Stacey Wheeler
DATA512_A7
DUE December 14, 2021

DATA512 Final Report

Introduction

Seasonal influenza continues to infect thousands of people every year. For many, this preventable illness is deadly. In order to control the spread of COVID-19, most state and local governments in the US put in place restrictions on gatherings and public spaces, enforced masking, and eventually encouraged residents to get vaccinated against the virus. One of the positive effects of these restrictions was that in trying to control the spread of COVID-19, they also minimized the spread of the seasonal flu.

Before the pandemic, the idea of restricting people's access to public spaces and engagement with the economy would have been considered a ridiculous proposition. But COVID-19 forced us to go to extreme measures to keep ourselves and others safe. Now we have the chance to see how these restrictions changed the landscape for another common and preventable disease by comparing the rates of flu transmission during times with high restrictions to times where no such restrictions were in place.

Background/Related Work

The CDC publishes data and other information about the flu every season to inform medical professionals and the public. In 2018, they conducted a study which concluded that between 3% and 11% of people in the US have symptomatic flu illness each season (Key Facts About Influenza (Flu)). Of these, between 12,000-50,000 deaths occur each year from the flu (Disease Burden of Flu). Children under 18 are in the group most likely to contract the flu with twice the likelihood of catching it compared to adults 65+. The flu shot is widely believed to be the best preventative measure against catching and spreading the flu, but everyday hygiene techniques are also effective at preventing the spread.

I am interested in learning how the COVID-19 pandemic impacted transmission of influenza. It makes sense that the measures put in place to control the spread of COVID-19 were also effective at mitigating transmission of influenza. I will address two main research questions in this analysis:

1. How did masking policies change the progression of confirmed COVID-19 cases from February 1, 2020 through October 15, 2021?
2. How did masking, quarantining, and social distancing change the transmission of seasonal influenza during the 2020-2021 flu season? My hypothesis is that when looking at the weekly incidence of influenza during the '20-'21 flu season, we will see at least a statistically significant difference from the seasons prior.

Methodology

To address the first research question, I used the COVID-19 data from John Hopkins University and US State and Territorial Public Mask Mandates From April 10, 2020 through August 15, 2021 by County by Day | Data | Centers for Disease Control and Prevention. In order to create my visualization for the COVID and mask data, both datasets were filtered to include only information about Oakland, Michigan. Then I used the daily case count to calculate a 7 day moving average and the cumulative cases. I did the same processing for the dataset on daily COVID-19 deaths, and then I merged the two dataframes on the date. Using the county population I calculated the susceptible population for each day by subtracting the cumulative deaths and cumulative cases from the total population.

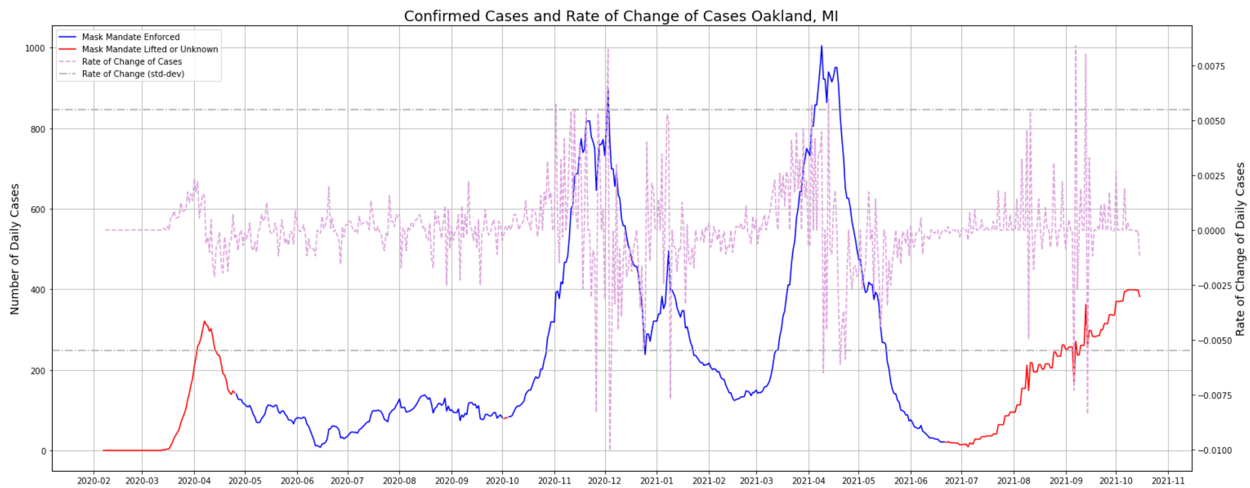
Some values of interest were calculated for each day using the following formulas:

Rate of Infection = (Cases moving average) / (Susceptible population)
Rate of Daily Infection = (Cases) / (Susceptible population)
Rate of Cases = (Daily difference in cases moving average) / (Total population)
Rate of Daily Cases = (Daily difference in cases) / (Total population)

To address my second research question, I downloaded a csv of flu data for the state of Michigan containing the number of specimens tested weekly, the number of specimens positive for influenza, and the percent positive by influenza type. The flu data sourced from the CDC is free and available for use as a public service.

I plotted the rate of flu transmission as a time series from 2017 through 2021. Visually I was able to see a clear change in flu transmission rates in the '20-'21 flu season than preceding years. To assess whether the change in flu transmission was actually statistically significant, I used a one-tail t-test to see if the transmission rate during COVID seasons was within a 95% confidence interval of the expected value.

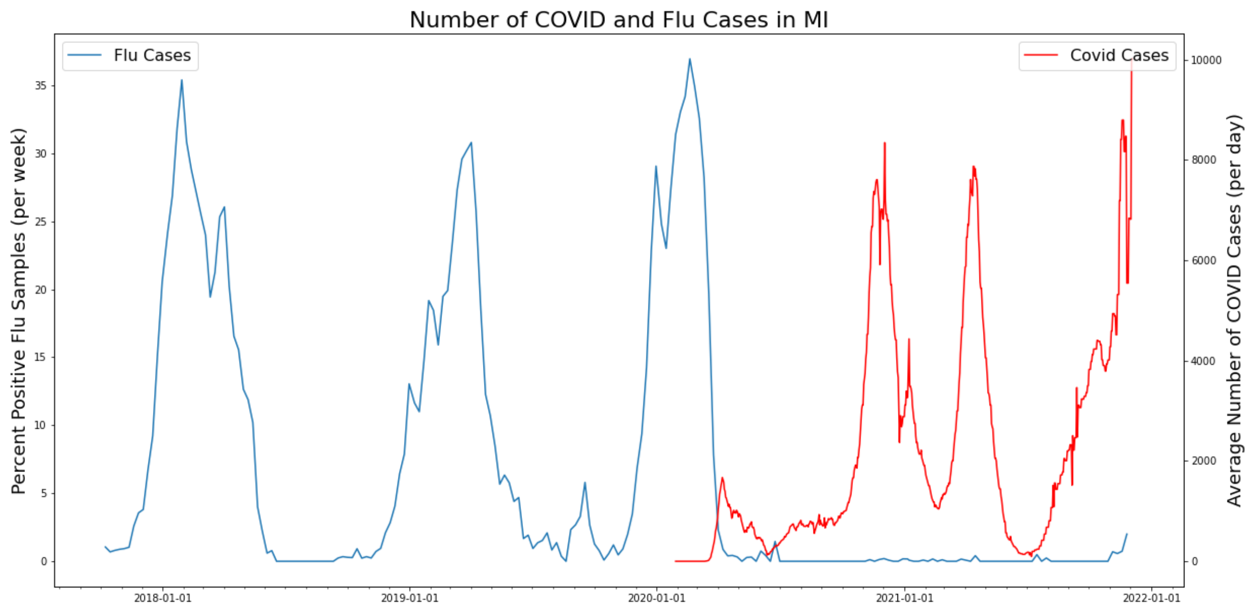
Findings



My visualization shows the number of confirmed daily COVID-19 cases in Oakland, Michigan from February 1, 2020 through October 15, 2021 (Number of Daily Cases on left axis). It also shows the rate of change in daily cases over this time frame (Rate of Change of Daily Cases on right axis). The confirmed case line is color coded according to when a mask mandate was in place: blue indicates that masking was enforced in the county, and red indicates that masking was not enforced or it is unknown. The rate of change of daily cases is shown with the plum line. A 95% confidence interval for the rate of change is also displayed with horizontal gray lines, which I calculated as 3 x std_dev(rate_of_cases).

The graph shows the cases 7 day moving average with the blue/red line and the 'Rate of Cases' with the plum. To understand the effectiveness of the masking policy in this county, we can observe what happened when the mandate was first enforced and when it was lifted. In April 2020, the county first put in place a masking policy that seems to have kept transmission at bay for approximately five months. October 2, 2020 was the first time the county lifted the mask mandate, and perhaps due to expected increase in transmission through the winter season it was immediately put back in place on October 5, 2020. Even though masks were required throughout the winter, the amount of gathering and travel that took place during the holidays is a confounding variable that caused huge spikes in transmission, as evidenced by the peaks in the blue lines as well as how many times the plum line goes outside the gray bars.

The peak in April 2021 may be due to spring break activities, but immediately following this peak there is a steep decline to the lowest daily case average since the pandemic began. Finally the mask mandate was lifted on June 22, 2021, and the cases began steadily rising again until the end of the time series on October 15, 2021. The last red portion of the graph is the strongest evidence for the efficacy of masking because even after vaccines became widely available, infections returned as soon as the mask mandate was lifted. Summer 2021 also coincided with the contagious Delta variant, so the new strain combined with no masking led to increasing infections.



This visualization shows the percent of samples that tested positive for the flu per week in Michigan from July 2017 through December 2021 (Percent Positive Flu Samples (per week) on left axis). It also shows the 7-day moving average number of confirmed daily COVID-19 cases in Michigan from February 2020 through December 2021 (Average Number of COVID Cases (per day) on right axis).

We can see visually that the 2020-2021 flu season had a significantly lower rate of incidence than in previous seasons. The ‘17-’18, ‘18-’19, and ‘19-’20 flu seasons all indicate a maximum weekly positive flu rate between 30-35% that occurred between January and March. The ‘20-’21 flu season shows no significant spike throughout the winter months that typically would see millions of cases throughout the US.

The null hypothesis for the t-test was that COVID restrictions had no effect on flu transmission. The result was p-value = 9e-15. The effect of the treatment (COVID restrictions) was so significant that we can confidently reject the null hypothesis and conclude that COVID restrictions did in fact have an effect of reducing the spread of the flu.

Discussion/Implications

The t-test confirms my assumption that COVID restrictions would also reduce the spread of the flu, but the question is whether or not that is useful information. COVID-19 is far more contagious and dangerous than the seasonal flu, even during particularly bad flu seasons. The restrictions put in place were intended to keep people safe and minimize the impact of the COVID. But people were not willing to accept that these restrictions were for their own safety and of others around them. We saw so much variance in the level of compliance with COVID restrictions across the US that it is difficult to tease out which measures helped the most and which did not at all. Compliance aside, our economy crumbled due to the sudden shift in consumer activity, and people suffered for it. Even though we could easily eliminate the flu each season with masking and social distancing, it is unrealistic to consider this a solution to the flu when we have decided as a country that flu is at an acceptable level of risk.

As time goes on, it will be interesting to study how the controversy around the COVID vaccine impacts flu vaccination rates as well. Some have speculated that the COVID vaccine will be rolled into the flu vaccine, or more likely it will be a separate annual booster that is tailored to the most prevalent strains that season. Even though vaccines have been widely used and tested for almost a century, people choose to abstain from getting vaccinated for various health and personal reasons. I predict that the hesitation for the COVID vaccine will also impact people's likelihood to get other vaccines, including the flu vaccine. For this reason, we may see higher rates of flu contraction in years to come.

Limitations

The main challenge I had when analyzing the flu data is regarding granularity of data. The flu data is available only on a state level, but the COVID case data and mask data is broken down by county. For COVID cases, I can aggregate by state instead of county, but the mask mandates differ by county. To address this issue and for consistency, I used county-level data to answer research question 1 and state-level data to answer question 2.

Another complication I encountered when working on research question 1 is figuring out how much of the population is susceptible at any given time. Factors that could impact who is susceptible include transmission rates of different variants, recovery times, likelihood of reinfection, breakthrough cases after vaccination, and a myriad of other unknown variables like self-isolation or non-compliance with masking that we have no way to measure. I chose to use a relatively simple formula to find the susceptible population since so many of these other factors are impossible to know (and likely balance each other out when taken altogether). My estimate for susceptible population is likely an underestimate of the actual number of susceptible people because once a person has recovered from the virus they return to the susceptible population; however for this purpose we will assume it is a close enough approximation due to the number of individuals who are also in the susceptible population but had very limited exposure due to self-imposed quarantine.

We also had meaningful discussions about the significance of the daily case count vs daily case rate, and what we learn when the rate of cases exceeds the 95% confidence interval. For example, we may expect the daily case count to increase, but an increase in the daily case rate is concerning because it indicates that infected individuals are infecting others more quickly Ultimately we determined that there are so many unknown factors and confounding variables (e.g. holiday season or new variants) that make it almost impossible to determine precisely how effective masks were for a given county.

Conclusion

This analysis aimed to answer two questions: (1) How did masking policies change the progression of confirmed COVID-19? And (2) How restrictions put in place for COVID change the transmission of seasonal influenza during the 2020-2021 flu season? The hypothesis for question 1 is that masks helped mitigate the spread of COVID-19, but probably not substantially due to the number of other variables at play. For question 2, the hypothesis is that COVID restrictions would drastically decrease the spread of seasonal flu. I found both of these hypotheses to be true: masks seemed to have some effect on controlling the spread of COVID, but not enough to have prevented a massive spike in cases over the 2020 holiday season. However, since the seasonal flu is less contagious than COVID, the restrictions that were in place practically eliminated the flu during the ‘20-’21 flu season.

References

Key Facts About Influenza (Flu) | CDC. (2021). Retrieved 10 December 2021, from https://www.cdc.gov/flu/about/keyfacts.htm

Burden of Influenza. (2021). Retrieved 14 December 2021, from https://www.cdc.gov/flu/about/burden/index.html

Data Sources

https://gis.cdc.gov/grasp/fluview/fluportaldashboard.html

https://www.kaggle.com/antgoldbloom/covid19-data-from-john-hopkins-university?select=RAW_global_deaths.csv

https://data.cdc.gov/Policy-Surveillance/U-S-State-and-Territorial-Public-Mask-Mandates-Fro/62d6-pm5i