The COVID-19 Pandemic: A Comparison of COVID-19 Test Positivity Rate by County

PHC6701 Project 3: Constructing a COVID-19 Data Analysis Pipeline

Seyedeh (Jasmine) Alemohammad, Natalie Goulett, Emdadul Haque, Michaela E. Larson, Ivan Pachon

2024-04-26

Table of contents

# Introduction

Coronaviruses (CoV) are a related group of RNA viruses that cause respiratory tract infections in humans.1 In 2019, a novel CoV, severe acute respiratory syndrome (SARS)-CoV-2, and its related disease, coronavirus disease 2019 (COVID-19), emerged as a potentially new global health threat. COVID-19 was first detected in China in December of 2019, but quickly spread to other countries across the globe. By the end of January of 2020, the World Health Organization (WHO) had declared a Public Health Emergency of Concern, and on March 11, 2020, the WHO officially declared COVID-19 a pandemic.2 As a group of viruses, CoVs range in severity. For example, Middle East respiratory syndrome-related (MERS)-CoV can kill up to 30% of people who are infected, whereas the common cold is relatively harmless.3 COVID-19 was most similar to the SARS-CoV outbreak in 2003, as it was able to cause both upper and lower respiratory tract infections. The pathogenicity of SARS-CoV-2 was impacted by both intrinsic viral factors and host-virus interactions. Virologically, SARS-CoV-2 was an evolutionarily more contagious and aggressive CoV given its affinity for ACE2 (a cellular entry receptor) that is highly expressed in the nasal epithelial cells, lungs, and nasal epithelial cells. Host factors that were critical for SARS-CoV-2 transmission and COVID-19 progression were related to existing insults to immune functioning, including the presence of other chronic illnesses such as diabetes and asthma, and immunosenescence due to aging and lifestyle factors.1

As of April 13, 2024, COVID-19 had claimed 1,188,991 lives in the United States alone.4 The death toll, especially during the initial stages of the pandemic, was largely due to poor emergency/disaster preparedness and surveillance systems, and incongruent responses across the United States that varied by state and sometimes by county. Surveillance systems that monitor the trends in COVID-19 test positivity rate and admissions and hospitalizations of infectious diseases like SARS-CoV-2 are critical for informing public health officials’ decisions. Early indicators of spread such as these can provide timely information to guide response efforts, including resource allocation, areas of particular need (“hotspots”), and informing policies and guidance related to restrictions (e.g., masking, business closures) to limit transmission, morbidity, and mortality. As such, the goal of this report was to examine key early indicators of COVID-19 spread at a community level to understand how our response to the pandemic may have changed over time.

## COVID-19 Surveillance Data

In response to the need for better surveillance during the COVID-19 pandemic, the Data Strategy and Execution Workgroup of the White House’s COVID-19 response team, made up of experts from the Department of Health and Human Services and the Centers for Disease Control and Prevention, began publishing Community Profile Reports (CPR) in December of 2020.5,6 The CPRs contains data on COVID-19 metrics for all regions, states, core-based statistical areas (CBSAs), and counties in the United States from December 17th, 2020 until February 23rd, 2023. Specifically, the CPRs include information about recent COVID-19 outcomes across the past 7 days (e.g., hospital admissions, mortality, test positivity rate, vaccinations), and additional context at the regional, state, county, and CBSA level (e.g., percent of people living in poverty, social vulnerability index scores, and percent of people that are racial/ethnic minorities). At irregular intervals, the CPR reported snapshots of recent COVID-19 outcomes from the last 7 days and how these outcomes changed from week to week. We used a set of the CDC’s CPR data that ranged from December 17, 2020 to July 07, 2022 from the `COVID Project’, which was created by a team of researchers from Florida International University and made available on GitHub.7 To quantify the strain that the COVID-19 pandemic placed on four specific counties in Florida, we were interested in the following metrics: total number of COVID-19 tests administered per day (reverse transcriptase-polymerase chain reaction [RT-PCR] tests and/or nucleic acid amplification tests [NAAT]), the proportion of positive COVID-19 tests, daily COVID-19 hospital admissions (confirmed and suspected), and number of inpatient beds occupied (hospital census).

# Methods

## Data Wrangling

385 workbooks from the COVID Project were cloned in RStudio using the COVID Project’s GitHub repository.7 Data for Miami-Dade, Broward, Palm Beach, and Suwannee County Florida were extracted across all workbooks regarding:

* Proportion of positive COVID-19 tests,
* Daily count of COVID-19 hospital admissions, and
* Daily census of hospital patients for any cause.

Each workbook was inspected to gather information on the column names used across all workbooks for the metrics on interest as they may have changed over time as the pandemic evolved. All data relevant for this report were extracted from the “Counties” sheet within each workbook. Utilizing the ::lubridate, ::tidyverse packages, we checked data on metrics of interest for missing values and harmonized them into six columns. Specifically, we extracted the “County” column that contained the county name from which the data originated. Data in this column was in the form: Miami-Dade County, FL. Information following the specific county name (county, FL) was removed. We then created a column that corresponded to the date of each file in YYYY-MM-DD format. The proportion of positive COVID-19 tests for the past 7 days was harmonized from three similar columns found in the workbooks: Viral (RT-PCR) lab test positivity rate - last 7 days (may exhibit anomalies due to delayed reporting), Viral (RT-PCR) lab test positivity rate - last 7 days (may be an underestimate due to delayed reporting), and NAAT positivity rate - last 7 days (may be an underestimate due to delayed reporting). To understand the context for the proportion of positive tests, we also harmonized 2 rows for the number of COVID-19 tests administered: Total RT-PCR diagnostic tests - last 7 days (may be an underestimate due to delayed reporting), and Total NAATs - last 7 days (may be an underestimate due to delayed reporting). To get the daily number of COVID-19 tests administered, we divided this column by 7. Daily count of COVID-19 hospital admissions was harmonized as the sum of Confirmed COVID-19 admissions - last 7 days and Suspected COVID-19 admissions - last 7 days, divided by 7. Lastly, to represent the daily census of hospital patients for any cause, we extracted the percent of inpatient beds occupied and multiplied it by the daily total number of inpatient beds (calculated by dividing the Total inpatient beds among hospitals reporting - last 7 days by 7).

## Examining the Proportion of Positive COVID-19 Tests Over Time

One metric used to gauge the early spread and prevalence of COVID-19 is the “positivity rate,” or the proportion of COVID-19 tests that are positive out of all tests performed. Metrics like the proportion of positive tests are important to track over time as they provide important information about the transmission of SARS-CoV-2, and can indicate (1) community spread of the virus, and (2) waves, or sustained peaks, of infections to monitor circulating strains and potential future impact in the healthcare system. To add context to the proportion of positive COVID-19 tests, we also created a metric to represent the number of tests performed. This metric was standardized to the mean number of tests administered throughout the reporting period for each respective county.

|  |
| --- |
| Hypotheses |
| 1. Peaks in the proportion of positive COVID-19 tests will align with known SARS-CoV-2 variant waves over the CPR reporting period; and 2. There will be differences in the trends of COVID-19 test positivity rate by county, and most notably between rural and urban counties. |

# Results

|  |
| --- |
| Figure 1: COVID-19 Test Positivity Rate in the Past 7 Days by Date |

To visualize the trends in the proportion of positive COVID-19 tests for the four counties over the reporting period, we constructed the above dot plot. In [Figure 1](#fig-positivity), we display the proportion of positive COVID-19 tests over the past seven days. The y-axis represents the average proportion of positive COVID-19 tests over the past 7 days. The x-axis displays the date the positivity rate was reported, ranging from December 17th, 2020 and July 7th, 2022. In [Figure 1](#fig-positivity), there are four counties being analyzed, all in the state of Florida: Broward, Miami-Dade, Palm Beach, and Suwannee County. Each county in the graph is symbolized by a color, highlighted in the legend. The data points in the graph represent each county’s COVID-19 positivity rate over the past 7 days. The transparency of the points signifies the standardized (to the mean) number of tests given in each respective county over the past 7 days. For each peak, there was an associated variant of the coronavirus that was the cause for most infections around that time. For example, between December 2021 and March 2022 the variant causing most infections was the Omicron variant, highlighted at the top of its peak. As we suspected, there were increases and decreases of the positivity rate over time, leading to “waves” in the graph. There also appear to be differences in the number of tests being performed and the test positivity rates between rural (Suwannee County) and urban (Miami-Dade, Palm Beach, and Broward Counties) counties. In particular, there was about a two week lag in the waves.

During the initial phase of the pandemic, which spanned from December 2020 to June 2021, various patterns emerged in the three urban countries. Test positivity rates in Miami-Dade and Palm Beach Counties increased dramatically between late December 2020 and early January 2021, most likely due to the appearance of the Alpha variation, COVID-19’s first major variant. Suwannee County, on the other hand, followed a different pattern throughout the same time period. Initially, test positivity rates rose in early 2021; however, this was followed by a significant and abrupt fall. This trend reflects a rapid fall in the proportion of positive tests compared to the total number of tests performed. Notably, the decrease in positive rates in Suwannee County occurred far faster than in Miami-Dade, Broward, and Palm Beach Counties, whose lines show slower descents and more fluctuations throughout this period. This variation highlights the complexities of managing and analyzing COVID-19 data at the county level. Another important note about the Alpha wave, is that this appears to be the point throughout the reporting period with the highest number of COVID-19 tests performed. This could be due to the massive mobilization of resources in the first few months of the pandemic that focused on gathering as much information about community level transmission of COVID-19 as possible.

Examining the Delta wave peaks from June to December 2021, we see similar trends as were observed in the Alpha wave across the 4 counties. Suwannee County had a significant increase in the proportion of positive tests beginning in June 2021, with its peak exceeding those of Miami-Dade, Broward, and Palm Beach counties. Notably during this time, Miami-Dade had the lowest proportion of positive tests but the highest number of tests performed during this wave compared to the other counties. During this wave we can also begin to see a pattern in testing emerge - more tests (indicated by darker points) are performed during the tail end (or declining portion) of each wave, which may indicate efforts to more accurately assess community-level spread and predict further strains on health care systems. This could also represent resource reallocation. For example, during COVID-19 peaks, more resources (money, personnel, etc.) could be allocated to hospital settings and treatment of COVID-19 infections as cases rise, and then back to testing as hospital systems become less overwhelmed. Furthermore, the amount of testing and reporting we see during the Delta wave was also likely impacted by Florida’s decision to stop daily reporting of certain COVID-19 metrics. This decision may have created a lag in data reporting when our source switched from the State of Florida Health Department to the CDC beginning in July 2021.8

Despite differences in peak height and timing, all three counties followed a similar pattern of rising positive rates throughout the Omicron wave, which lasted from December 2021 to March 2022. The proportion of positive COVID-19 tests was the highest during this wave compared to all other waves. The high rates of positive tests during this wave was likely due to the Omicron variant’s higher transmissibility compared to Delta and Alpha. Although more transmissible, Omicron infections have shorter incubation periods and result in less severe disease than Delta and Alpha, which might explain why the proportion of positive tests during this time doesn’t appear to last as long and also falls more rapidly. Moreover, the large increase in Omicron cases in Florida is likely also associated with the loosening of COVID-19 restrictions at the state level, the momentary ban of mask mandates in schools, and the lack of vaccine mandates which created circumstances conducive to viral spread.9

The final wave before the end of the reporting period was caused by the Omicron BA.2 variant. Although this variant was likely more transmissible than the other variants, it caused less serious illness. This wave was predicted to be the largest wave in the United States, but is not reflected as such in our data. This is partially because the COVIDProject data do not capture the entire wave, and the peaks in percent positivity are only visible for the three counties with the largest populations (Miami-Dade, Broward, and Palm-Beach Counties). Despite this, the proportion of positive tests that are observed in our four counties could be due to decreased testing due to the mild symptoms caused by this strain as well as the wind-down in testing overall (and the use of at-home tests).

# Discussion

The COVID-19 pandemic is the third most deadly modern-day pandemic behind the Spanish Flu pandemic of 1918 and the ongoing HIV/AIDS pandemic. In this report, we modeled the proportion of positive tests across four counties in Florida from December 17, 2020 to July 07, 2022. Consistent with the trends of pandemics, we saw increases and decreases in the proportion of positive tests over time that were associated with new virus variants, and may also be related to pandemic-related regulations, vaccination, and testing procedures. Generally, every time the proportion of positive tests dropped below 10%, it would eventually start increasing again (creating a wave pattern). Each of these waves lasted around 4 months. The patterns that we observed provided key insight to public health officials during the pandemic by acting as early indicators of community spread of COVID-19 and potential impending influxes in the use of hospital resources.

## Public Health Impact of the COVID-19 Pandemic

The COVID-19 pandemic unveiled serious flaws in our nation’s pandemic preparedness. As the virus spread, and more variants emerged, more people fell seriously ill, resulting in increased utilization of healthcare services. Both the government and healthcare sectors were not fully prepared to deal with the magnitude of the COVID-19 pandemic, both from a fiscal point of view (money, personnel, etc.) and in terms of materials (e.g., personal protective equipment, ventilators, etc.). As a result, we experienced shortages of personal protective equipment, nurses, doctors, and life-saving equipment in the healthcare system.

At the community level, the COVID-19 pandemic disproportionately affected racial and ethnic minorities. The mortality rate of COVID-19 for Black/African American individuals in the United States is estimated to be 33% higher than Hispanic/Latino individuals, and more than 50% higher compared to white individuals.10 Black/African American individuals were also over represented in the number of COVID-19 cases as well. Disparities in the morbidity and mortality of COVID-19 were likely the result of inequities in income, access to care, and health insurance rates that were exacerbated by the pandemic. Furthermore, prior to the pandemic, racial and ethnic minorities had higher rates of comorbid conditions that put them at greater risk for infection, and potentially worse COVID-19 outcomes.11

During the COVID-19 pandemic, rates of mental health conditions also increased presenting another urgent public health problem during an ongoing crisis. The World Health Organization estimates that the prevalence of anxiety and depressive disorders have increased 25% globally since the outbreak of the COVID-19 pandemic began.12 The mental health crisis during the COVID-19 pandemic was largely driven by the measures taken to control the spread of the virus. Most notably, social isolation due to lockdowns and quarantines to slow the spread and reduce exposure likely contributed most to this crisis. Other contributions to the increase in mental health conditions arose due to the closure of businesses that resulted in substantial loss of income and homelessness for already vulnerable populations, and an economic downturn. Mental health was also likely impacted by the initial lack of scientific information on how this novel coronavirus could affect human health, followed by a deluge of information that may not have been relevant or, at times, was completely incorrect - also termed an “infodemic.”13

The shortcomings of our response during the pandemic that may have inadvertently increased disparities in COVID-19-related morbidity and mortality, along with chronic diseases like depression and anxiety, highlight the need for better pandemic preparedness at a local and national scale.

## Limitations

Our report has a few limitations that are worth discussion. First, the healthcare infrastructure of Suwannee County differed from that of the other three counties that we examined. For example, Suwannee County does not have a full-service hospital within the county. Therefore, very little data was available regarding the strain of COVID-19 on its healthcare system, particularly COVID-19 admissions. This could have also affected the number of COVID-19 tests being performed in the county, and could have resulted in an underestimation of community-level COVID-19 infections. Second, our data only spans from December 2020 to July 2022, while the CDC’s Community Profile Report includes data until February 23rd, 2023. This limited temporal scope may affect the thoroughness and accuracy of our report, and miss any waves that could have occurred after the Omicron BA.2 wave around June of 2022. As a result, no conclusions can be made about the Omicron BA.2 variants’ impact on Suwannee County until more data are made available. Third, our report could benefit from the inclusion of national COVID-19-related and demographic data to better understand the impact of COVID-19 across various regions, demographic groups, and healthcare systems. Incorporating a broader range of data sources would enhance the validity and generalizability of our findings. Future research should explore additional factors influencing COVID-19 transmission dynamics, such as vaccination coverage, adherence to (non)pharmaceutical interventions, and other social determinants of health, which could offer valuable insights into the complex dynamics of COVID-19 transmission and response efforts at both local and national levels.

## Conclusion

Our report sheds light on the dynamic nature of COVID-19 transmission within four counties in Florida, offering valuable insights into the pandemic’s progression and the effectiveness of response measures. During the initial phase of the pandemic, disparities in testing and positivity rates emerged between urban and rural counties. While urban areas experienced dramatic increases in positivity rates, rural counties exhibited unique trends, suggesting varying impacts and response needs. Subsequent waves, notably the Delta and Omicron waves, presented challenges and opportunities for public health intervention. We observed fluctuations in testing practices, likely influenced by resource allocation and policy changes. Despite differences in peak heights and timing, all counties experienced surges during the Omicron wave, underscoring the variant’s heightened transmissibility. Notably, the Omicron BA.2 variant, while anticipated to be the largest wave, did not manifest as such in our data, possibly due to decreased testing and milder symptoms associated with this strain. This highlights the importance of continued vigilance and robust surveillance, even amidst evolving variants and changing testing dynamics. Our findings underscore the necessity of adaptive and localized strategies in pandemic response, considering the unique characteristics and needs of diverse communities. Furthermore, they emphasize the critical role of data-driven decision-making and the importance of ongoing monitoring to guide effective public health interventions.Looking ahead, our analysis underscores the importance of maintaining vigilance and adapting response efforts to address emerging challenges, such as variant dynamics and shifting testing practices. By leveraging data insights and fostering collaboration, we can enhance our ability to mitigate the impact of COVID-19 and safeguard public health.

# References

1. Umakanthan S, Sahu P, Ranade AV, Bukelo MM, Rao JS, Abrahao-Machado LF, Dahal S, Kumar H, Kv D. Origin, transmission, diagnosis and management of coronavirus disease 2019 (COVID-19). Postgrad Med J. 2020 Dec;96(1142):753-758. doi: 10.1136/postgradmedj-2020-138234. Epub 2020 Jun 20. PMID: 32563999; PMCID: PMC10016932.
2. Cucinotta D, Vanelli M. WHO Declares COVID-19 a Pandemic. Acta Biomed. 2020 Mar 19;91(1):157-160. doi: 10.23750/abm.v91i1.9397. PMID: 32191675; PMCID: PMC7569573.
3. Hui DS, Azhar EI, Kim YJ, Memish ZA, Oh MD, Zumla A. Middle East respiratory syndrome coronavirus: risk factors and determinants of primary, household, and nosocomial transmission. Lancet Infect Dis. 2018 Aug;18(8):e217-e227. doi: 10.1016/S1473-3099(18)30127-0. Epub 2018 Apr 18. PMID: 29680581; PMCID: PMC7164784.
4. Centers for Disease Control and Prevention. COVID Data Tracker. Atlanta, GA: U.S. Department of Health and Human Services, CDC; 2024, April 26. https://covid.cdc.gov/covid-data-tracker
5. White House COVID-19 Team, Joint Coordination Cell, Data Strategy and Execution Workgroup. COVID-19 Community Profile Report. 2023, July 25. https://healthdata.gov/Health/COVID-19-Community-Profile-Report/gqxm-d9w9/about\_data
6. Duvall, K. COVID-19 Open Data - Our Data is Your Data. 2024, December 18. https://healthdata.gov/stories/s/Blog-COVID-19-Open-Data-Our-Data-is-Your-Data/9c45-dw4i
7. Rodriguez, A., Odom, G. COVIDProject. 2021, August 03. https://github.com/annyrodriguez/COVIDProject
8. Dyer O. Florida loses legal battle to keep covid data secret BMJ 2023; 383 :p2419 doi:10.1136/bmj.p2419
9. Rivkees SA. The shifting impact and response to COVID-19 in Florida. Front Public Health. 2024 Feb 19;12:1351690. doi: 10.3389/fpubh.2024.1351690. PMID: 38439756; PMCID: PMC10909809.
10. Vasquez Reyes M. The Disproportional Impact of COVID-19 on African Americans. Health Hum Rights. 2020 Dec;22(2):299-307. PMID: 33390715; PMCID: PMC7762908.
11. Kabarriti R, Brodin NP, Maron MI, Guha C, Kalnicki S, Garg MK, Racine AD. Association of Race and Ethnicity With Comorbidities and Survival Among Patients With COVID-19 at an Urban Medical Center in New York. JAMA Netw Open. 2020 Sep 1;3(9):e2019795. doi: 10.1001/jamanetworkopen.2020.19795. PMID: 32975574; PMCID: PMC7519416.
12. Brunier, A., Drysdale, C. COVID-19 pandemic triggers 25% increase in prevalence of anxiety and depression worldwide. 2022, March 02. https://www.who.int/news/item/02-03-2022-covid-19-pandemic-triggers-25-increase-in-prevalence-of-anxiety-and-depression-worldwide
13. Kupcova I, Danisovic L, Klein M, Harsanyi S. Effects of the COVID-19 pandemic on mental health, anxiety, and depression. BMC Psychol. 2023 Apr 11;11(1):108. doi: 10.1186/s40359-023-01130-5. PMID: 37041568; PMCID: PMC10088605.