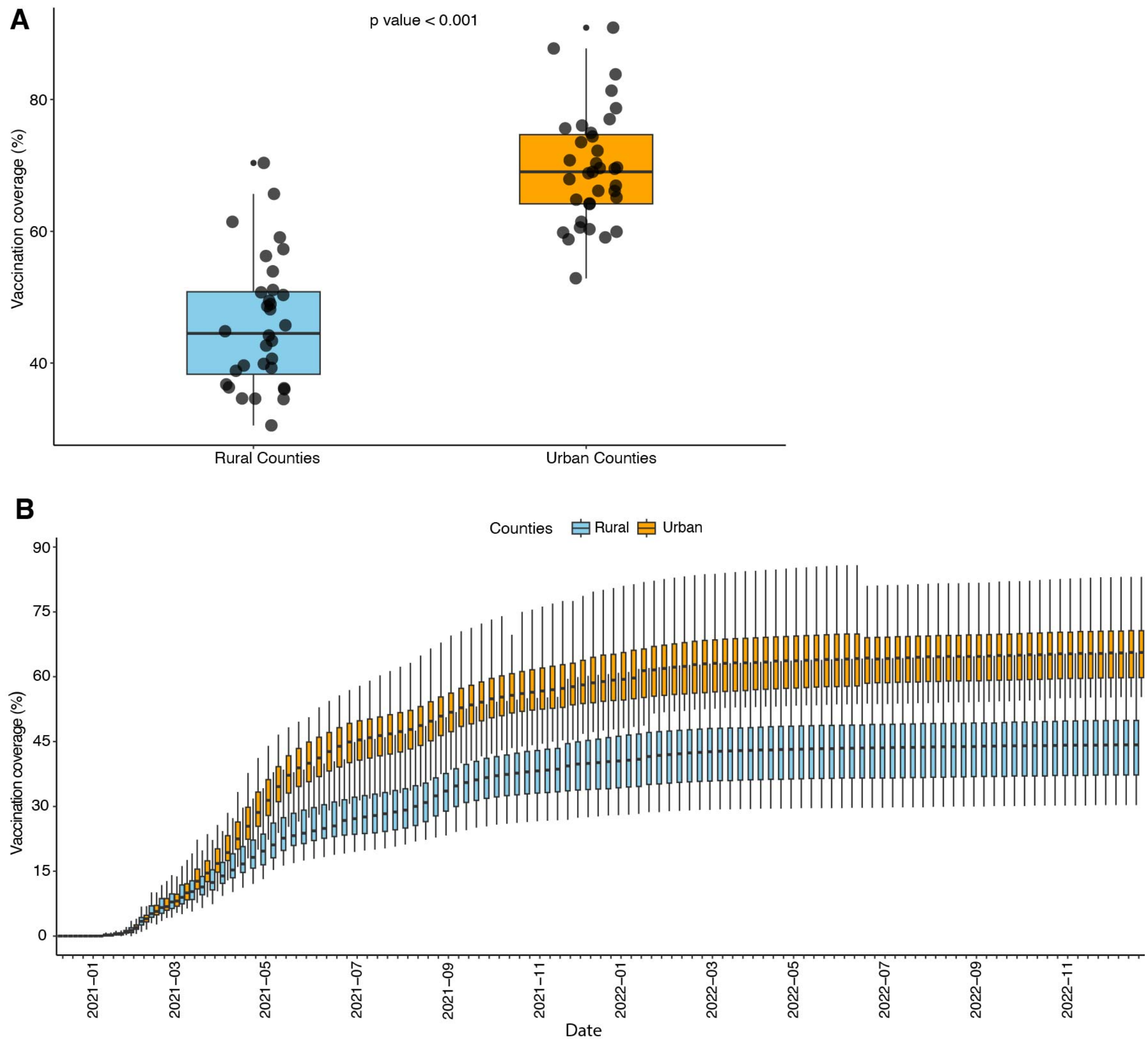


Figure 7: *The variation of vaccination coverage (A) in urban (orange) and rural (sky-blue) in Florida counties as of December 2022. (B) Trend of COVID-19 vaccination coverage in urban and rural counties.*

3.4 Vaccination disparities between urban and rural counties in Florida

To understand the role of vaccination on the trajectory of the COVID-19 pandemic in Florida, we analyzed the difference in vaccination coverage in urban and rural counties. We observed a significantly lower vaccination coverage in rural counties compared to urban counties (***Figure 7A***). We also investigated vaccination coverage trends throughout the study. A persistent pattern became evident: Urban counties (Mean:

B

medRxiv preprint doi: https://doi.org/10.1101/2024.01.26.24301823; this version posted January 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

69.48, 95% CI: 66.55-72.41) consistently had ~ 20% higher vaccination coverage than rural counties (Mean: 45.94, 95% CI: 42.39-49.50) (***Figure 7B***). However, there were several rural counties that were outliers, exhibiting higher vaccination rates than certain urban counties.

3.5 Correlation analysis

Correlation analysis was used to evaluate the degree of correlation between selected influencing factors and COVID-19 case rate, mortality rate, and CFR. A significant positive correlation was observed between case rate and individuals below poverty (r = 0.36, P = 0.003) and percent of the population of black race (r = 55, P < 0.001); significant negative correlation was observed between case rate and median age (r = -0.58, P < 0.001), percent of population insured (r = -0.32, p=0.008) and COVID-19 vaccination coverage (r = -0.26, P = 0.03). We also found significant positive correlation between mortality rate and percent of rural population (r = 0.53, P = <0.001), percent of population with obesity (BMI>30) (r = 0.34, P = 0.006), percent of population with diabetes (r = 0.43, P < 0.001), and percent of population who are smokers (r = 0.48, P < 0.001). In contrast, a significant negative correlation was observed between the mortality rate and median household income (r = -0.52, P <0.001) and vaccination coverage (r = -0.51, P < 0.001). The CFR is significantly correlated with median age (r = 0.43, P <0.001), rural population (r = 0.41, P <0.001), diabetes (r = 0.50, P <0.001), and smoking (r = 0.39, P = 0.001) (**Table 2**).

medRxiv preprint doi: https://doi.org/10.1101/2024.01.26.24301823; this version posted January 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Table 2: *Univariate correlation analysis with predictor variables for case rate, mortality rate, and CFR.*

Variable	Case Rate		Mortality Rate		Case Fatality Rate	
	Correlation Coefficient	P value	Correlation Coefficient	P value	Correlation Coefficient	P value
Median age	-0.58	<0.001***	0.11	0.367	0.43	<0.001***
Aged	-0.62	<0.001***	0.12	0.332	0.44	<0.001***
Median income	*-0.2*	*0.11*	*-0.52*	*<0.001****	*-0.47*	*<0.001****
High school diploma	0.19	0.123	0.55	<0.001***	0.49	<0.001***
Rural population	0.24	0.056	0.53	<0.001***	0.41	<0.001***
Unemployment Rate	-0.15	0.221	-0.18	0.143	-0.12	0.303
Poverty rate	0.36	0.003**	0.29	0.017*	0.15	0.237
Below poverty level	0.34	0.004**	0.26	0.032*	0.13	0.282
Hospital beds	-0.17	0.179	-0.24	0.049*	-0.13	0.282
Acute care beds	-0.13	0.289	-0.17	0.178	-0.09	0.471
Health insurance	-0.32	0.008**	-0.24	0.0.042*	0.02	0.819
Vaccination one dose	-0.26	0.035*	-0.51	<0.001***	-0.39	0.001**
Fully vaccinated	-0.26	0.035*	-0.51	<0.001***	-0.38	0.001**
Adult obesity	0.29	0.019*	0.34	0.006**	0.23	0.067
Diabetes	0.03	0.807	0.43	<0.001***	0.5	<0.001***
Smoking	0.1	0.423	0.48	<0.001***	0.39	0.001**
Black	0.55	<0.001***	-0.07	0.591	-0.24	0.047*
Hispanic	0.03	0.833	-0.29	0.017*	-0.33	0.006**

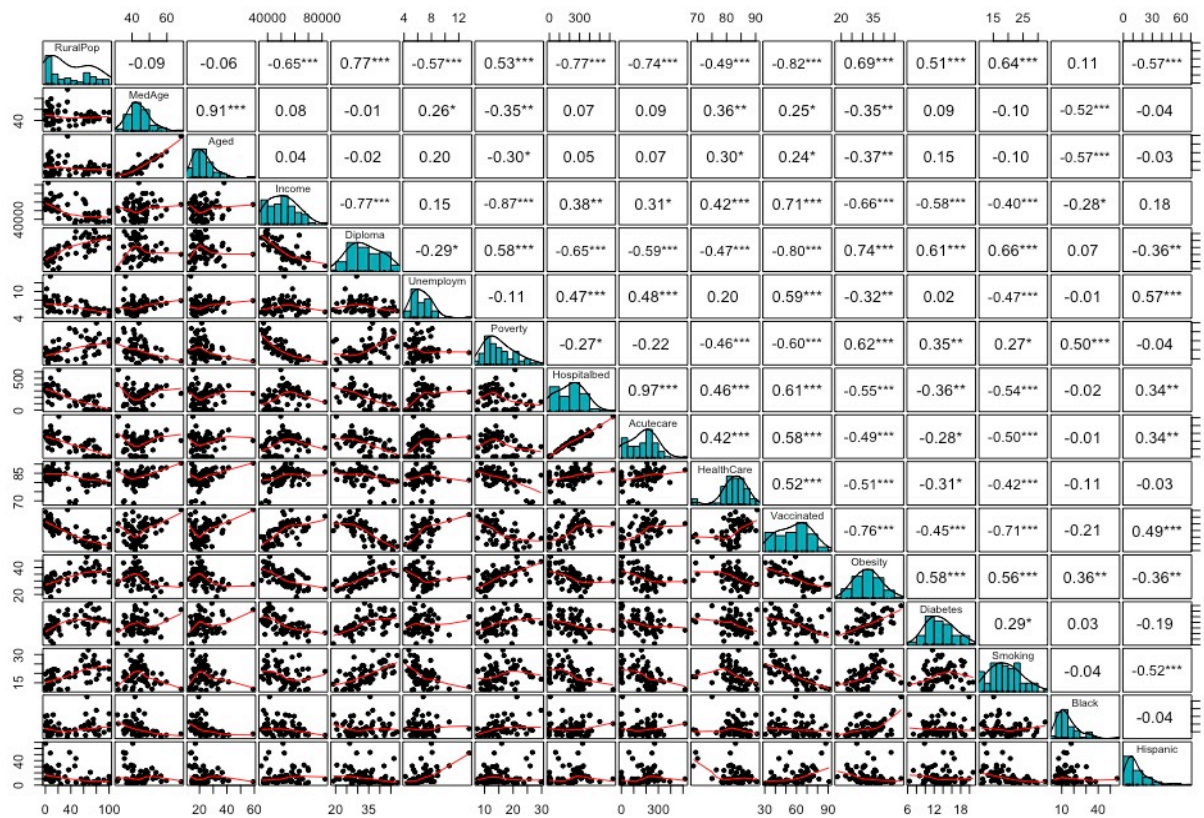
*Notes: * p < 0.05 significant at 0.05 level*
*** p < 0.01 significant at 0.01 level*
**** p < 0.001 significant at 0.001 level*

3.6 Multivariate analysis

The spread and collinearity of covariates were assessed through histograms, bivariate scatterplots, and Spearman correlation coefficients. Expectedly, the strongest correlation existed between median household income and poverty rate as well as the percentage of the population aged over 65 years and median age (**Supplementary Figure S1**). COVID-19 vaccination with one dose and fully vaccinated was also positively correlated, as a result, we retained fully vaccinated in the final analysis. Hospital and acute care beds, both proxies for healthcare access, showed a strong positive correlation. However, in the multivariate regression analysis, one of the

medRxiv preprint doi: https://doi.org/10.1101/2024.01.26.24301823; this version posted January 26, 2024. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

variables was omitted due to collinearity. Collinearity among covariates was further investigated in the multiple linear regression models by measuring VIF.



Supplementary Figure S1: *Correlation matrix of covariates. Spearman's correlation coefficient was used to measure the correlation between predictor variables. The diagonal shows the distribution of each variable in the histogram. On the top are the values of the correlations with significance measures (asterisks), and the bivariate scatter plots are shown at the bottom. p-value: *** < 0.001; ** <0.01; * < 0.05.*

medRxiv preprint doi: https://doi.org/10.1101/2024.01.26.24301823; this version posted January 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Supplementary Table S1. *Multivariable regression models of the association between socioeconomic characteristics and risk factors with the vaccination coverage for COVID-19. Vaccination coverage is measured per 100 people, and data was included up to December 2022.*

Variables	Adjusted R square	Estimate	Std. Error	t value	Pr(>|t|)	VIF
Intercept		-5.56	9.14	-0.61	0.545	
Median Income		< 0.001	< 0.001	6.57	< 0.001***	1.91
Unemployment Rate		2.03	0.69	2.93	0.004**	2.07
Rural population (%)	0.85	-0.15	0.04	-4.07	< 0.001***	2.69
% of black population		0.32	0.09	3.55	< 0.001***	1.37
Aged over 65 years		0.64	0.11	5.81	< 0.001***	1.43
% of Hispanic population		0.23	0.07	3.10	0.003**	1.79

In the final model for case rate, the percentage of the population in counties with rural blocks, percent of black population, percent of obese people, Hispanic population, and percent of people aged over 65 years remained in the final stepwise model. The R^2 of the final model including five these covariates was 0.42 (P < 0.01), indicating that they explain 42% of variation in the case rate in Florida. The low value of VIF (< 3) suggests that the final five variables do not have significant multi-collinearity and have substantial explanatory power to explain maximum model variation (***Table 3***).

For the mortality rate, the final selected model included median age, percent of people in the rural area, percent of black population, vaccination coverage, percent of population who smokes, and percent of Hispanic population, which explained 37% of variation of mortality rate in Florida. The result of collinearity diagnostics (VIF < 5) indicated that the collinearity between predictor variables was acceptable in this multiple linear regression model (***Table 3***).

Based on the stepwise multiple linear regression analysis, median age, acute care beds rate vaccination coverage, percent of the population who smokes, percent of the diabetes population were found to explain 65% of variation in the CFR in Florida (***Table 3***). The collinearity diagnostics (VIF < 3) suggested that the predictor variables are relatively independent of each other in the final model.

medRxiv preprint doi: https://doi.org/10.1101/2024.01.26.24301823; this version posted January 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

3.7 Spatial regression

We compared the final OLS models for case rate, mortality rate, and CFR with SLM and SEM to assess spatial association between the predictor covariates and outcome variables. SLM considers spatial lag effects, which reflect how the CFR in each county is influenced by the COVID-19 outcomes in neighboring counties. SEM addresses spatial autocorrelation, a measure of unexplained spatial variation in CFR. For the case rate and mortality rate, coefficients of the OLS and SEM models are closer than SLM (**Table 3**). However, AIC values were lowest for the OLS model compared to the spatial models, indicating that case and mortality rate can be explained better by the OLS model (**Table 4**).

We then explored SLM and SEM models to account for spatial dependencies in CFR. Notably, the estimates and significance of the covariates in the SLM model remained consistent with the OLS model, indicating that the spatial lag effects did not significantly alter the relationships between predictors and CFR. The SLM also estimated a spatial lag coefficient (rho) of 0.26, informing the extent to which neighboring counties' CFR values influence each other. As with SLM models, the predictor variables of SEM exhibit estimates and significance similar to the OLS model. Even after accounting for these predictor variables, the SEM estimated a spatial error coefficient (lambda) of 0.31, indicating persistent spatial variation in CFR (**Table 3**). The AIC values further informed our model selection, with the SLM having the lowest AIC (29.9), indicating the best trade-off between model fit and complexity. The SEM followed closely with an AIC of 31.2, while the OLS model had the highest AIC at 32.3 (**Table 4**). Therefore, spatial dependencies play a role in explaining variations in COVID-19 CFR among Florida counties.

***Table 3**: Multivariable regression models of the association between socioeconomic characteristics and risk factors with the COVID-19 outcome variables. The three outcome variables included are case rate, death rate, and case fatality rate. All are calculated per 100,000 people, and data was included up to December 2022.*

Dependent Variable	***Variables***	***OLS***				***Spatial lag model***			***Spatial Error model***		
		Adjusted R square	**Estimate**	**Std. Error**	**(VIF)**	**Rho**	**Estimate**	**Std. Error**	**Lambda**	**Estimate**	**Std. Error**
	(intercept)		44912***	6163			43657***	7560		44861***	5870.21
	Rural Population		96.05**	30.53	2.15		94.27***	29.82		96.73***	29.02
	Black		199.70*	82.88	1.34		195.17*	80.69		200.80*	78.80
Case Rate	Obese	0.42	-486.22**	168.63	2.32	0.04	-482.07**	161.49	-0.02	-485.75**	160.94
	Age over 65 years		-372.39***	101.07	1.42		-370.53***	96.66		-372.62***	95.99
	Hispanic		220.59***	58.65	1.3		219.72***	55.99		220.28***	55.55
	(intercept)		33.51	213.04			-25.39	209.78		17.37	201.87
	Median Age		10.95***	2.97	1.82		10.86***	2.81		11.61***	2.96
	Rural Population		-0.34	0.83	3.48		-0.50	0.79		-0.64	0.78
Mortality rate	Black	0.37	2.90	1.90	1.54	0.13	2.85	1.79	0.21	3.09	1.80
	Vaccination		-5.40**	1.97	4.18		-5.30**	1.86		-5.75**	1.89
	Smoking		10.03*	4.43	2.44		10.06*	4.16		10.46*	4.25
	Hispanic		2.53	1.39	1.59		2.66*	1.30		3.07*	1.40
	(intercept)		-0.81	0.45			-1.03*	0.43		-0.56	0.45
	Median Age		0.05***	0.01	1.27		0.04***	0.01		0.05***	0.01
	Acute Care Beds		0.001***	< 0.001	1.49		0.001***	< 0.001		0.001***	<0.001
Case Fatality Rate	Vaccination	0.65	-0.02***	0.004	2.96	0.26	-0.01***	<0.001	0.31	-0.02***	<0.001
	Smoking		0.02*	0.01	2.02		0.02*	0.01		0.02*	0.01
	Diabetes		0.04**	0.01	1.37		0.03**	0.01		0.03*	0.01

Notes: * p < 0.05 significant at 0.05 level
** p < 0.01 significant at 0.01 level
*** p < 0.001 significant at 0.001 level

medRxiv preprint doi: https://doi.org/10.1101/2024.01.26.24301823; this version posted January 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Table 4: Comparison of AIC values in OLS, SLM and SEM models for case rate, mortality rate and CFR

Outcome variables	**OLS**	**SLM**	**SEM**
Case Rate	1350.8	1352.7	1352.8
Mortality Rate	836.8	838.2	837.8
Case Fatality Rate	32.3	29.9	31.2

medRxiv preprint doi: https://doi.org/10.1101/2024.01.26.24301823; this version posted January 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

4. Discussion

Socioeconomically disadvantaged communities often bear a greater burden during epidemics (Meara et al., 2008; Ottersen et al., 2014), and characteristically, disparities were prevalent during the COVID-19 pandemic (Burström and Tao, 2020; Karmakar et al., 2021). In Florida, links between geographic factors and COVID-19 outcomes have previously been identified (Backer et al., 2022; Joshi et al., 2022; Khan and Odoi, 2023). Here, we further explore these associations by comprehensively investigating the interplay between demographic, socioeconomic, and public health factors and differences in case rate, mortality rate, and CFR. Through our analysis, we observed significant temporospatial variations in county-level COVID-19 outcomes, which are explained in part by variation in socioeconomic and health-related risk factors as well as vaccination coverage.

From the onset of the pandemic until December 2022, Florida's epidemic followed the US national trend; however, state level data aggregation obscured county-level variation. Subsequent comparison of the COVID-19 case rate, mortality rate, and vaccination coverage between rural and urban counties across pandemic waves revealed several notable trends. Rural counties experienced higher case rates in the first and third COVID-19 waves than their urban counterparts. Moreover, a concerning trend emerged in the mortality and case fatality rates. During the pandemic's second and fourth wave, rural counties experienced significantly higher mortality than urban counties. In addition, a significantly higher cumulative CFRs among rural counties compared to urban counterparts was observed, a finding that is consistent with other US-based studies (Pro et al., 2020; Iyanda et al., 2022). Variation in CFR was also observed across the five waves, highlighting differences in epidemic trajectories between rural and urban counties that corresponded with differences in vaccine uptake. Specifically, CFR remained elevated among rural counties during the fourth and fifth waves after the introduction of COVID-19 vaccines, which have been shown to be highly efficacious in reducing mortality (Mohammed et al., 2022). Mirroring the trend in CFR, all-cause mortality, which is robust to potential variation in COVID-19 mortality reporting among rural and urban counties, followed the same pattern. Together, these observations elucidate a broader and potentially longer-lasting impact of the pandemic

medRxiv preprint doi: https://doi.org/10.1101/2024.01.26.24301823; this version posted January 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

to population health. Detrimental collateral effects of the pandemic on the diagnosis and management of chronic health conditions have been described previously (Hacker et al., 2021), and these effects often disproportionally impact rural communities with higher rates of comorbidities. Here, we find that a number of county level predictors are associated with variation in COVID-19 outcomes.

Higher rates of SARS-CoV-2 transmission, limited access to tertiary healthcare, high prevalence of commodities, and decreased vaccine uptake may contribute to observed differences in COVID-19 outcomes. In general, rural populations tend to be un- or under-insured, older, more likely to live farther away from tertiary medical facilities and have underlying health conditions, which increase their likelihood for adverse outcomes (Ullrich and Mueller, 2021). Indeed, here we find that median income, adult obesity, smoking rates, and diabetes were associated with higher county-level mortality rates. Obesity and diabetes are well-recognized risk factors for COVID-19 severity and mortality (Dietz and Santos-Burgoa, 2020; Fang et al., 2020; Grasselli et al., 2020a; Kim et al., 2021), as adipose tissue may be directly or indirectly involved in interaction with SARS-CoV-2 pathogenesis (Kruglikov et al., 2020). Smoking also is associated with increased risk of poor COVID-19 outcomes due to the potential impact on lung health (Lacedonia et al., 2021; Razjouyan et al., 2022; Patanavanich et al., 2023). In univariate analysis, number of acute care and hospital beds and the proportion uninsured in a county were inversely associated with all COVID-19 outcomes, suggesting that access to care contributed to the observed disparities.

Race has also been a consistent risk factor associated COVID-19 outcomes throughout the pandemic. We found a significant positive association between the percentage of Black persons in a county with COVID-19 case rate and CFR. Racial disparities in COVID-19 outcomes have also been reported in multiple studies (Golestaneh et al., 2020; Hooper et al., 2020; Magesh et al., 2021), and the relationship between race and economic status is well established, with African Americans living mostly in low-income neighborhoods that are densely populated, leading to more frequent person-to-person contact and increased COVID-19 transmission (Phelan and Schneider, 1996; Shah et al., 2020). Furthermore, the African American populations have higher prevalence of

medRxiv preprint doi: https://doi.org/10.1101/2024.01.26.24301823; this version posted January 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

cardiovascular disease, obesity, diabetes and hypertension, which may contribute to higher symptomatic COVID-19 cases (Reyes, 2020; Yancy, 2020). Interestingly, while previous studies have reported increased COVID-19 case and mortality rates among Hispanic/Latinx population in US (Glance et al., 2021; Khanijahani, 2021; De Ramos et al., 2022), we did not observe a significant association between the percent Hispanic population in a county and increased morbidity and mortality in our study. However, we found that both the percentage Black and Hispanic population was significantly and positively associated with county-level immunization rates. As higher immunization rates were found to be protective in multivariable models of mortality and case fatality rates, this may partially explain the lack of association between Hispanic percentage and COVID-19 outcomes, while raising questions about why the percentage Black population remained associated with increased morbidity and mortality. Further research is required to better understand the factors to the observed variation in COVID-19 outcome in Latinx and Black populations. Overall, pandemic preparedness should recognize the substantial impact of systemic flaws within the structural framework of American society, encompassing issues related to healthcare accessibility, and the availability of community-level resources (Hawkins et al., 2020).

Our analysis also indicates that rural counties had slower vaccine uptake and significantly lower vaccination coverage than urban counties, which may have resulted in excess mortality among rural counties during later pandemic waves. Indeed, county vaccination rates remained significantly associated with mortality rates when controlling for a number of other social and demographic variables. These findings align with multiple recent studies conducted in the US (Hernandez et al., 2022; Saelee et al., 2022; Sun and Monnat, 2022). Several factors may contribute to vaccination coverage variation. First, health care access in rural counties remains a challenge. People in rural areas were more likely to complain about shortage of healthcare facilities and hospitals than those in urban areas. Access to COVID-19 vaccination might be difficult for rural residents in this situation (Ullrich and Mueller, 2021). Second, there are potential differences in knowledge and attitudes regarding the severity of COVID-19 infection and implementation of COVID-19 mitigation strategies across urban-rural areas, which are

medRxiv preprint doi: https://doi.org/10.1101/2024.01.26.24301823; this version posted January 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

influenced by different political ideologies and sociocultural identities (Ullrich and Mueller, 2021). Third, there is historically higher vaccine hesitancy in rural areas than in urban areas for routine vaccines, which likely contributed to reduced vaccination coverage in rural areas. For example, during the pandemic, rural adults in the U.S. were reported to be nearly three times more likely than urban adults to respond that they "definitely won't" get vaccinated against COVID-19 (Sparks et al., 2022). Our multivariate analysis of county level predictors of vaccination coverage indicates that all of these factors may have contributed to low vaccination rates. As we demonstrate that increased county-level vaccination coverage was associated with decreased mortality rate and CFR, a targeted approach is needed to close the gap in vaccination coverage between rural and urban communities by increasing vaccine confidence.

While this study provides valuable insights into the associations between various factors and COVID-19 outcomes, several limitations warrant consideration. The use of cross-sectional data limits the ability to infer causal relationships. Additionally, other unmeasured factors, such as variation in county-level interventions, population mobility, and cultural practices, may also shape COVID-19 outcomes. In Florida in particular, there was significant variation in county-level mitigation strategies such as non-pharmaceutical interventions due to the absence of state-level policies. Therefore, our analysis did not account for difference in mask mandates, stay-at-home orders, or school-based interventions. Last, an underlying assumption of our analysis is that case, death, and vaccine reporting was consistent among Florida counties. In particular, differences in testing capacity between rural and urban counties may have led to underreporting of cases among rural counties and a subsequent over-estimation of the CFR (Souch and Cossman, 2021). However, our sensitivity analysis suggests that our results are robust to modest differences in testing and case reporting, and our assessment of all-cause mortality in addition to COVID-19 mortality supports the observed disparities.

medRxiv preprint doi: https://doi.org/10.1101/2024.01.26.24301823; this version posted January 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

5. Conclusion

In conclusion, the findings from this study contribute to our understanding of the multifaceted nature of COVID-19 outcomes in Florida and accentuate the vulnerability of rural communities to health crises. Spatial variation and associations between demographic, socioeconomic, and health factors emphasize the need for tailored and equitable public health strategies addressing infectious diseases and the broader determinants of health. In particular, increasing vaccination rates among rural counties would result in a measurable impact. The insights gained from this study have the potential to inform evidence-based interventions aimed at reducing health disparities and enhancing overall population resilience in the face of public health crises.

Authorship contribution

Sobur Ali: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing
Taj Azarian: Conceptualization, Project administration, Resources, Supervision, Writing – review & editing

Funding

There was no funding associated with this study.

medRxiv preprint doi: https://doi.org/10.1101/2024.01.26.24301823; this version posted January 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

6. References

Anselin L, Syabri I, Kho Y (2009) GeoDa: an introduction to spatial data analysis. In: Handbook of applied spatial analysis: Software tools, methods and applications, pp 73–89. Springer.

Backer S, Rezene A, Kahar P, Khanna D (2022) Socioeconomic determinants of COVID-19 incidence and mortality in Florida. Cureus 14.

Bilal U, Tabb LP, Barber S, Diez Roux A V (2021) Spatial inequities in COVID-19 testing, positivity, confirmed cases, and mortality in 3 US cities: an ecological study. Ann Intern Med 174:936–944.

Bivand R, Hauke J, Kossowski T (2013) Computing the J acobian in G aussian Spatial Autoregressive Models: An Illustrated Comparison of Available Methods. Geogr Anal 45:150–179.

Braveman P, Gottlieb L (2014) The social determinants of health: it's time to consider the causes of the causes. Public Health Rep 129:19–31.

Burström B, Tao W (2020) Social determinants of health and inequalities in COVID-19. Eur J Public Health 30:617–618.

Cao Y, Hiyoshi A, Montgomery S (2020) COVID-19 case-fatality rate and demographic and socioeconomic influencers: worldwide spatial regression analysis based on country-level data. BMJ Open 10:e043560.

CDC (2023) COVID-19 Vaccinations in the United States,County. Centers for Disease Control and Prevention. Available at: https://data.cdc.gov/Vaccinations/COVID-19-Vaccinations-in-the-United-States-County/8xkx-amqh [Accessed August 1, 2023].

De Ramos IP, Lazo M, Schnake-Mahl A, Li R, Martinez-Donate AP, Roux AVD, Bilal U (2022) COVID-19 outcomes among the Hispanic population of 27 large US cities, 2020–2021. Am J Public Health 112:1034–1044.

Dietz W, Santos-Burgoa C (2020) Obesity and its implications for COVID-19 mortality. Obesity 28:1005.

Egede LE, Walker RJ (2020) Structural racism, social risk factors. and COVID-19—A dangerous convergence for Black Americans. N. Engl. J. Med ….

Fang L, Karakiulakis G, Roth M (2020) Are patients with hypertension and diabetes mellitus at increased risk for COVID-19 infection? lancet Respir Med 8:e21.

FLHealthCHART (2023) Community Health Assessment Resource Tool Set, Powered by Florida's Bureaus of Community Health Assessment and Vital Statistics. Available at: https://www.flhealthcharts.gov/charts/default.aspx [Accessed August 1, 2023].

FloridaHealth (2023) Florida Rural Counties (Last Upadate: February 2023). Available at: chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.floridahealth.gov/programs-and-services/community-health/_documents/Rural_Counties_February_2023.pdf [Accessed August 1, 2023].

medRxiv preprint doi: https://doi.org/10.1101/2024.01.26.24301823; this version posted January 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Getis A (2009) Spatial autocorrelation. In: Handbook of applied spatial analysis: Software tools, methods and applications, pp 255–278. Springer.

Glance LG, Thirukumaran CP, Dick AW (2021) The unequal burden of COVID-19 deaths in counties with high proportions of Black and Hispanic residents. Med Care 59:470.

Golestaneh L, Neugarten J, Fisher M, Billett HH, Gil MR, Johns T, Yunes M, Mokrzycki MH, Coco M, Norris KC (2020) The association of race and COVID-19 mortality. EClinicalMedicine 25.

Grasselli G, Greco M, Zanella A, Albano G, Antonelli M, Bellani G, Bonanomi E, Cabrini L, Carlesso E, Castelli G (2020a) Risk factors associated with mortality among patients with COVID-19 in intensive care units in Lombardy, Italy. JAMA Intern Med 180:1345–1355.

Grasselli G, Zangrillo A, Zanella A, Antonelli M, Cabrini L, Castelli A, Cereda D, Coluccello A, Foti G, Fumagalli R (2020b) Baseline characteristics and outcomes of 1591 patients infected with SARS-CoV-2 admitted to ICUs of the Lombardy Region, Italy. Jama 323:1574–1581.

Hacker KA, Briss PA, Richardson L, Wright J, Petersen R (2021) Peer reviewed: COVID-19 and chronic disease: the impact now and in the future. Prev Chronic Dis 18.

Hawkins RB, Charles EJ, Mehaffey JH (2020) Socio-economic status and COVID-19–related cases and fatalities. Public Health 189:129–134.

Hernandez I, Dickson S, Tang S, Gabriel N, Berenbrok LA, Guo J (2022) Disparities in distribution of COVID-19 vaccines across US counties: A geographic information system–based cross-sectional study. PLoS Med 19:e1004069.

Hooper MW, Nápoles AM, Pérez-Stable EJ (2020) COVID-19 and racial/ethnic disparities. Jama 323:2466–2467.

Islam N, Lacey B, Shabnam S, Erzurumluoglu AM, Dambha-Miller H, Chowell G, Kawachi I, Marmot M (2021) Social inequality and the syndemic of chronic disease and COVID-19: county-level analysis in the USA. J Epidemiol Community Heal 75:496–500.

Iyanda AE, Boakye KA, Lu Y, Oppong JR (2022) Racial/ethnic heterogeneity and rural-urban disparity of COVID-19 case fatality ratio in the USA: a negative binomial and GIS-based analysis. J racial Ethn Heal disparities 9:708–721.

Joshi M, Di Y, Bhattacharyya S, Chatterjee S (2022) Changes over Time in Association Patterns between Estimated COVID-19 Case Fatality Rates and Demographic, Socioeconomic and Health Factors in the US States of Florida and New York. COVID 2:1417–1434.

Karmakar M, Lantz PM, Tipirneni R (2021) Association of social and demographic factors with COVID-19 incidence and death rates in the US. JAMA Netw open 4:e2036462–e2036462.

Kathe NJ, Wani RJ (2021) Determinants of COVID-19 case fatality rate in the United States: spatial analysis over one year of the pandemic. J Heal Econ Outcomes Res 8:51.

Khan MM, Odoi EW (2023) Geographic disparities in COVID-19 testing and outcomes in Florida. BMC Public Health 23:1–13.

Khanijahani A (2021) Racial, ethnic, and socioeconomic disparities in confirmed COVID-19 cases and deaths in the United States: a county-level analysis as of November 2020. Ethn Health

medRxiv preprint doi: https://doi.org/10.1101/2024.01.26.24301823; this version posted January 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

26:22–35.

Khanijahani A, Iezadi S, Gholipour K, Azami-Aghdash S, Naghibi D (2021) A systematic review of racial/ethnic and socioeconomic disparities in COVID-19. Int J Equity Health 20:1–30.

Kim L, Garg S, O'Halloran A, Whitaker M, Pham H, Anderson EJ, Armistead I, Bennett NM, Billing L, Como-Sabetti K (2021) Risk factors for intensive care unit admission and in-hospital mortality among hospitalized adults identified through the US coronavirus disease 2019 (COVID-19)-associated hospitalization surveillance network (COVID-NET). Clin Infect Dis 72:e206–e214.

Kim SJ, Bostwick W (2020) <? covid19?> Social vulnerability and racial inequality in COVID-19 deaths in Chicago. Heal Educ Behav 47:509–513.

Kruglikov IL, Shah M, Scherer PE (2020) Obesity and diabetes as comorbidities for COVID-19: Underlying mechanisms and the role of viral–bacterial interactions. Elife 9:e61330.

Lacedonia D, Scioscia G, Santomasi C, Fuso P, Carpagnano GE, Portacci A, Mastroianni F, Larizza G, Sabato E, Profilo E (2021) Impact of smoking, COPD and comorbidities on the mortality of COVID-19 patients. Sci Rep 11:19251.

Magesh S, John D, Li WT, Li Y, Mattingly-App A, Jain S, Chang EY, Ongkeko WM (2021) Disparities in COVID-19 outcomes by race, ethnicity, and socioeconomic status: a systematic review and meta-analysis. JAMA Netw open 4:e2134147–e2134147.

Marmot M, Allen J (2020) COVID-19: exposing and amplifying inequalities. J Epidemiol Community Heal 74:681–682.

Martins-Filho PR, Quintans-Júnior LJ, de Souza Araújo AA, Sposato KB, Tavares CSS, Gurgel RQ, Leite DCF, de Paiva SM, Santos Jr HP, Santos VS (2021) Socio-economic inequalities and COVID-19 incidence and mortality in Brazilian children: a nationwide register-based study. Public Health 190:4–6.

Meara ER, Richards S, Cutler DM (2008) The gap gets bigger: changes in mortality and life expectancy, by education, 1981–2000. Health Aff 27:350–360.

Middya AI, Roy S (2021) Geographically varying relationships of COVID-19 mortality with different factors in India. Sci Rep 11:7890.

Mohammed I, Nauman A, Paul P, Ganesan S, Chen K-H, Jalil SMS, Jaouni SH, Kawas H, Khan WA, Vattoth AL (2022) The efficacy and effectiveness of the COVID-19 vaccines in reducing infection, severity, hospitalization, and mortality: a systematic review. Hum Vaccin Immunother 18:2027160.

Neiderud C-J (2015) How urbanization affects the epidemiology of emerging infectious diseases. Infect Ecol Epidemiol 5:27060.

Onder G, Rezza G, Brusaferro S (2020) Case-fatality rate and characteristics of patients dying in relation to COVID-19 in Italy. Jama 323:1775–1776.

Ossimetha A, Ossimetha A, Kosar CM, Rahman M (2021) Socioeconomic disparities in community mobility reduction and COVID-19 growth. In: Mayo Clinic Proceedings, pp 78–85. Elsevier.

Ottersen OP, Dasgupta J, Blouin C, Buss P, Chongsuvivatwong V, Frenk J, Fukuda-Parr S,

medRxiv preprint doi: https://doi.org/10.1101/2024.01.26.24301823; this version posted January 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Gawanas BP, Giacaman R, Gyapong J (2014) The political origins of health inequity: prospects for change. Lancet 383:630–667.

Patanavanich R, Siripoon T, Amponnavarat S, Glantz SA (2023) Active smokers are at higher risk of COVID-19 death: a systematic review and meta-analysis. Nicotine Tob Res 25:177–184.

Pebesma E, Bivand R (2023) Spatial data science: With applications in R. CRC Press.

Phelan TJ, Schneider M (1996) Race, ethnicity, and class in American suburbs. Urban Aff Rev 31:659–680.

Pro G, Hubach R, Wheeler D, Camplain R, Haberstroh S, Giano Z, Camplain C, Baldwin JA (2020) Differences in US COVID-19 case rates and case fatality rates across the urban-rural continuum. Rural Remote Health 20:6074.

Ramírez-Aldana R, Gomez-Verjan JC, Bello-Chavolla OY (2020) Spatial analysis of COVID-19 spread in Iran: Insights into geographical and structural transmission determinants at a province level. PLoS Negl Trop Dis 14:e0008875.

Razjouyan J, Helmer DA, Lynch KE, Hanania NA, Klotman PE, Sharafkhaneh A, Amos CI (2022) Smoking status and factors associated with COVID-19 in-hospital mortality among US veterans. Nicotine Tob Res 24:785–793.

Revelle W (2023) Psych: procedures for psychological, psychometric, and personality research. Northwestern University (2020). Evanston, Illinois, R Packag.

Reyes MV (2020) The disproportional impact of COVID-19 on African Americans. Health Hum Rights 22:299.

Reyes R, Ahn R, Thurber K, Burke TF (2013) Urbanization and infectious diseases: general principles, historical perspectives, and contemporary challenges. Challenges Infect Dis:123–146.

RStudio T (2020) RStudio: integrated development for R. Rstudio Team, PBC, Boston, MA URL http//www rstudio com.

Saelee R, Zell E, Murthy BP, Castro-Roman P, Fast H, Meng L, Shaw L, Gibbs-Scharf L, Chorba T, Harris LQ (2022) Disparities in COVID-19 vaccination coverage between urban and rural counties—United States, December 14, 2020–January 31, 2022. Morb Mortal Wkly Rep 71:335.

Shah M, Sachdeva M, Dodiuk-Gad RP (2020) COVID-19 and racial disparities. J Am Acad Dermatol 83:e35.

Silhavy R, Silhavy P, Prokopova Z (2017) Analysis and selection of a regression model for the use case points method using a stepwise approach. J Syst Softw 125:1–14.

Souch JM, Cossman JS (2021) A commentary on rural-urban disparities in COVID-19 testing rates per 100,000 and risk factors. J Rural Heal 37:188.

Sparks G, Hamel L, Kirzinger A, Stokes M, Brodie M (2022) COVID-19 Vaccine Monitor: Differences in Vaccine Attitudes Between Rural, Suburban, and Urban Areas.

medRxiv preprint doi: https://doi.org/10.1101/2024.01.26.24301823; this version posted January 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Sun F, Matthews SA, Yang T-C, Hu M-H (2020) A spatial analysis of the COVID-19 period prevalence in US counties through June 28, 2020: where geography matters? Ann Epidemiol 52:54–59.

Sun Y, Monnat SM (2022) Rural-urban and within-rural differences in COVID-19 vaccination rates. J Rural Heal 38:916–922.

Tan AX, Hinman JA, Magid HSA, Nelson LM, Odden MC (2021) Association between income inequality and county-level COVID-19 cases and deaths in the US. JAMA Netw open 4:e218799–e218799.

Team RC (2021) R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. (No Title).

The New York Times (2021) Coronavirus (Covid-19) Data in the United States. New York Times Available at: https://github.com/nytimes/covid-19-data [Accessed July 10, 2023].

Ullrich F, Mueller K (2021) COVID-19 cases and deaths, metropolitan and nonmetropolitan counties over time (update). Policy File 2020.

USCensus (2022) United States Census Bureau, Cartographic Boundary Files. Available at: https://www.census.gov/geographies/mapping-files/time-series/geo/cartographic-boundary.html [Accessed August 1, 2023].

Wickham H, Chang W, Wickham MH (2016) Package 'ggplot2.' Creat elegant data Vis using Gramm Graph Version 2:1–189.

Yancy CW (2020) COVID-19 and african americans. Jama 323:1891–1892.