Effectiveness of Covid-19 Vaccination [8]

Rajan Chaturvedi
Courant Institute of Mathematical Sciences
New York University
New York, NY, USA
rsc9044@nyu.edu

ABSTRACT

Vaccination is the most important key element against the fight with Covid-19 pandemic. This is logical assumption given that the natural herd immunity within the people takes many years to develop and till then many lives will be lost. Vaccination against the Covid-19 was started from December 14, 2020[1], and it is continuing till now. Post development of vaccines, the question of effectiveness and longevity of efficacy of the current vaccines against the new variants of the Coronavirus has arisen. By analysis and visualization of daily new cases of Covid-19 across the USA and Covid-19 Vaccination trend, I have evidence that initial vaccination does had success in curbing the new cases but after a year, new case soared to a new high which indicates that efficacy of vaccines is fading readily, and supplementary shots of vaccines (Booster shots) may be necessary for now but may not be enough. Hence, there is a need to develop more improved and promising vaccines, which will be effective against the existing as well as new variants of coronavirus in long term. Without such an improving, it will be much harder to eliminate the Covid-19 pandemic.

INTRODUCTION

Coronavirus disease 2019 (COVID-19) is defined as illness caused by a novel coronavirus called severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2; formerly called 2019-nCoV), which was first identified amid an outbreak of respiratory illness cases in Wuhan City, Hubei Province, China. It was initially reported to the WHO on December 31, 2019. On January 30, 2020, the WHO declared the COVID-19 outbreak a global health emergency. On March 11, 2020, the WHO declared COVID-19 a global pandemic, its first such designation since declaring H1N1 influenza a pandemic in 2009.[2] In USA, there has been more than 80 million Covid positive cases and more than 1 million American have lost their lives due to Covid-19.[3] US Vaccination effort has been going on since more than a year and focus has been on the stubborn persistence of the pandemic which is fueled by new variants and incomplete or partial vaccination.

The initial vaccination saw success in preventing the spread of the disease but between November 27, 2021, and January 12, 2022, a total of 886,774 eligible persons infected with the omicron variant, 204,154 eligible persons infected with the delta variant, and 1,572,621 eligible test-negative controls were identified. At all-time points investigated and for all combinations of primary course and booster vaccines, vaccine effectiveness against symptomatic disease was higher for the delta variant than for the omicron variant. No effect against the omicron variant was noted from 20 weeks after two ChAdOx1 nCoV-19 doses, whereas vaccine effectiveness after two BNT162b2 doses was 65.5% (95% confidence interval [CI], 63.9 to 67.0) at 2 to 4 weeks, dropping to 8.8% (95% CI, 7.0 to 10.5) at 25 or more weeks. Among ChAdOx1 nCoV-19 primary course recipients, vaccine effectiveness increased to 62.4% (95% CI, 61.8 to 63.0) at 2 to 4 weeks after a BNT162b2 booster before decreasing to 39.6% (95% CI, 38.0 to 41.1) at 10 or more weeks. Among BNT162b2 primary course recipients, vaccine effectiveness increased to 67.2% (95% CI, 66.5 to 67.8) at 2 to 4 weeks after a BNT162b2 booster before declining to 45.7% (95% CI, 44.7 to 46.7) at 10 or more weeks. Vaccine effectiveness after a ChAdOx1 nCoV-19 primary course increased to 70.1% (95% CI, 69.5 to 70.7) at 2 to 4 weeks after an mRNA-1273 booster and decreased to 60.9% (95% CI, 59.7 to 62.1) at 5 to 9 weeks. After a BNT162b2 primary course, the mRNA-1273 booster increased vaccine effectiveness to 73.9% (95% CI, 73.1 to 74.6) at 2 to 4 weeks; vaccine effectiveness fell to 64.4% (95% CI, 62.6 to 66.1) at 5 to 9 weeks.[4]

It can be concluded that the primary immunization with two doses of ChAdOx1 nCoV-19 or BNT162b2 vaccine provided limited protection against symptomatic disease caused by the omicron variant. A BNT162b2 or mRNA-1273 booster after either the ChAdOx1 nCoV-19 or BNT162b2 primary course substantially increased protection, but that protection waned over time.[4]

In this study, we will analyze effectiveness of current vaccines against the new as well as the existing variants of the coronavirus. The study is a walk-through of the methods, data and techniques applied to analyze the past and current trends and final conclusions that can be drawn from these trends. The results of this study will help the citizens as well as the governments to decide whether more effective vaccine with long term protection is required or not.

METHODOLOGY

Data Sources:

To perform the analysis, I used two major data sources. One to find state-wise daily new cases of Covid-19 and other to find the state-wise vaccination progress.

- Daily New-Cases of Covid-19 Dataset: This dataset is managed by the CDC (Centers for Disease Control and Prevention). It contains state-wise daily new cases, overall cases, deaths, probable cases, and probable deaths etc. of each date. It contains data from December 2019 to till date.[5]
- Vaccination Dataset: This dataset is managed by *Our World in Data publication* of *Global Change data lab*. This dataset also contains state-wise vaccination data of each date. The data consists of daily total vaccination, count of fully vaccinated and partially vaccinated people, and total vaccination count till date etc. [6]

Study Design:

To study this data, I used Python language on Jupyter notebook. Following are the steps taken in overall study of data.

Data Wrangling: First the raw data was imported into the Jupyter notebook using Pandas library. Both datasets (Daily New Cases and Vaccination) are stored into separate dataframes. Then from both dataframes I removed unnecessary columns which were not required for the study. Then I removed null values using dropna() function in both dataframes. The date column's datatype was not consistent in both dataframes. I fixed the data type of both dataframes by using Pandas datetime function (pd.to_datetime). Also, in Daily new cases dataframe format of the new cases count was not consistent. I fixed that by converting the new_case column's data type to integer. After this, both dataframes were sorted by dates. Lastly for predictive modeling and graphical analysis, I prepared a new dataframe by combining both dataframes. The new dataframes consists of the information of overall daily new cases and overall vaccination count of USA for the duration between January 15, 2021, to April 15, 2022.

Data Visualization: The datasets of the new cases as well as of the Vaccination was state-wise. Visualization on the map of USA made more sense to me. Because in such a visualization we can keep track of trends, not only country-wise but also state-wise. To achieve this, I used Geopandas library of the Python language. I used the shape file of the USA.[7] To visualize the new cases first, I segmented the data into months, and I took the average number of new cases of each month of each state and visualized it on the map of USA in orange-red color. Similarly, for Vaccinations I took the overall vaccination count of each state by the end of each month and visualized the same on the Map of USA in blue color. In both visualizations, increase in color density represents increase in the count or vice-versa.

Predictive Modeling: To grasp the relationship between daily new cases count and overall vaccination count, I used linear regression. I applied the linear regression in the Python by using *statsmodels.api* library and produced the summary of the results by using *stats* function from *scipy* library. Also using *scipy* library I calculated the *Pearson correlation coefficient* between the new cases count and total vaccination count.

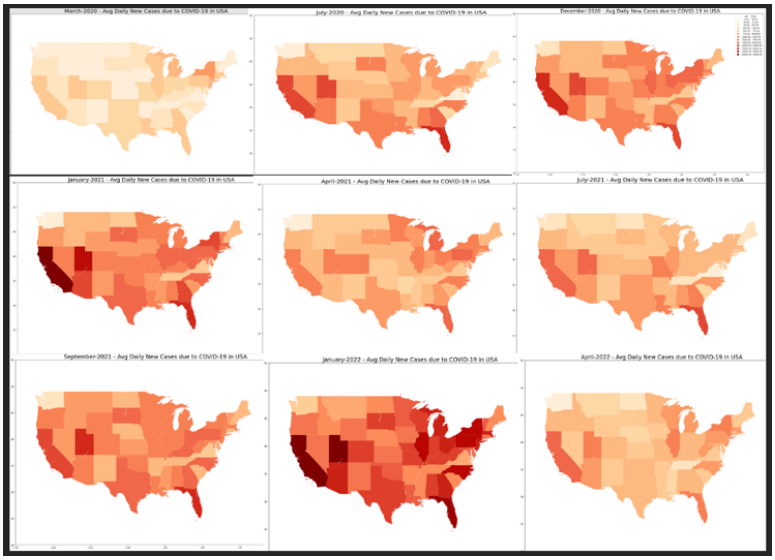
Graphical Analysis: To perform the graphical analysis between daily new cases count and total vaccination count, I plotted two graphs: scatterplot using *seaborn python library* and other is line graph using *Matplotlib* library in Python.

RESULTS

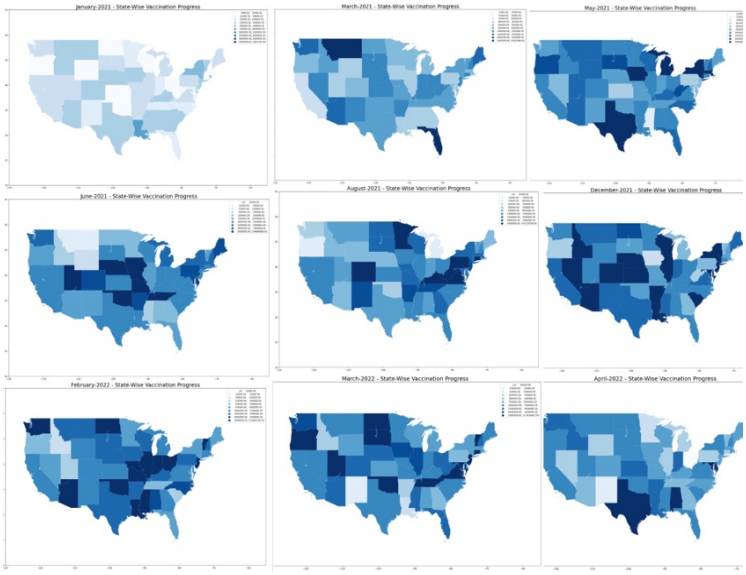
Data Visualization:

Following is the output I received in the data visualization step.

- Average of new covid-19 cases of the specific months.



- Vaccination trend throughout the USA in the specific months.



Predictive Modeling:

- Following is the output summary of linear regression of New covid-19 cases count vs Fully vaccinated people count.

OLS Regression Results

Dep. Variable:	new_case	R-squared:	0.040
Model:	OLS	Adj. R-squared:	0.038
Method:	Least Squares	F-statistic:	16.90
Date:	Mon, 02 May 2022	Prob (F-statistic):	4.78e-05
Time:	13:50:02	Log-Likelihood:	-5368.7
No. Observations:	404	AIC:	1.074e+04
Df Residuals:	402	BIC:	1.075e+04
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	4.608e+04	1.67e+04	2.757	0.006	1.32e+04	7.89e+04
people_fully_vaccinated	0.0004	0.000	4.111	0.000	0.000	0.001

Omnibus:	356.743	Durbin-Watson:	0.226
Prob(Omnibus):	0.000	Jarque-Bera (JB):	6946.477
Skew:	3.844	Prob(JB):	0.00
Kurtosis:	21.803	Cond. No.	3.76e+08

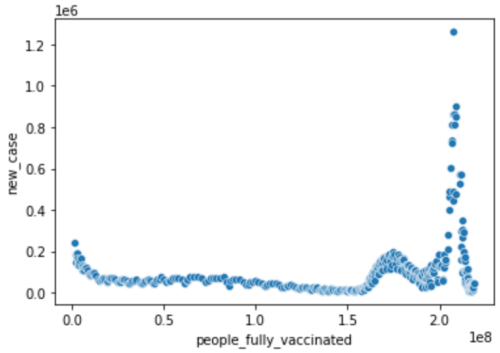
- Pearson correlation coefficient - New covid-19 cases count vs Fully vaccinated people count:

```
0.20085284013359417 4.780774577476342e-05
```

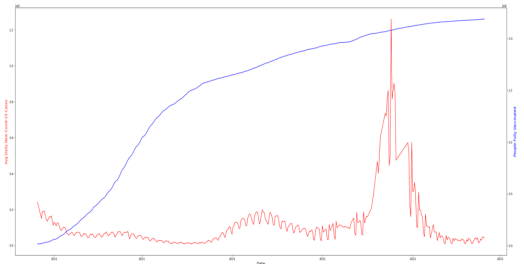
Graphical Analysis:

Graphical plot of New covid-19 cases counts vs Fully vaccinated people counts:

- Scatterplot



- Line Graph



DISCUSSION

Analysis of two map visualizations indicated that initially when the vaccination program began in January 2021, average daily new cases count began to go down. This is evidenced by reduction in color density of new cases map as the color density of vaccination map increased. But this color density reduction trend of new cases reversed from the month of September 2021, and it can be observed that color density of new cases map reached a darker density in the month of January 2022 even though vaccination trend kept on increasing in its color density. But again, in the months March and April 2022, fading color density can be seen in new cases of covid-19. The change in the trend can be attributed to the administration of the Booster dose of the vaccines.

In linear regression between the daily new Covid-19 cases and fully vaccinated people, R-squared value turned out to be 0.040. Also, the Pearson correlation coefficient is 0.20.

In the graphical analysis between the daily new Covid-19 cases and fully vaccinated people we plotted one scatter plot and other the line plot. Scatter plot indicated a larger deviation from the best fit and most of the line equation has the 0 slope. In the line plot it can be clearly observed that though with the vaccination, new cases went down initially but it reached very higher value during the last quarter of 2021 and starting of 2022. After that February 2022, we can again see the downward trend in the new cases.

CONCLUSION

The analysis of the visualizations, graphs and relationship statistics indicates that the vaccines are not successfully able to bring down over all new cases count in long run. It is evidenced by spike in new cases after September 2021 reaching an all-time high in January 2022. The downward trend in the new cases post February 2022 can be attributed to administration of the Booster doses. But as current vaccine has showcased its short-lived effect in 2021, it seems Booster doses are not the permanent solution to stop this pandemic. A weak relationship between the new covid-19 cases and current vaccination program is also indicated by the low R-Square value (0.040) and Pearson correlation coefficient (0.20). Hence, current vaccines are required to be improved such that they provide long term protection against the existing as well as the new variants of the coronavirus.

REFERENCES

[1] COVID-19 vaccination in the United States; https://en.wikipedia.org/wiki/COVID-19_vaccination_in_the_United_States

[2] Coronavirus Disease 2019 (COVID-19); https://emedicine.medscape.com/article/2500114-overview#:~:text=Coronavirus%20disease%202019%20(COVID%2D19)%20is%20defined%20as%20illness,City%2C%20Hubei%20Province%2C%20China.

[3] Worldometers; https://www.worldometers.info/coronavirus/country/us/

[4] Covid-19 Vaccine Effectiveness against the Omicron (B.1.1.529) Variant; https://www.nejm.org/doi/full/10.1056/NEJMoa2119451

[5] CDC; https://data.cdc.gov/Case-Surveillance/United-States-COVID-19-Cases-and-Deaths-by-State-o/9mfq-cb36

[6] Our World in Data; https://ourworldindata.org/us-states-vaccinations

[7] United States Census Bureau; https://www.census.gov/geographies/mapping-files/time-series/geo/carto-boundary-file.html
[8] Fall-2021-RBDA-Project; https://github.com/rajan0112/Fall-2021-RBDA-Project