Data512  
Qiaohui Sun

**Will Vaccination Help Reduce Spread of Covid-19?**

**El Paso County**

**Introduction**

The spread of SARS-CoV-2, the causative agent of COVID-19, has resulted in an unprecedented global public health and economic crisis. The outbreak was declared a pandemic by the World Health Organization on 11 March 2020, and development of COVID-19 vaccines has been a major undertaking in fighting the disease. It is estimated that a novel COVID-19 vaccine will need to be accepted by at least 55% of the population to provide herd immunity, with estimates reaching as high as 85% depending on country and infection rate. Reaching these required vaccination levels should not be assumed given well-documented evidence of vaccination hesitancy across the world, which is often fueled by online and offline misinformation surrounding the importance, safety, or effectiveness of vaccines. There has been widely circulating false information about the pandemic on social media platforms, such as 5G mobile networks are linked with the virus, vaccine trial participants have died after taking a COVID-19 vaccine, and the pandemic is a conspiracy or a bioweapon. Such information can build on pre-existing fears, seeding doubt and cynicism over new vaccines, and threatens to limit public uptake of COVID-19 vaccines.   
  
Therefore, in this research, we would like to further investigate the impact of mask mandate and Covid 19 vaccine administration on the spread of the disease in El Paso County in Colorado State. We also strive towards increasing the awareness of the importance of Covid-19 vaccination and help change the perspective of unvaccinated people towards the need for vaccination.

**Background/Related Work**

It has long been believed that the COVID-19 vaccine can reduce the risk of serious illness and hospitalization. But until recently, there were questions about whether they could reduce the spread of the virus.

This is an important issue and a public health policy formulated around this issue. For example, the UK has mandated that all social care workers be vaccinated against COVID to protect the vulnerable groups they care for; they will do the same for NHS staff. In Italy, from October 15th, workers must show their employer a vaccination certificate, test negative for COVID or recover from a recent infection. Anyone who fails to do this can be suspended without pay. Other countries are also taking similar measures.

A large study led by a team at Oxford University, especially investigating delta variants, has shown that both the Pfizer and AstraZeneca vaccines reduce disease transmission. This study examined the contacts of approximately 150,000 people followed from nearly 100,000 early cases of COVID. The first COVID-positive cases are a mix of vaccinated and unvaccinated, and the goal is not only which group is most likely to be infected with the virus, but also Pfizer. It was to see if the vaccine or the AstraZeneca vaccine was the most effective in reducing the infection.

The latest study examined the effect of vaccines on transmission more directly. It analyzed testing data from 139,164 close contacts of 95,716 people infected with SARS-CoV-2 between January and August 2021 in the United Kingdom, when the Alpha and Delta variants were competing for dominance. The authors found that the vaccine provided some protection against infection and subsequent infections, but Delta diminished its effectiveness. After being fully vaccinated, those who had “breakthrough” delta infections were almost twice as likely to be infected with the virus as those who were infected with alpha. And that was in addition to the higher risk of breakthrough infections caused by Delta than those caused by Alpha.

Although most clinical trials of the COVID-19 vaccine show that the vaccine can prevent the disease, some test results also provide clues that the vaccine may prevent infection. Larry Corey, a vaccinologist at the Fred Hutchinson Cancer Research Center in Seattle, Washington, said a vaccine that is highly effective in preventing people from getting infected from the start will help reduce transmission.

In this research, we would like to see if vaccination will help reduce the spread of the disease by using the data from El Paso County in Colorado States. The questions we were trying to answer are:

1. How did mask mandate policy help to reduce the spread of the disease?
2. How did Covid-19 vaccine administration help reduce the spread of the disease?
3. Did vaccination help with the decrease in daily Covid death?

**Methodology**

In this analysis, we will first use visualization to see if there is any potential association between mask mandate policy and the daily confirmed cases. Then we will use hypothesis testing with ordinary least-squares method to investigate the impact of Covid-19 vaccination administration on the spread of the disease.

The null hypothesis for the second and third research are as follows:

1) There is no association between the number of total vaccinated people and total confirmed cases in El Paso County.  
2) There is no association between the number of total vaccinated people and the number of daily newly confirmed cases in El Paso County.  
3) There is no association between the number of total vaccinated people and the increase in total death (caused by Covid19) in El Paso County.

Ordinary least-squares (OLS) model assumes that the analysis is fitting a model of a relationship between one or more explanatory variables. It also assumes that a continuous or at least interval outcome variable exists that minimizes the sum of squared errors, where an error is the difference between the actual and the predicted value of the outcome variable. The most common analytical method that utilizes OLS models is linear regression (with a single or multiple predictor variables).   
  
In our analysis, we will be conducting two Linear Regression analysis. In the first one, the predictor variable will be the total number of confirmed cases and the independent variable will be the number of fully vaccinated people. This model is to investigate if there is any association between the number of fully vaccinated people and the total confirmed cases in El Paso County. In the second one, the predictor variable will be the number of daily confirmed cases and the independent variable will be the number of fully vaccinated people. This model is to investigate if there is any association between the number of fully vaccinated people and daily confirmed cases in El Paso County. The third analysis is with total Covid-19 death rate as a predictor variable and the total number of fully vaccinated people as an independent variable to investigate if there is any association between the Covid-19 death rate and the daily increase of fully vaccinated people.

The dataset does not contain any patient’s personal information but only numbers of cumulative confirmed cases, covid death and the number of vaccinated people. Also, we can assume that when people made the vaccination reservation, they have signed the information consent the data will be used for scientific research. Thus, we think OLS regression is a reasonable method to conduct this research.

**Findings**

**Findings From EDA**

Firstly, we would like to see if the mask mandatory policy helps reduce the total as well as daily confirmed Covid cases via a line chart.

Chart, histogram

Description automatically generated

The red color represents that mask mandate order was issued and blue line depict that the mask mandate order was not issued or was lifted. The dash line represents the daily increase rate of the confirmed cases, and the solid line shows the cumulative total confirmed cases. From the chart, we can clearly see that the mask mandate brought down confirmed positive cases over a shorter duration of time and reduced daily increase rate in El Paso, Colorado.

From the result above, we can say that mask mandate might be helpful to bring down the daily increase rate, but we also think the development and administration of Covid 19 vaccination can significantly contribute to this effort. Thus, in the following analysis, we would like to further investigate the impact of Covid 19 vaccine administration on the spread of the disease.

Next, we will see how vaccination contributes to prevent the spread of the disease.   
Chart, histogram

Description automatically generated

In the chart above, the blue color represents the daily confirmed cases and the green color represent the number of daily fully vaccinated people. Before vaccinations were available, for example from Nov.2020 to Jan.2021, increased daily confirmed cases led to higher daily covid death. We also noticed that the number of daily increased fully vaccinated increased significantly in the beginning of Aug.2021. In details, the percentage of people who got fully vaccinated jumped from 15.5% to 47.1% in two days. Considering the number of facilities and number of nurses who would help with administrating the vaccination wouldn't be changed so tremendously, we think this data looks problematic. But before making any decision, we would like to see how the graph look like when excluding these two days' data.

Chart, histogram

Description automatically generated

The above figure is to show if vaccination contributes to prevent the spread of the disease by removing the data from 2021-08-01 to 2021-08-02. From the graph, we can see that between Apr.2021 to Aug.2021, with more vaccinations, daily confirmed cases decreased. However, after Sep.2021, as the number of daily vaccinated people decreased, the daily confirmed cases increased.

Since couldn't find any evidence that can explain why the number of fully vaccinated people got increased by 213% just in two days, and including these data made it difficult to tell if there is any association between the number of daily confirmed cases and the number of daily fully vaccinated. We will exclude these two days moving forward.

Next, we will check if there is any potential association between daily vaccinated and daily death.

Chart, histogram

Description automatically generated

The red solid line represents the daily death and the green dash line represent the number of daily vaccinated people. From the chart, we can see that before the vaccination was available, from example from Nov.2020 to Jan.2021, increased daily confirmed cases led to higher daily covid death. However, after the vaccination began, these is no clear association between daily confirmed cases and daily covid death.

Therefore, from the above, we can see that there might be an association between daily vaccinated and daily confirmed cases whereas there may not be any association between daily vaccinated and daily death.

**Findings From Correlation Analysis**

Moving forward, we will apply OLS for hypothesis testing to see if there is any association between the number of total vaccinated and daily confirmed cases as well as the number of total vaccinated and total death. But before we start with the OLS model, we would like to see if any of the parameters have a strong correlation between each other.

Chart

Description automatically generated

From the correlation matrix, we can see that most parameters are not highly correlated. In particular, ‘Series\_Complete\_Yes’ means the total number of vaccinated people, ‘mean\_cases’ represents the 7-day rolling total confirmed cases and ‘death’ represents the total covid death. From the heatmap, we can see that there is a relatively strong correlation between the total number of vaccinated people with total confirmed cases (0.86), and daily confirmed cases (0.58) and total covid death (0.83).

**Findings From Hypothesis Testing**

Therefore, we will apply the OSL for hypothesis testing with significance level equals 0.05.

**Hypothesis Testing 1 Results & Interpretations**

The first hypothesis assumed that there is no correlation between the total fully vaccinated people and the number of total confirmed cases in El Paso County. Thus, we use the 7-day rolling total confirmed cases as the predictor variable and the total number of vaccinated people as independent variable for OLS model. The results show below:

A screenshot of a computer

Description automatically generated with low confidence

From the results, we reject the hypothesis 1 because p value is smaller than 0.05. Thus, we conclude there is an association between the total vaccinated people and the number of total confirmed cases in El Paso County. In detail, if 1 more people got vaccinated every day, then the number of total confirmed cases will be increased by 0.1926.

**Hypothesis Testing 2 Results & Interpretations**

The second hypothesis assumed that there is no correlation between the total fully vaccinated people and the number of daily newly confirmed cases in El Paso County. Thus, we use the daily confirmed cases as the predictor variable and the total number of vaccinated people as independent variable for OLS model. The results show below:

A screenshot of a computer

Description automatically generated with low confidence

From the results, we reject the hypothesis 2 because p value is smaller than 0.05. Thus, we conclude there is an association between the total vaccinated people and the number of daily confirmed cases in El Paso County. In detail, if 1 more people got vaccinated every day, then the number of daily confirmed cases will be increased by 0.0004. However, the R squared value is only 0.054, which means this model does not explain the data well. Thus, even though the results show statistically significant, we may not want to draw any conclusion at this moment.

**Hypothesis Testing 3 Results & Interpretations**:

The third hypothesis assumed that there is no correlation between the total fully vaccinated people and the number of total deaths in El Paso County. Thus, we use the total covid deaths as the predictor variable and the total number of vaccinated people as independent variable for OLS model. The results show below:

Table

Description automatically generated

From the results, we reject the hypothesis 3 because p value is smaller than 0.05. Thus, we conclude there is an association between the total vaccinated people and the number of total covid deaths in El Paso County. In detail, if 1 more people got vaccinated every day, then the number of total deaths will be increased by 0.0013.

**Discussion/Implications**

Current vaccines are effective at preventing severe disease and deaths from COVID-19, but our findings suggest that vaccination alone is not sufficient to prevent transmission of the disease. The results show that there is a potential positive association between the total number of vaccinated people and the total number of confirmed cases and the total deaths, but the results should not be taken too literally or used for prediction because the data is skewed, and the model assumptions were not well met.

The R squared value, especially for hypothesis testing 2 and 3 were very low, which indicates that the model used in the analysis cannot explain the data well. Therefore, for future research, we should also explore other hypothesis testing methods to better fit assumptions for looking at the control of Covid-19 epidemic around the world.

**Limitations**

There are four main limitations for this research. First, the model assumptions are not well met due to the skewed data so the results of the model can be used to understand basic associations but should not be taken too literally or used for prediction. Other model should be applied before drawing any conclusions.

Second, the associations found in this analysis may not translate to other counties, states, and countries.

Third, this analysis does not completely capture the associations the completely vaccinated and confirmed cases and total covid death.

Finally, other influencing factors such as recovery data and more vaccination data could be considered for further analysis. The lack of vaccination data from Dec.2020 to May 24th, 2021 may not very well explain the impact of vaccine administration on reducing the spread of Covid-19 regarding other factors such as stay-at-home order, mask mandate policy, etc.

**Conclusion**

This analysis investigated the following questions:

1. How did mask mandate policy help to reduce the spread of the disease?
2. How did Covid-19 vaccine administration help reduce the spread of the disease?
3. Did vaccination help with the decrease in daily Covid death?

In conclusion, this analysis looked at vaccination effectiveness on reducing the spread of corona virus at a county level and found that there is a possible positive association between the number of fully vaccinated people and confirmed cases and with daily confirmed cases and total covid deaths.

**References**

Eyre, D. W., Taylor, D., Purver, M., Chapman, D., Fowler, T., Pouwels, K. B., Walker, A. S., & Peto, T. E. A. (2021, January 1). *The impact of SARS-COV-2 vaccination on Alpha & Delta Variant Transmission*. medRxiv. Retrieved December 13, 2021, from https://www.medrxiv.org/content/10.1101/2021.09.28.21264260v1.full.

Elliott, P., Haw, D., Wang, H., Eales, O., Walters, C. E., Ainslie, K. E. C., Atchison, C., Fronterre, C., Diggle, P. J., Page, A. J., Trotter, A. J., Prosolek, S. J., The COVID-19 Genomics UK (COG-UK) Consortium, Ashby, D., Donnelly, C. A., Barclay, W., Taylor, G., Cooke, G., Ward, H., … Riley, S. (2021, January 1). *React-1 round 13 final report: Exponential Growth, high prevalence of SARS-COV-2 and vaccine effectiveness associated with Delta variant in England during May to July 2021*. medRxiv. Retrieved December 13, 2021, from https://www.medrxiv.org/content/10.1101/2021.09.02.21262979v1.

Mallapaty, S. (2021, February 19). *Can covid vaccines stop transmission? scientists race to find answers*. Nature News. Retrieved December 13, 2021, from https://www.nature.com/articles/d41586-021-00450-z.

**Data Sources**

* [COVID-19 Vaccinations in the United States and per County](https://data.cdc.gov/Vaccinations/COVID-19-Vaccinations-in-the-United-States-County/8xkx-amqh)
* [COVID-19 data from John Hopkins University](https://www.kaggle.com/antgoldbloom/covid19-data-from-john-hopkins-university?select=CONVENIENT_us_deaths.csv)
* [U.S. State and Territorial Public Mask Mandates](https://data.cdc.gov/Policy-Surveillance/U-S-State-and-Territorial-Public-Mask-Mandates-Fro/62d6-pm5i)