

# Correlation of Vaccine Hesitancy with COVID-19 Deaths and Cases

Tzu Yu Huang

2022-12-01

## Introduction

Vaccine hesitancy, which is the reluctance or refusal to vaccinate despite the availability of vaccines, has been proposed by the World Health Organization (WHO) as one of the ten threats to global health in 2019 <https://www.who.int/news-room/spotlight/ten-threats-to-global-health-in-2019>, even before the global pandemic of COVID-19.

Vaccine is currently one of the most cost-effective ways of preventing disease outbreak and reducing disease consequence, and consequently disease's impact on economic. As of 2019, 2-3 million deaths a year are prevented by vaccination. If global coverage of vaccinations is improved, an estimated 1.5 million of death could be further avoided.

As for the reason of vaccines hesitancy are complex, could be a result of numerous factors including lack of education, knowledge of how vaccines are produced, distrust of government or authorities, or even just fear of needles. Its existence dates back the time vaccines were invented, however, it is only recently that it started to resurface to the discussion of mainstream media.

After the global outbreak of COVID-19 and the invention of COVID-19 vaccines, vaccine hesitancy plays an even more important role than before. We can observe vaccines being not just a measure of disease prevention, but also a mean of politics, yet we can also see a subset of population refused to get vaccinated even with the resources at hand.

This report is not to discuss the cause of vaccine hesitancy or the legitimacy of the claims by certain "anti-vaccinationism", but to see if vaccine hesitancy actually impacts the death and case number of COVID-19 on a state level in the US.

# Methods

## Datasets

A total of 3 datasets will be used in this project. The first two are from US Center for Disease Control and Prevention, and the third one are from United States Census Bureau (an official website of the United State Government).

1. The first dataset is “Vaccine Hesitancy for COVID-19: County and local estimates” from <https://data.cdc.gov/Vaccinations/Vaccine-Hesitancy-for-COVID-19-County-and-local-es/q9mh-h2tw> Read in the data using through API web scrapping.
2. The second dataset is “United States COVID-19 Cases and Deaths by State over Time” from <https://data.cdc.gov/Case-Surveillance/United-States-COVID-19-Cases-and-Deaths-by-State-o/9mfq-cb36> Read in the data using through API web scrapping.
3. And for the comparison between states summary statistics to be reasonable and effective, I need the population data of different places. “County Population Totals: 2020-2021”. Downloaded from <https://www.census.gov/data/tables/time-series/demo/popest/2020s-counties-total.html>

No missing data in the key variables of interest are found in the datasets. After reviewing the data, there are a total number of 60 states in the COVID-19 death and cases dataset, after investigation, those state include US territory oversea such as Guam and American Samoa, for this investigation, I’ll just focus on non-overseas US states. A variety of plots are plotted to answer the reasearch question, such as scatter, line plot, and heat map.

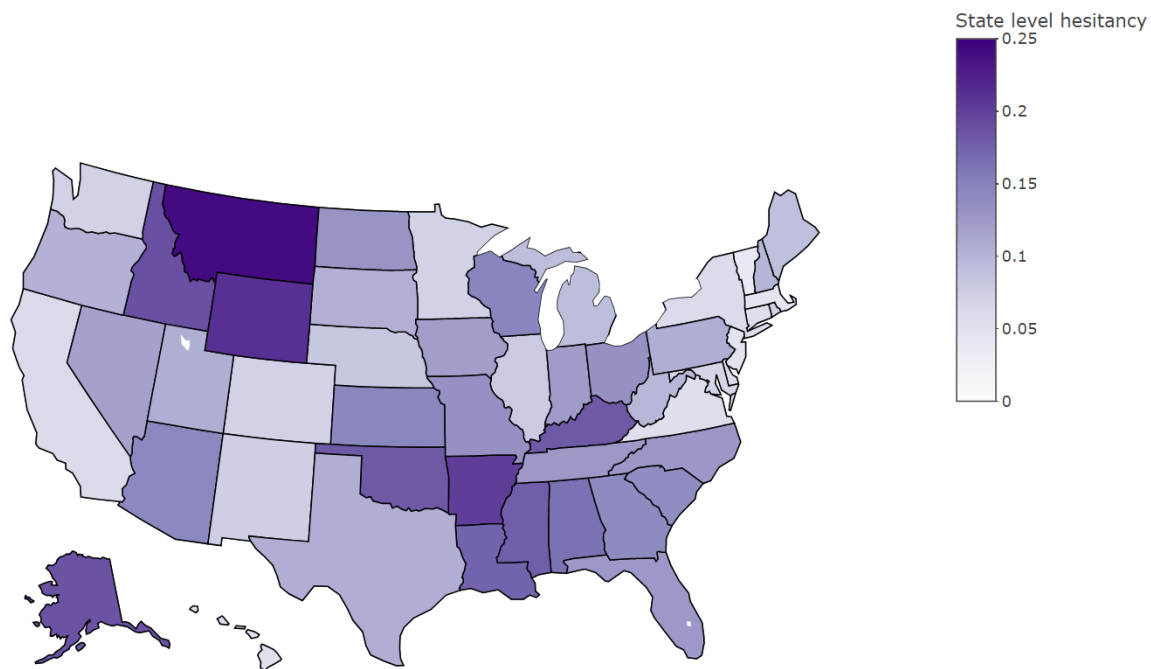
## Estimated state vaccine hesitancy level

The state vaccine hesitancy level is estimated by calculating the weighted average of the county hesitancy level, with population information from the third dataset as weight. While the state population us calculated by summing the counties population within each state.

## Results

First, we look at the distribution of estimated hesitancy level for each state. It seems that Montana, Wyoming, Alaska, Arkansas are the few states with the higher estimated state vaccine hesitancy level among all the states within the US (Figure 1). An interactive version of this plot can be found on my website. <https://rawcdn.githack.com/tzuyuhua/PM566-finalproject/bc7d92e2def05c61b190ebbe0656976028f4a998/index.html>

Figure 1. State level hesitancy



Next, we look at the Population-normalized COVID-19 deaths/cases over time (Figure 2). We are trying to see if states with higher estimated state vaccine hesitancy level would also show a stronger surge of COVID-19 cases and deaths per person (state level case number and death number divided by state population) over time. We can observe that for both Alaska and Arkansas, two state that were previously observed with higher estimated state vaccine hesitancy level among all the states within the US, shows a strong surge of cases and deaths compare to other state. However, this trend does not hold for Montana, while it only shows an average level of surge of COVID-19 cases and deaths compare to other states. An interactive version of this plot can be found on my website. <https://rawcdn.githack.com/tzuyuhua/PM566-finalproject/bc7d92e2def05c61b190ebbe0656976028f4a998/index.html>

Figure 2a. Population-normalized COVID-19 cases over time

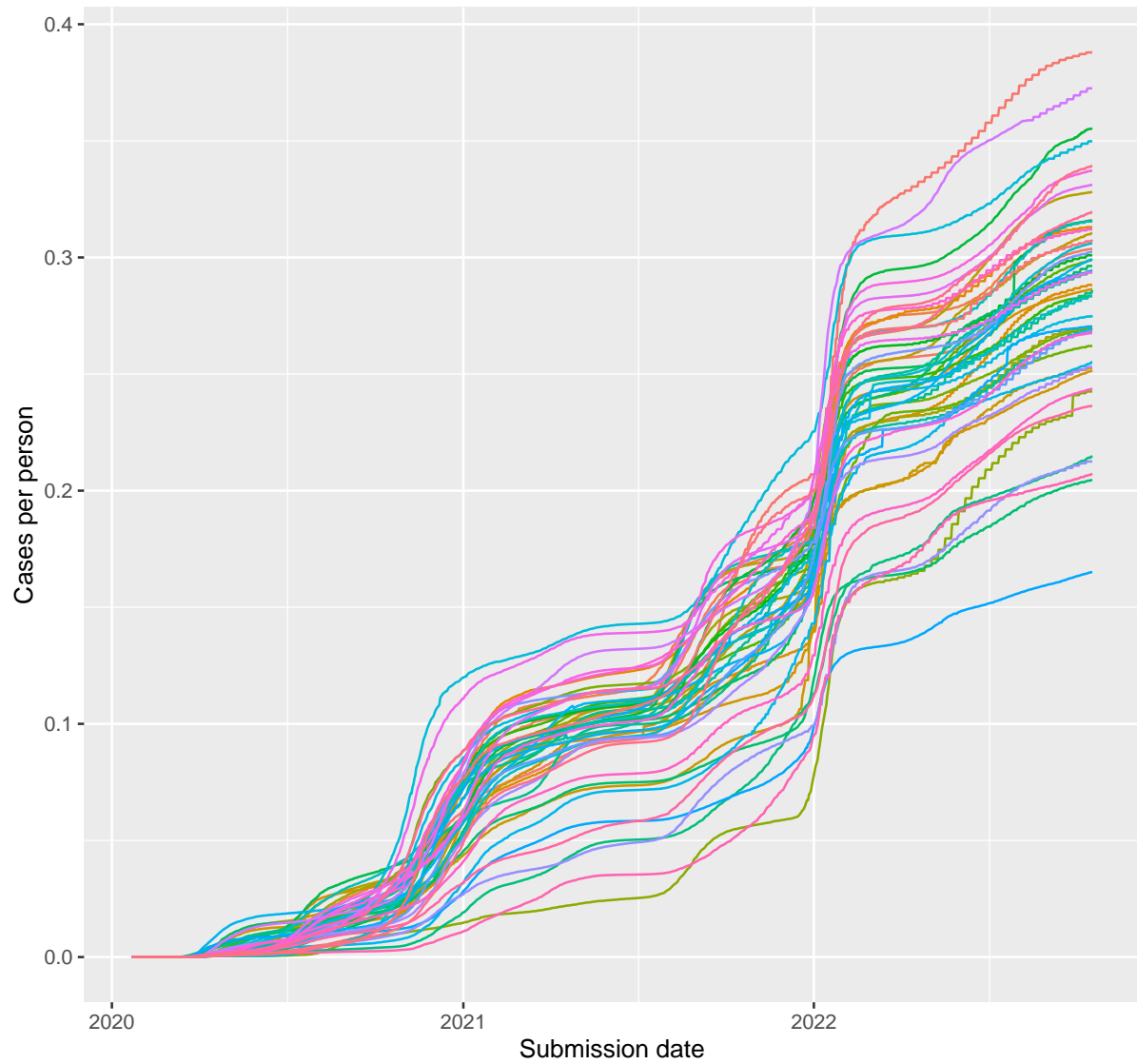
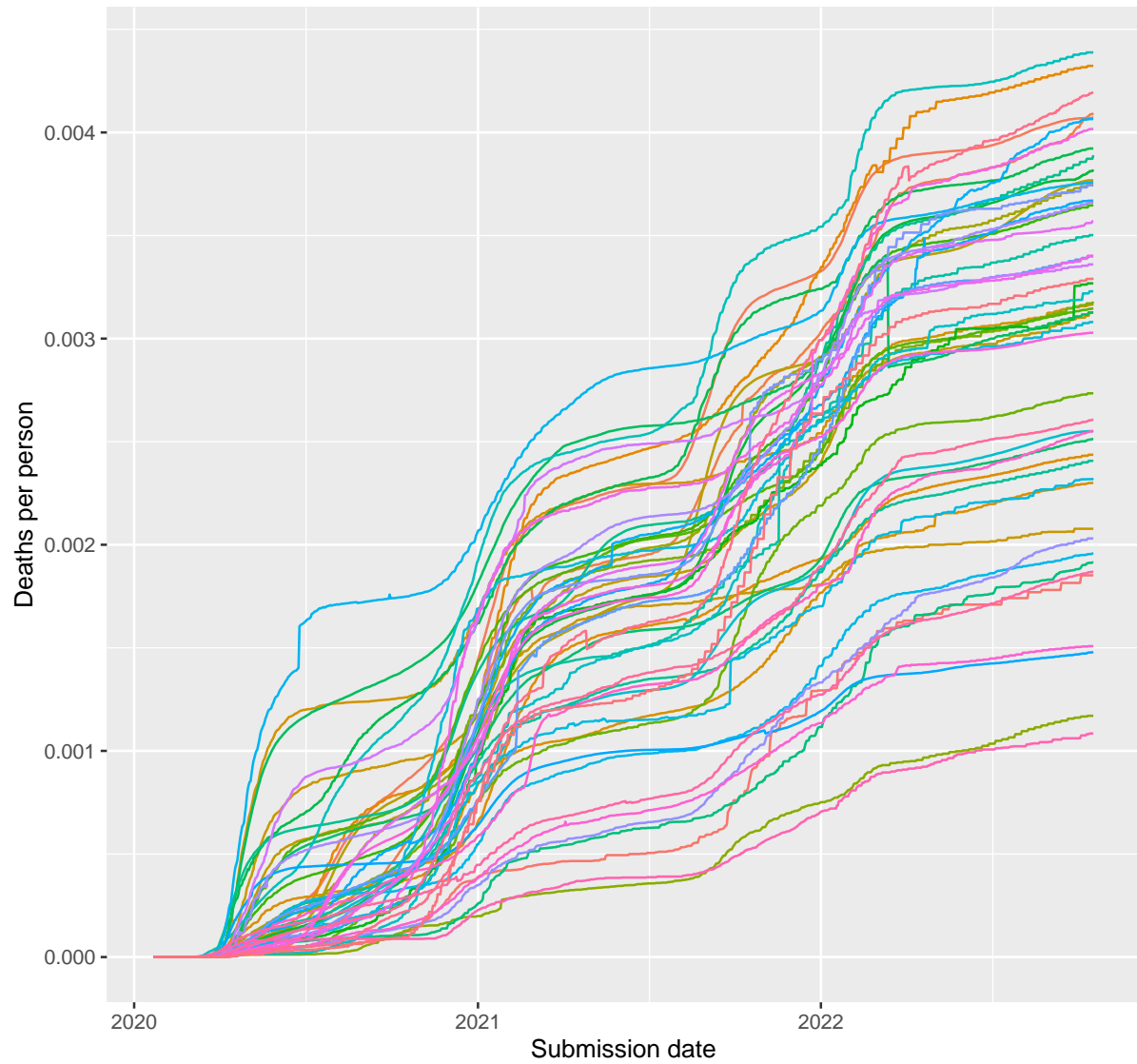


Figure 2b. Population-normalized COVID-19 deaths over time



AK	GA	ME	NM	SC
AL	GU	MI	NV	SD
AR	HI	MN	NY	TN
AS	IA	MO	NYC	TX
AZ	ID	MP	OH	UT
CA	IL	MS	OK	VA
CO	IN	MT	OR	VI
CT	KS	NC	PA	VT
DC	KY	ND	PR	WA
DE	LA	NE	PW	WI
FL	MA	NH	RI	WV
FSM	MD	NJ	RMI	WY

Third, we examine the relationship between total cases, total deaths, and with the diameter of the data point as the estimated state hesitancy level. We are expecting to see a positive correlation between total cases and total deaths (controlled for state population), which is shown in Figure 4. We are also trying to see if data points closer to the upper right hand corner (more deaths and cases per person), would also be bigger in diameter, showing correlation with vaccine hesitancy level. And that is generally the case in Figure 4, although Alaska is interesting, with high cases per person and high estimated hesitancy level, but relatively low deaths per person. An interactive version of this plot can be found on my website. <https://rawcdn.github.com/tzuyuhua/PM566-finalproject/bc7d92e2def05c61b190ebbe0656976028f4a998/index.html>

Figure 3.

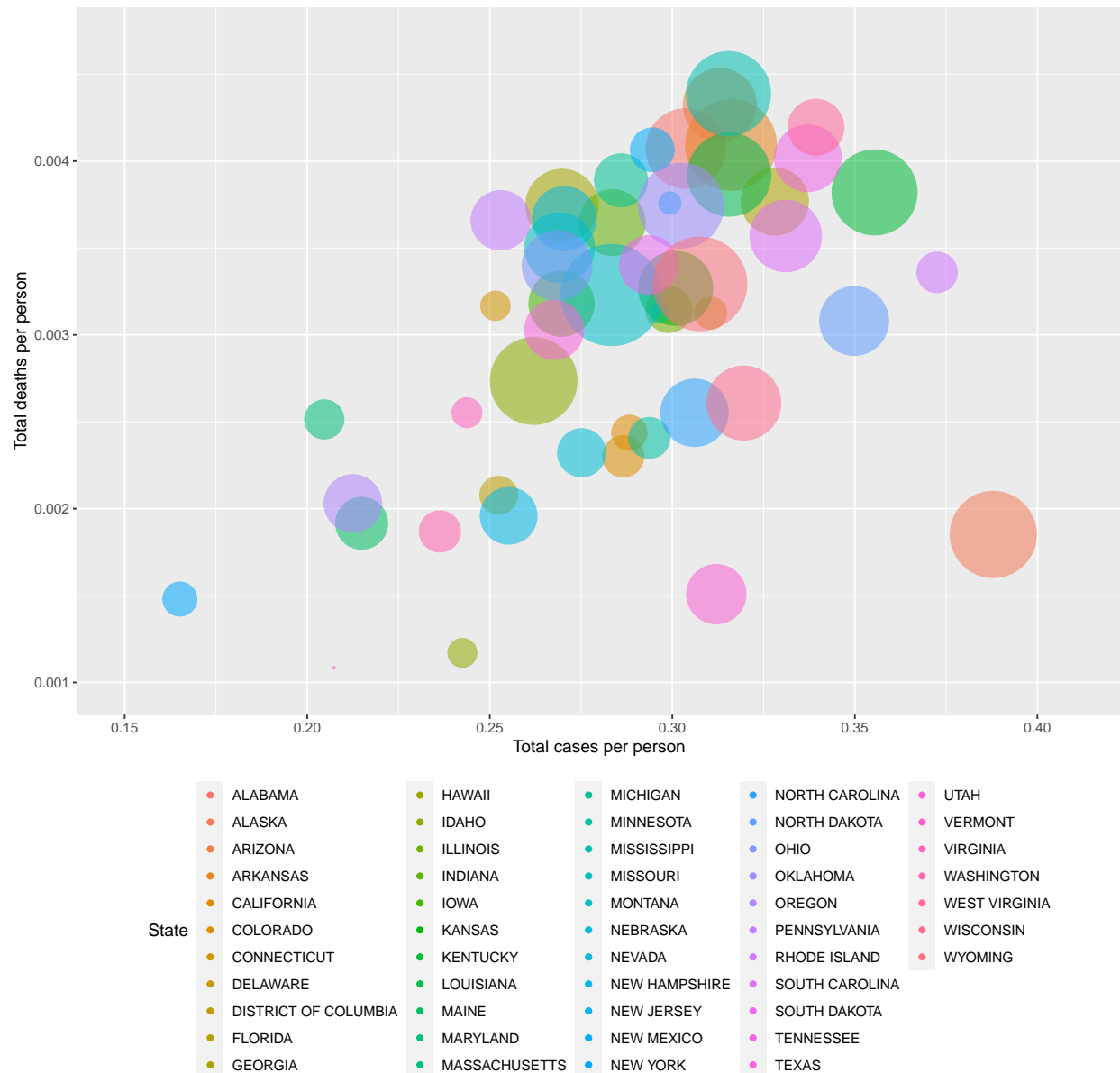


Figure 4a.

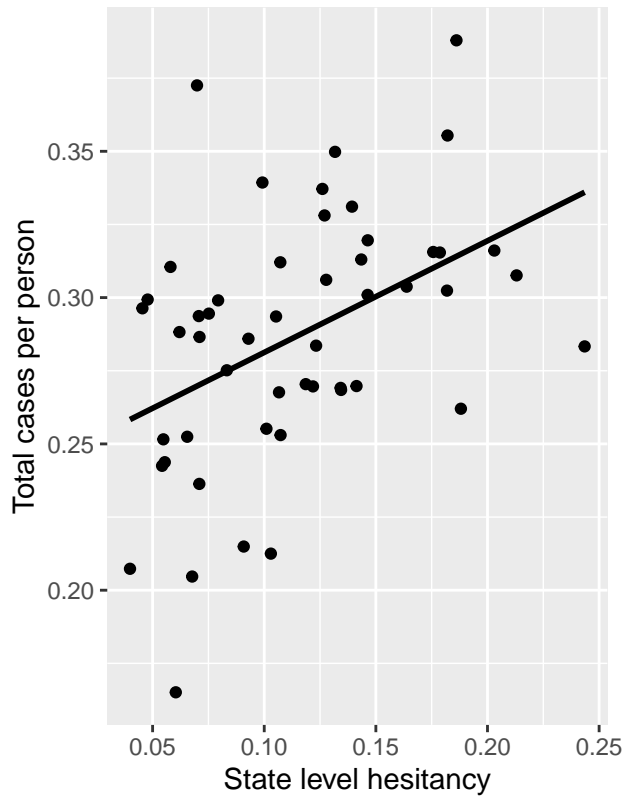
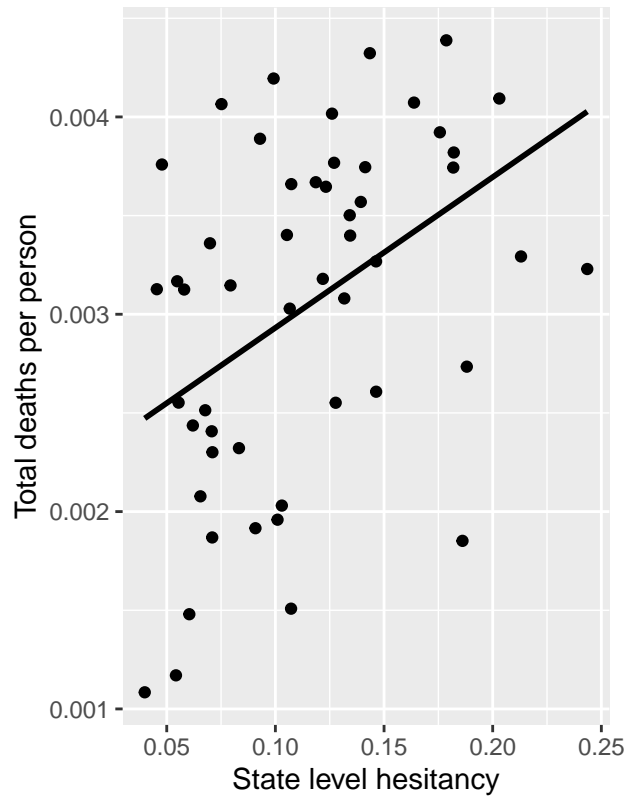


Figure 4b.



Now that's take a look at the relationship between hesitancy and It does appear that there is some positive correlation between vaccine hesitancy and total number of cases (proportion to state population size) (Figure 4a). From the above graph, there is also a positive correlation between vaccine hesitancy and total number of deaths (proportion to state population size). And the slope appears to be even steeper.(Figure 4b)

## Conclusion

Overall, vaccine hesitancy does show to have a positive correlation with death and cases, and death especially, which is in accordance with the general science community consensus, that vaccine can slow down the rate of COVID-19 transmission, and reduce COVID-19 symptom severity (increase protection against hospitalization and death in people). However, more confounding variables (etc, age, masks or population density) should be included to better understand the relationship.