Nicole Riggio
12/14/21

A7 – Project Report

Introduction

As of the 2020 US Census, Marion County, IN had almost 1M residents – representing nearly 1M residents and containing just under 15% of the state's total population [1]. Marion County also contains Indianapolis, the state capitol and largest metropolitan area in the state. As a result, it is not surprising that Marion County recorded the first COVID-19 case in the state of Indiana on March 6, 2020. This analysis seeks to gain insight into how a real or perceived threat from COVID-19 in a community can impact outcomes.

We will first explore the relationship between any public masking mandates and confirmed COVID-19 cases before investigating how outcomes – specifically increased hospitalizations from COVID-19 may have impacted deaths in the area and vaccination rate among the population. Answering these questions can help influence public messaging regarding safety measures (i.e. public masking or vaccination) as well as inform the COVID-19 response in hospitals such as anticipating staffing shortages or determining when to cancel elective procedures. Successful adoption of COVID-19 mitigation strategies or countermeasures is shown to have an impact on the spread of COVID-19 in a community and gaining a better understanding of how this has worked in practice in Marion County, IN can be useful for local public health officials and community members alike.

Background & Related Work

In June 2020, the University of Iowa published a study in the Health Affairs journal which analyzed the relationship between COVID-19 case rates and mask mandates in 15 states [2]. This study focused on community or public-use of face masks between April 8, 2020 and May 15, 2020, and found that they potentially averted over 200k COVID-19 cases in the study population. The authors compared this model to one where masks were only required for employees during working time and found no significant decline in COVID-19 growth rates. As a result, this study adds to the evidence that public masking mandates had a significant impact on reducing COVID-19 spread where required. We are interested to see if this trend holds true for Marion County – the county did not have a public mask mandate until July 27, 2020 and did not see a major COVID-19 spike until mid-to-late November 2020. It is possible that residents did not perceive a real threat from COVID-19 since their first major wave did not occur until well after areas like New York City became overwhelmed by the pandemic.

Other related research is on the effectiveness of vaccines – in a medRXiv preprint, a team of researchers found in a survey of healthcare workers and long-term care facility residents that COVID-19 spread declined in the weeks following their first vaccine while spread continued to grow for their unvaccinated peers [3]. The study noted that vaccinated individuals had

significantly lower 14-day hospitalization rates when propensity matched to unvaccinated individuals. This could indicate that a populations willingness to receive the COVID-19 vaccine could have an impact on infection rate across their community – highlighting the importance of public health messaging regarding COVID-19 vaccination and the ability of vaccines to prevent hospitalization from COVID-19.

The CDC published a report on the Impact of Hospital Strain on Excess Deaths During the COVID-19 Pandemic – specifically focusing on the United States from July 2020-21 [4]. This report asserts that intensive care unit (ICU) bed use at 75% occupancy is associated with an estimated additional 12k excess deaths 2 weeks later while an occupancy in excess of 100% could lead to 80k excess deaths across the country. The Indiana Department of Homeland Security stated in a memo that when a hospital is on diversion, it means that the "current emergency patient load exceeds the Emergency Department's ability to treat additional patients promptly" [5]. There are anecdotal reports from local news that Marion County, IN hospitals were operating at capacity and at one point 11 of 13 area hospitals were on diversion [6][7]. The importance of maintaining acceptable ICU occupancy levels can have an impact on health outcomes for anyone needing ICU care regardless of their COVID-19 infection or vaccination status.

It is also important to note that ICU occupancy is more often limited by maintaining safe staffing levels rather than actual bed availability. A Washington Post article from August 2021 highlights this phenomenon for six Broward County hospitals in Florida [8]. Better understanding the lag between hospitalizations and deaths from COVID-19 can help to anticipate staffing needs. In this case, hospitals had enough beds but not nearly enough nurses to staff them and was scrambling to hire travel nurses as they dealt with the highest rate of COVID-19 admissions in the country.

Additional research on the perception of COVID-19 in Marion County, IN revealed that the state government had blocked vaccine mandates and refused to issue a second mask mandate in the wake of rising COVID-19 infections in mid-to-late November 2021 that forced the cancellation of elective procedures in Marion County hospitals [9][10]. Other research on location conditions confirmed that all Indiana residents age 16+ became eligible for vaccination against COVID-19 on March 31, 2021 [11].

Methodology

Data for this analysis was gathered from a variety of sources including a Kaggle repository of US confirmed COVID-19 cases & deaths from John Hopkins University, a CDC dataset of masking mandates, and a NYTimes survey on mask compliance. Additional data was gathered from the Marion County COVID Public Information Dashboard and the Indiana Department of Health (IDOH) related to COVID-19 hospitalizations, ICU occupancy, and vaccination status. All data is

updated daily by county – aside from the IDOH hospital occupancy rates which is aggregated at the state-level and the Marion County hospitalization count which is computed as a 7-day rolling average. Data sources are cited below and all data is available via an Attribution 4.0 International or Creative Commons license which allow for re-distribution and re-use. The Marion County dashboard is produced by the Marion County Public Health Department (MCPHD) and is only available as a visualization. The WebPlotDigitizer tool was used to extract the 7-day rolling hospitalization average due to COVID-19 in Marion County and validated by spot checking extracted values with the original dashboard [12].

To investigate the impact of a public masking mandate on COVID-19 in Marion County, IN – visualizations were produced showing the timeline of the mask mandate as compared to increases in COVID-19 confirmed cases and infection rate among the population. This was compared to changes in daily infection rate to determine if any notable changes took place while the mandate was in effect. Two major waves of COVID-19 infection in the community were identified for further investigation.

Continued analysis involves the removal of stationarity through smoothing of the time-series and attempts to calculate Case Fatality Ratio (CFR) over time. Case Fatality Ratio represents the proportion of cases that eventually result in death from a disease "once an epidemic has ended" – this is not information we presently have available and is difficult to track over time with the presence of only aggregate data. An article in Acta Biomedica proposes calculated CFR over time by dividing the deaths on a given day with the cases on that day minus some constant lag [13]. This methodology was investigated to determine if a robust, representative value for this constant could be found. This analysis would given local officials details on how patient outcomes were impacted by increases in hospitalization rate.

Since another hypothesis was that the relationship would be stronger between COVID-19 hospitalizations & deaths than cases & deaths – Pearson correlation was used to determine which two signals had stronger global synchrony [14]. An Augmented Dickey-Fuller (ADF) test will be used to validate assumptions on stationarity and correct for any seasonality or trend before the correlation is performed [15, 16]. This analysis will also employ Time Lagged Cross-Correlation (TLCC) to identify the anticipated lag between hospitalizations & deaths. Unlike the more straightforward Pearson R correlation, TLCC provides more insight into signal dynamics and can tell us which signal leads the other and by how much [14].

This analysis will look at how spikes in COVID-19 impacted ICU occupancy in Marion County and it this trend held over the course of the pandemic. To investigate trend – a linear fit on overall ICU occupancy will be calculated. Finally – we will attempt to quantify any change in behavior through an analysis of COVID-19 vaccination rates over time. We expect to see an initial rush on vaccinations once they become available but are specifically interested in how vaccination rate changes as the initial frenzy subsided.

Overall, concern should be taken to ensure that the assumptions made in this analysis are well documented to better understand the limitations of any conclusions drawn. It is important to be cognizant of potential confounding effects that may change how any relationships may change throughout the course of the pandemic. For example, it is possible that treatment protocols or COVID-19 variants with higher infections ability may change timelines or skew results. We should note that we have an ethical obligation to ensure overly simplistic conclusions such as “masks don’t work” or “most people don’t die from COVID” should be stated cautiously with appropriate caveats as to not be misinterpreted.

Findings

Preliminary analysis found that while Marion County had a public mask mandate from July 27, 2020 until April 5, 2021 the first major COVID-19 wave in the did not peak until December 2021 [Figure 1].

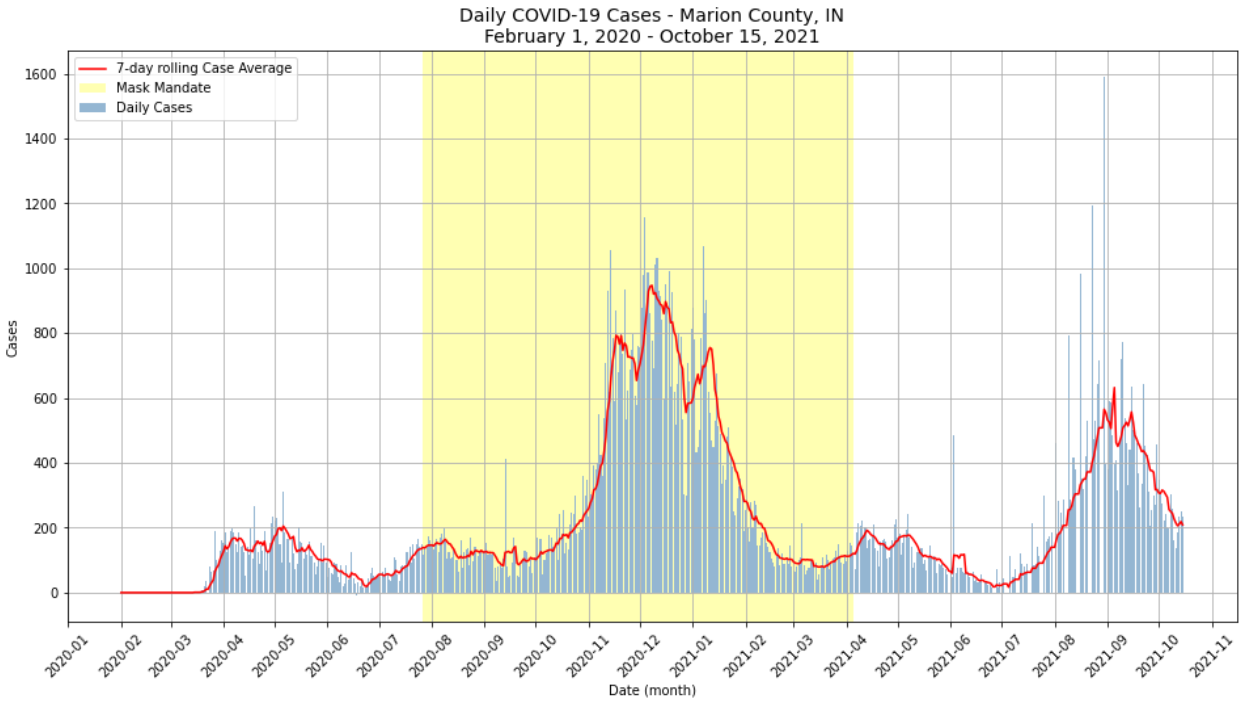

Figure 1

This analysis also found that while the proposed formula for calculating Case Fatality Ratio as described above was an overly simplistic solution. There is a non-constant lag between cases and deaths – with early reporting showing a higher proportion of deaths early on followed by additional fluctuations throughout the pandemic. This is potential indicative of early testing shortages. As seen in Figure 2, there is a point beginning in November 2020 where these lines do become parallel with some lag – however, they begin to separate even before vaccines are available.

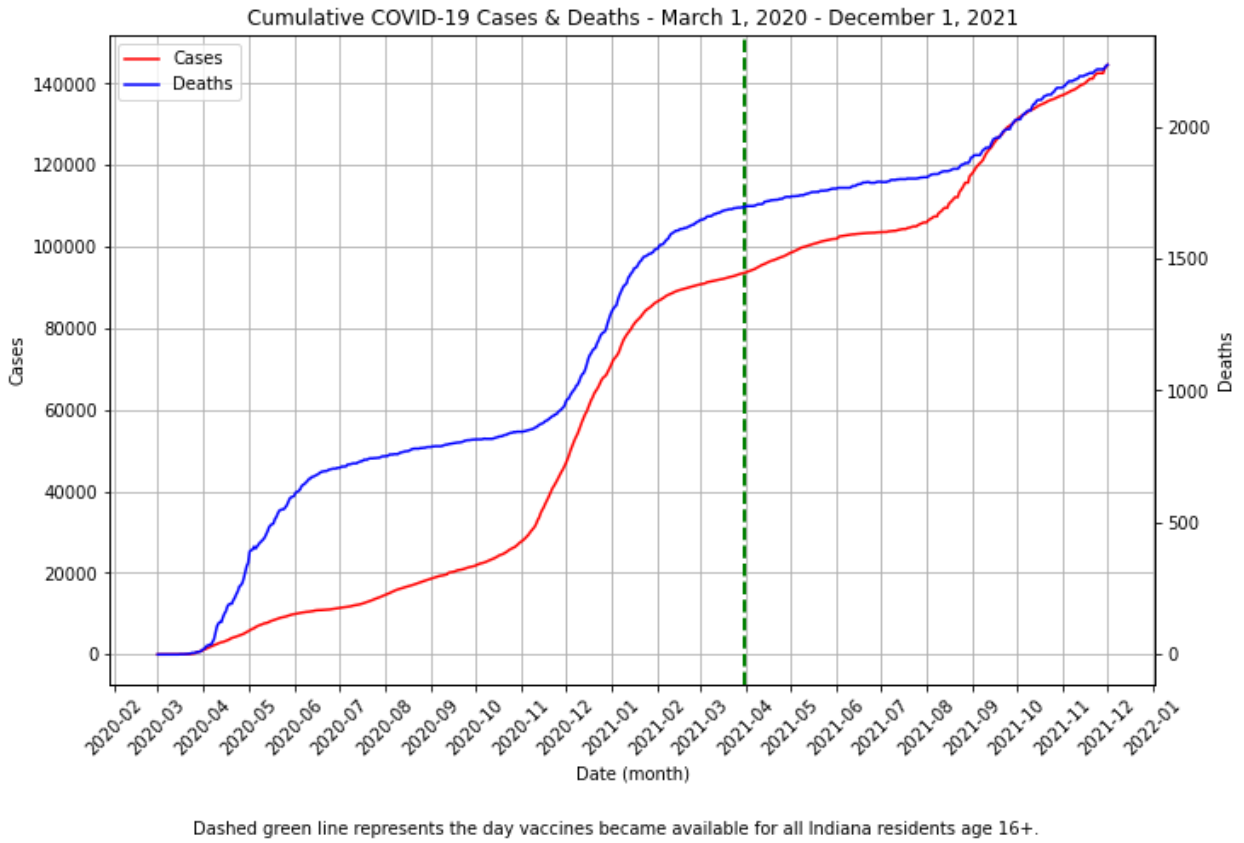

Figure 2

Despite CFR over time being difficult to accurately measure given available data, we are still able to review whether cases or hospitalizations were a more robust metric for predicting COVID-19 deaths. Due to expected reporting inconsistencies on related to day-of-the-week, smoothing the time series through a rolling 7-day average made the data stationary and allowed for the calculation of Pearson R. As expected, this result confirmed that there was greater global synchrony between hospitalizations and deaths [Figure 3].

	Daily Count p-value	Rolling Avg. p-value
Cases	0.231	0.097
Hospitalizations	N/A	0.037
Deaths	0.184	0.030

	Pearson R	p-value
Cases-Deaths	0.533	<0.001
Hospitalizations-Deaths	0.631	<0.001

Figure 3

By using Time-Lagged Cross Correlation, the offset between hospitalizations and deaths that maximized Pearson R was found to be hospitalizations leading by 16 days [Figure 4]. This maximized the correlation at 0.852 which is quite strong and can be used to help inform disease progression and projected hospital occupancy.

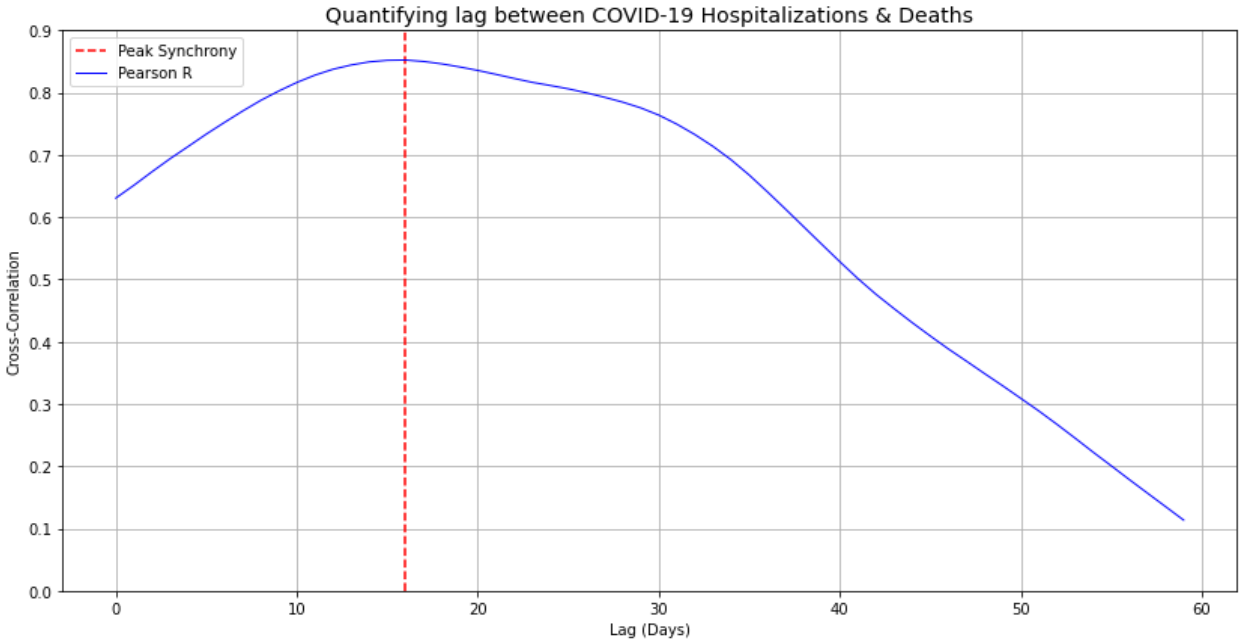

Figure 4

A linear fit of state-level ICU occupancy over time showed an increasing trend in overall occupancy rates throughout the pandemic [Figure 5]. Although local hospitals deployed mitigation strategies such as cancelling elective procedures, they were unable to offset the influx of COVID-19 patients. Occupancy rates can be misleading as less-critical patients may be admitted if there is space to accommodate them. Nevertheless, Indiana saw increased occupancy over the course of the pandemic with the second wave having higher total capacity and during times of peak COVID-19 surpassed the 75% threshold that has been shown to correlate with excess deaths.

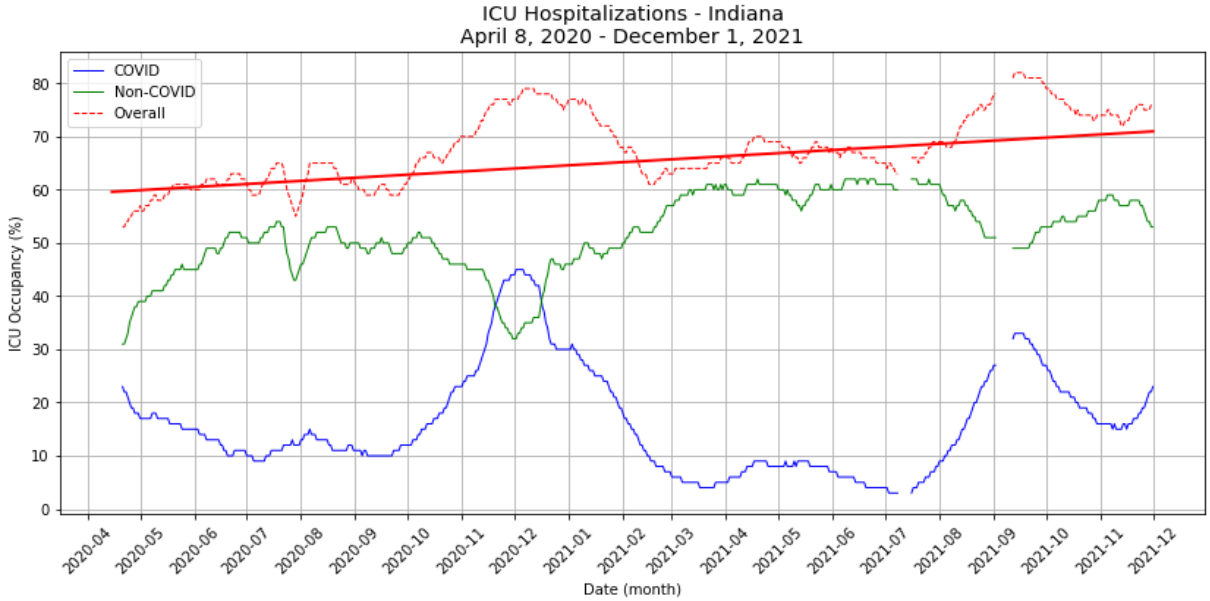

Figure 5

There is also some evidence that a surge of COVID-19 cases that peaked in late August 2021 caused a slight uptick in vaccination rate among Marion County residents [Figure 6]. It is difficult to tell if this trend holds without further data, but it could be a sign that local news of increased cases and hospital occupancy rates rising drove individuals to get vaccinated.

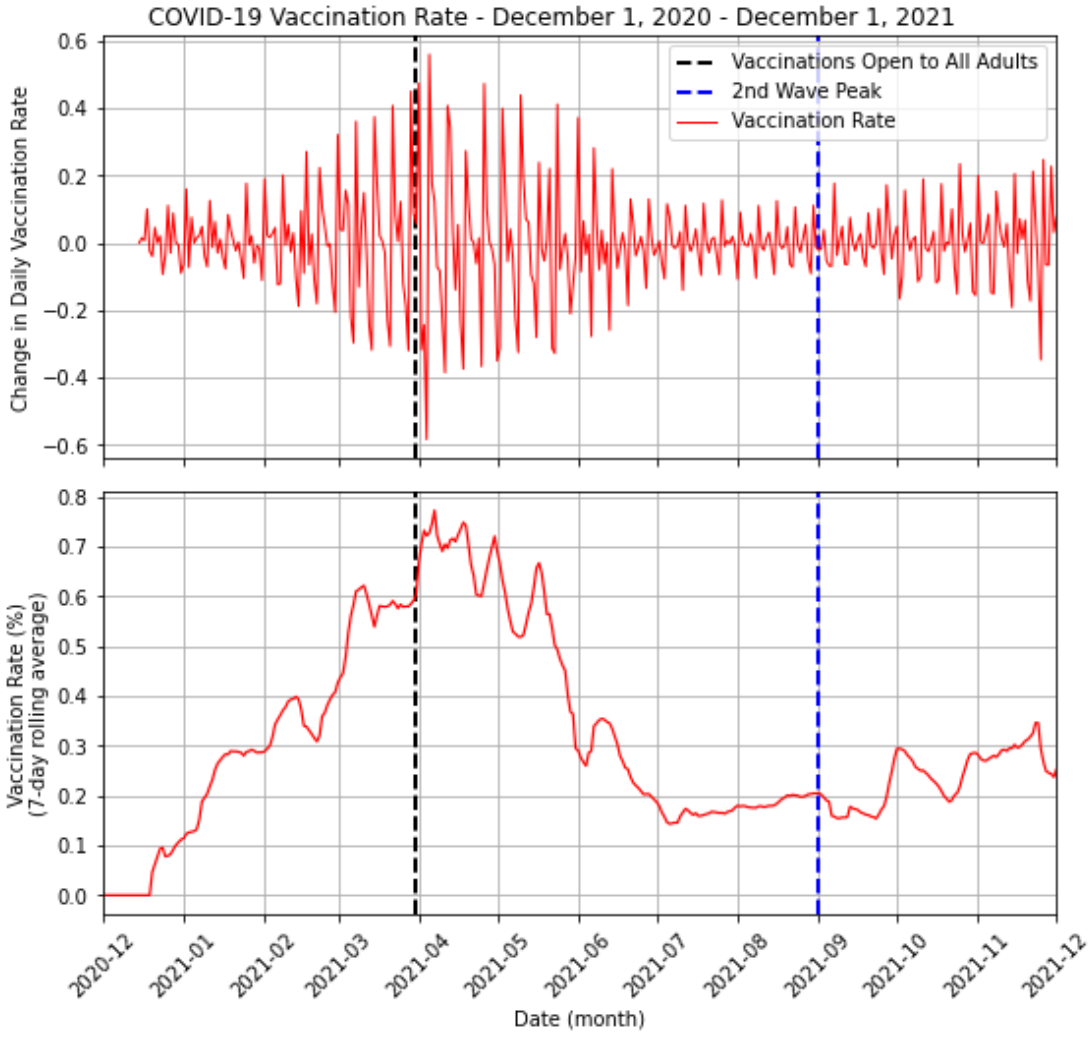

Figure 6

Discussion & Implications

Future research could expand upon this analysis to better inform hospital staffing decisions and public health messaging related to mitigation strategies such as public masking and vaccination. This analysis did not account for rates of mask compliance or social distancing but with additional data could be expanded upon to create a holistic picture of how Marion County and similar communities can reduce the spread of COVID-19.

It was important to highlight that early testing shortages or undetected asymptomatic transmission clouded analysis for determining change in CFR over time. It might be useful for a hospital or public health office that has access to individual data to perform this analysis by tracking outcomes for some random subset of individuals with confirmed COVID-19 cases.

I found it interesting that there was some slight uptick in vaccinations as the second COVID-19 wave was occurring in Marion County. The increase in vaccination rate included all vaccines (i.e. 1st or 2nd dose recipients) since there was some anecdotal evidence building that "the one -dose problem is real" indicating that some Americans skipped their second vaccination [17]. It could be important to investigate if this increase in vaccinations was due to individuals who got a first dose earlier on finally getting a second dose months later or if these were first time vaccine recipients. Although the governor of Indiana has blocked vaccine mandates where possible, it would be important to investigate if vaccine mandates for Federal employees had any impact on vaccine rates in the state. Having a better understanding of what leads people to get vaccinated can help local health and government officials promote better messaging around the important of taking precautions were possible. For example, if it proves significant that Marion County residents were more likely to get vaccinated after local news reported on hospital diversions and spiking cases, then accessibility of public health dashboard explaining the threat of COVID-19 to the community could prove valuable.

It is also possible to extend this analysis to additional counties or states – do we get the same or similar results for neighboring counties? Is there an urban/rural divide that impacts outcomes?

Limitations

As discussed previously, there were assumptions made in this analysis and some limitations to be aware of when drawing any conclusions. Most data used was fairly complete – however some tables, specifically related to hospital occupancy rates or vaccination were zero or null when we would have expected some value. In cases like this, a forward-fill of the time series was deployed when computing a rolling average or finding the correlation. This likely did not have much impact on the results. For vaccination data, the data dictionary itself makes clear that any values less than 5 were suppressed for privacy concerns.

All data used was licensed for re-use with proper attribution – the main concern with data collection was the use of WebPlotDigitizer to extract county hospitalization data from a dashboard, this data was spot-checked to verify accuracy [12]. This data was licensed for public use and there is some indication that the Marion County Public Health Department could be contacted to share the underlying data but given the timeline of this project extracting it directly from the dashboard proved easier.

Additional assumptions were made when verifying the hypothesis that hospitalizations would prove to be a more robust predictor of COVID-19 deaths than confirmed cases – because the Augmented Dickey-Fuller test and Pearson R correlation assume no null values, any missing case, death, or hospitalization data was imputed. There were relatively few missing values in these datasets, and similar results were achieved by imputing values with either a 0 or forward filling the last non-null value available. There is also potential for bias in peak-picking where the spikes COVID-19 cases occurred – this was generally done by identifying the local maxima on the smoothed signal. However, it is possible that smoothing the data in this way shifted the true peak.

Other limitations include assumptions around calculating rates – for example, infection rate or case fatality rate – in some cases the calculation and assumptions felt robust enough and in-line with known standards that further analysis was appropriate, and in other cases another approach was chosen to tackle the hypothesis. To calculate infection rate it was assumed that all cases were properly identified by a confirmed positive COVID-19 test result (i.e. no false positives or false negatives and all individuals who were infected got tested) and that all case cycles were equal in duration. Individuals were assumed to have "active cases" for 3 days prior to and 10 days following a confirmed test – this aligns with CDC guidance to get tested 3-5 days post exposure and remain in quarantine for 10 days post diagnosis.

When calculating Case Fatality Rate over time – the inability to track individuals was a major limitation – the assumption that all cases were equal in duration and resulted in death or recovery on the same day post confirmed case status was too general to be useful. As shown in Figure 2, even with aggregate data it is clear these assumptions do not hold up and this is a major opportunity for further exploration. These were the main limitations when conducting this exercise – any additional assumptions made are outlined in analysis itself.

Conclusion

This analysis can better inform messaging from public health officials on COVID-19 mitigation strategies and provide hospitals with additional information to on ICU occupancy and staffing needs as the pandemic continues. It also identifies areas for further exploration to better achieve the stated goals – many of which are achievable for local or state officials with access to non-aggregate data.

A 16-day lag was identified as the peak synchrony between COVID-19 hospitalizations and deaths – this can aid officials in projecting ICU occupancy and avoid the 75% occupancy threshold that has been shown to correlation with excess deaths for all patients. There is some evidence for an uptick in vaccinations following the second wave of COVID-19 in Marion County that may be independent of vaccine mandates or other requirements. This could indicate that

local reporting on hospital conditions and rising case levels could have an impact on community behavior.

The human-centeredness of this analysis lies in the attempt to measure human behavior in response to a quantifiable rise in cases and the anecdotal nature of local news reporting. The intended outcomes are to help decision-makers better predict and prepare for future conditions while also being aware of the consequences for anyone needs hospital care if COVID-19 is able to overwhelm the system.

References

- [1] US Census: Indiana
- [2] Community Use of Face Masks and COVID-19
- [3] FDA-authorized COVID-19 vaccines are effective per real-world evidence
- [4] Impact of Hospital Strain on Excess Deaths
- [5] IN DHS: EMS Ambulance Diversions
- [6] US News: Indiana Hospitals at Capacity
- [7] WTHR: Marion County Hospitals on Diversion
- [8] FL hospital shortages related to staffing concerns
- [9] US News: Indiana Governor Stymies Vaccine Mandates
- [10] IN Governor rejects second mask mandate
- [11] IndyStar: Indiana Vaccine Eligibility
- [12] WebPlotDigitizer
- [13] Challenges in COVID-19 Case Fatality Ratio Calculation
- [14] Pearson R & Time-Lagged Cross Correlation
- [15] Stationarity in Time Series
- [16] Augmented Dickey-Fuller Test
- [17] Vox: The One-Dose Problem

Data Sources

- CDC: Mask Mandates by County
- IDOH: COVID-19 Bed & Vent Usage by Day
- IDOH: COVID-19 Vaccinations by County by Date
- Kaggle: RAW_us_deaths.csv via JHU
- Kaggle: RAW_us_confirmed_cases.csv via JHU
- Marion County COVID-19 Dashboard

- NYTimes: Mask Compliance
- Statesman Journal: Hospital Capacity by County